ISSN 1818-0442
DOI: 10.21293/1818-0442



ДОКЛАДЫ

Томского государственного университета систем управления и радиоэлектроники

2026 • Том 29, № 2

Доклады ТУСУР. 2026 • Том 29, № 2

Министерство науки и высшего образования Российской Федерации

ДОКЛАДЫ
ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ

2026, том 29, № 2

Периодический научный журнал

Выходит 4 раза в год

Основан в 1997 г.

ISSN 1818-0442

DOI: 10.21293/1818-0442

Редакционная коллегия

В.М. Рулевский, д.т.н., профессор, ректор Томского государственного университета систем управления и радиоэлектроники (ТУСУР), научный руководитель направления НИИ АЭМ ТУСУРа, Томск (***гл. редактор***).

А.А. Шелупанов, чл.-корр. РАН, д.т.н., профессор, президент ТУСУРа, заслуженный работник высшей школы РФ, дважды лауреат премии Правительства РФ в области образования, трижды лауреат премии Правительства РФ в области науки и техники, Томск (***зам. гл. редактора***).

С.П. Куксенко, д.т.н., профессор, проф. каф. телевидения и управления, проректор по НРиИ ТУСУРа, Томск (***зам. гл. редактора***).

Б.А. Беляев, д.т.н., профессор, зав. лабораторией ЭиСВЧЭ Института физики им. Л.В. Киренского СО РАН, заслуженный изобретатель России, Красноярск.

В.А. Богуш, д.ф.-м.н., профессор, ректор Белорусского государственного университета информатики и радиоэлектроники, Минск.

С.А. Гаврилов, д.т.н., профессор, ректор НИУ «Московский институт электронной техники», лауреат Премии Правительства РФ в области образования, Москва.

Т.Р. Газизов, д.т.н., профессор, зав. каф. телевидения и управления ТУСУРа, Томск.

Е.Б. Грибанова, д.т.н., доцент, профессор каф. автоматизированных систем управления ТУСУРа, Томск.

А.М. Заболоцкий, д.т.н., профессор, директор института радиоэлектронной техники ТУСУРа, Томск.

Д.П. Зегжда, чл.-корр. РАН, д.т.н., профессор, директор института кибербезопасности и защиты информации Санкт-Петербургского политехнического университета, Санкт-Петербург.

А.М. Кориков, академик международной академии наук, профессор, д.т.н., проф. каф. автоматизированных систем управления ТУСУРа, заслуженный деятель науки РФ, почетный работник науки и техники РФ, почетный работник высшего профессионального образования РФ, Томск.

Ю.Н. Кульчин, академик РАН, д.ф.-м.н., вице-президент РАН, научный руководитель Института автоматики и процессов управления Дальневосточного отделения РАН, Владивосток.

П.С. Ложников, д.т.н., профессор, проректор по НРиИД Омского государственного технического университета (ОмГТУ), зав. каф. комплексной защиты информации, гл.н.с. научно-исследовательской лаборатории информационной безопасности ОмГТУ, Омск.

Н.Д. Малютин, д.т.н., профессор, проф. каф. конструирования узлов и деталей радиоаппаратуры ТУСУРа, гл.н.с. НИИ систем электрической связи ТУСУРа, Томск.

А.С. Марков, д.т.н., с.н.с., президент ЗАО «НПО «Эшелон», зав. каф. НИЯУ МИФИ, проф. каф. МГТУ им. Н.Э. Баумана, лауреат Премии Правительства РФ в области науки и техники, почетный работник науки и высоких технологий РФ, Москва.

В.Ш. Меликян, чл.-корр. НАН Республики Армения, д.т.н., профессор, директор учебного департамента ЗАО «Синопсис Армения», заслуженный деятель науки Республики Армения, Ереван.

А.А. Мицель, д.т.н., профессор, проф. каф. автоматизированных систем управления ТУСУРа, Томск.

Д.Ю. Муромцев, д.т.н., профессор, проректор по научно-инновационной деятельности Тамбовского государственного технического университета, Тамбов.

С.Д. Одинцов, д.ф.-м.н., профессор, иностранный член Норвежской академии наук, проф.-исследователь Института космических наук (ICREA), Барселона.

А.В. Осипов, д.т.н., доцент каф. промышленной электроники ТУСУРа, Томск.

Н.А. Ратахин, академик РАН, д.ф.-м.н., советник директора, гл.н.с. Института сильноточной электроники СО РАН, Томск.

В.К. Сарьян, академик НАН Республики Армения, д.т.н., профессор, проф. каф. радио и информационных технологий факультета радиотехники и кибернетики МФТИ, заслуженный работник связи РФ, лауреат Государственной премии РФ в области науки и техники, лауреат Премии Правительства РФ в области науки и техники, Москва.

А.Р. Сафин, д.ф.-м.н., профессор РАН, рук. лаб. антиферромагнитной спинтроники, профессор, зам. директора по научной работе ФГБУН Института радиотехники и электроники им. В.А. Котельникова РАН; проф. каф. формирования и обработки радиосигналов НИУ «МЭИ», Москва.

И.А. Ходашинский, д.т.н., профессор, проф. каф. компьютерных систем в управлении и проектировании, зав. лаб. интеллектуальных систем (ЛИС) каф. КСУП ТУСУРа, Томск.

В.В. Шайдуров, академик РАН, д.ф.-м.н., профессор, директор Института вычислительного моделирования СО РАН Федерального исследовательского центра «Красноярский научный центр Сибирского отделения Российской академии наук» (ФИЦ КНЦ СО РАН), зав. каф. вычислительных и информационных технологий Сибирского федерального университета, Красноярск.

С.М. Шандаров, д.ф.-м.н., профессор., проф. каф. электронных приборов ТУСУРа, заслуженный работник высшей школы РФ, член Оптического общества Америки (OSA), член Международного НТО IEEE/LEOS инженеров по электротехнике и электронике, действительный член Оптического общества им. Д.С. Рождественского, Томск.

Ю.А. Шурыгин, д.т.н., профессор, советник при ректорате по комплексным вопросам функционирования университета, зав. каф. компьютерных систем в управлении и проектировании ТУСУРа, научный руководитель НИИ АЭМ ТУСУРа, заслуженный деятель науки РФ, лауреат Премии Правительства РФ в области образования, Томск.

Адрес издателя, редакции: 634050, г. Томск, пр. Ленина, 40. МК, каб. 310/2, тел. +7 (382-2) 70-15-82, 51-21-21. Эл. почта: journal@tusur.ru

Свидетельство о регистрации СМИ выдано Федеральной службой по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия: ПИ № ФС 77-19130

Учредитель: Томский государственный университет систем управления и радиоэлектроники

Подписной индекс 20648 в каталоге агентства «Урал-Пресс»: газеты и журналы.

Верстка, техническое редактирование, подготовка оригинал-макета В.М. Бочкаревой. Корректор В.Г. Лихачева.

Адрес типографии: ООО «Дельтаплан». 634041, г. Томск, ул. Тверская, 81, т. 435-400, 434-600

Подписано в печать 12.10.2026. Выход в свет 15.10.2026. Формат 60×84 1/8. Печ. л. 12,5. Тираж 500. Заказ 12. Цена 900 руб.

Editorial board

Viktor M. Rulevskiy — Dr. Sc. Eng., Professor, Rector of Tomsk State University of Control Systems and Radioelectronics (TUSUR), Scientific Director of the Research Institute of AEM TUSUR, Tomsk (***Editor-in-Chief***).

Alexander A. Shelupanov — Corr. Member of the RAS, Dr. Sc. Eng., Professor, President of TUSUR, Honored Worker of Higher School of the Russian Federation (RF), Honorary Worker of Science and Technology of RF, Laureate of RF Government Prize in Education, Three times Laureate of RF Government Prize in Science and Technology, Tomsk (***Deputy Editor in Chief***).

Sergey P. Kuksenko — Dr. Sc. Eng., Professor, Prof., Dept. of Television and Control, TUSUR Vice-Rector for Research and Innovation of TUSUR, Tomsk (***Deputy Editor-in-Chief***).

Boris A. Belyaev — Dr. Sc. Eng., Professor, Head of the Electromagnetic Computing and Microwave Electronics Laboratory, L.V. Kirensky Institute of Physics SB RAS, Honored Inventor of Russia, Krasnoyarsk.

Vadim A. Bogush — Dr. Phys.-Math. Sci., Professor, Rector of the Belarusian State University of Informatics and Radioelectronics, Minsk.

Sergei A. Gavrilov — Dr. Sc. Eng., Professor, Rector of the National Research University «Moscow Institute of Electronic Technology», Laureate of RF Government Prize in Education, Moscow.

Talgat R. Gazizov — Dr. Sc. Eng., Professor, Head of the Dept. of Television and Control, TUSUR, Tomsk.

Ekaterina B. Gribanova — Dr. Sc. Eng., Associate Professor, Prof., Dept. of Automated Control Systems, TUSUR, Tomsk.

Alexander M. Zabolotsky — Dr. Sc. Eng., Professor, Director of the Institute of Radioelectronic Engineering, TUSUR, Tomsk.

Dmitry P. Zegzhda — Corr. Member of the RAS, Dr. Sc. Eng., Professor, Director of the Institute of Cybersecurity and Information Protection, St. Petersburg Polytechnic University, St. Petersburg.

Anatoly M. Korikov — Academician of the International Academy of Sciences, Dr. Sc. Eng., Professor, Prof., Dept. of Automated Control Systems, TUSUR, Honored Scientist of RF, Honorary Worker of Science and Technology of RF, Honorary Worker of Higher Professional Education of RF, Tomsk.

Yury N. Kulchin — Academician of the RAS, Dr. Phys.-Math. Sci., Vice-President of the RAS, Scientific Director of the Institute of Automation and Control Processes FEB RAS, Vladivostok.

Pavel S. Lozhnikov — Dr. Sc. Eng., Professor, Vice-Rector for Research and Innovation of Omsk State Technical University (OmSTU), Head of Department «Complex Information Security», OmSTU, Omsk.

Nikolay D. Malyutin — Dr. Sc. Eng., Professor, Chief Researcher, Research Institute of Electrical Communication Systems, Prof., Dept. of Design of Units and Components for Radioelectronic Systems, TUSUR, Tomsk.

Alexey S. Markov — Dr. Sc. Eng., Senior Researcher, President at NPO Echelon, Laureate of RF Government Prize in Science and Technology, Honorary Worker of Science and High Technologies of RF, Moscow.

Vazgen Sh. Melikyan — Corr. Member of the National Academy of Sciences of the Republic of Armenia, Dr. Sc. Eng., Professor, Director of the Educational Department of ZAO Synopsys Armenia, Honored Scientist of the Republic of Armenia, Yerevan.

Arthur A. Mitsel — Dr. Sc. Eng., Professor, Prof., Dept. of Automated Control Systems, Academician of the International Academy of Sciences, TUSUR, Tomsk.

Dmitry Yu. Muromtsev — Dr. Sc. Eng., Professor, Vice-Rector for Research, Tambov State Technical University, Tambov.

Sergey D. Odintsov — Dr. Phys.-Math. Sci., Professor, Foreign Member of the Royal Academy Norway and the European Physical Society, Research Professor, Institute of Space Sciences (ICREA), Barcelona, Spain.

Alexander V. Osipov — Dr. Sc. Eng., Associate Professor of the Department of Industrial Electronics TUSUR, Tomsk.

Nikolay A. Ratakhin — Academician of the RAS, Dr. Phys.-Math. Sci., Advisor to the Director, Chief Researcher, Institute of High Current Electronics, SB RAS, Tomsk.

Vilyam K. Saryan — Academician of the National Academy of Sciences of the Republic of Armenia, Dr. Sc. Eng., Professor, Prof., Dept. of Radio and Information Technologies, Faculty of Radio Engineering and Cybernetics, Moscow Institute of Physics and Technology (National Research University), Laureate of the State Prize of the Russian Federation in the field of Science and Technology, Laureate of RF Government Prizes in the field of Science and Technology, Moscow.

Ansar R. Safin — Dr. Phys.-Math. Sci., Professor of the RAS, Head of the Antiferromagnetic Spinelectronics Lab, Professor, Kotelnikov Institute of Radioengineering and Electronics of the RAS, Prof., Dept. of Formation and Processing of Radio Signals, National Research University "MPEI", Moscow.

Ilya A. Hodashinsky — Dr. Sc. Eng., Professor, Prof., Dept. of Computer Control and Design Systems, Head of the Intelligent Systems Laboratory, TUSUR, Tomsk.

Vladimir V. Shaidurov — Corr. Member of the RAS, Dr. Phys.-Math. Sci., Professor, Director, Institute of Computational Modeling SB RAS, Chief Researcher, Computational Mathematics Department, ICM SB RAS, Federal Research Center, Krasnoyarsk Scientific Center, SB RAS, Krasnoyarsk.

Stanislav M. Shandarov — Dr. Phys.-Math. Sci., Professor, Prof., Dept. of Electronic Devices, TUSUR, Honored Worker of Higher Education of the Russian Federation, Member of the Optical Society of America (OSA), Member of the IEEE/LEOS International Scientific and Technical Society of Electrical and Electronic Engineers, Full Member of the D.S. Rozhdestvensky Optical Society, Tomsk.

Yury A. Shurygin — Dr. Sc. Eng., Professor, Advisor to the Rector's Office on Complex Issues of University Functioning, TUSUR, Head of the Department of Computer Control and Design Systems, TUSUR, Scientific Director of the Research Institute of Automation and Electromechanics (RIAE

Содержание

ЭЛЕКТРОНИКА, РАДИОТЕХНИКА И СВЯЗЬ

Власов С.В., Жечев Е.С.
Методы улучшения добротности и снижения частотной дисперсии параметров интегральных катушек индуктивности: обзор и постановка исследовательских задач 7

Юзвик Д.А., Степанов М.А.
Метод формирования нескольких минимумов электромагнитного поля в требуемых точках пространства 16

Мурманский М.С., Томских А.Е., Жечев Е.С., Суровцев Р.С., Скорняков И.А.
Способы подавления перекрестных помех в связанных линиях передачи: обзор 24

Зотов Ф.А., Меркулов А.А., Цитович А.А.
Проектирование вторичного источника питания для АФАР Ka-диапазона 39

Чинь С.Б., Газизов Т.Р.
Влияние аппроксимации оптимальной токовой сеткой на массу и моностатическую эффективную площадь рассеяния параболической рефлекторной проводно-сеточной антенны 48

Скорик А.Д.
Использование мониторинга помех и неоднородностей ионосферы для повышения надежности коротковолновой связи 57

Кравченко А.Н., Покаместов Д.А.
Методы кодирования случайных кодов с малой плотностью проверок на четность 64

Безрукова И.В., Семенов Э.В.
Моделирование и особенности нелинейного отклика малоразмерных электронных устройств при интермодуляционном методе сверхширокополосного зондирования 74

УПРАВЛЕНИЕ, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАТИКА

Слядников Е.Е., Турчановский И.Ю.
Математическая модель неравновесных структурных переходов в аморфных металлических сплавах при термомеханических воздействиях 85

Богатов Н.М., Володин В.С., Григорьян Л.Р., Коваленко М.С.
Численное решение уравнения Пуассона в кремниевой *n–i–p*-наноструктуре с учетом поверхностных состояний 93

Сапунов А.О., Немирович-Данченко М.М.
Сравнительный анализ применения информационного моделирования в отечественной и зарубежной практике управления данными жизненного цикла объекта 99

Бабаев С.И., Костров Б.В., Муратов Е.Р., Тарасов А.С., Тарасова В.Ю.
Алгоритм уточнения совмещения разнородных изображений на основе ключевых точек 107

Коломиец В.В.
Автоматическое индексирование баз данных с использованием обучения с подкреплением 116

Захаров М.Е., Чернышов М.Н., Дмитриев В.М., Захаров К.Е., Ульев Л.М.
Методика нелинейного предиктивного управления технологическими комплексами нефтегазоподготовки 129

Свинцов М.М., Мышляев Л.П., Макаров Г.В., Грачев В.В.
Математические модели объектов управления с координатно-параметрическими связями 135

Чуков К.И., Фаерман В.А.
Программный комплекс для моделирования оптического тракта системы квантовой коммуникации 140

Штука С.Е., Боршевников А.Е.
Модификация нейросетевого алгоритма высоконадёжной биометрической аутентификации по радужной оболочке глаза 148

Бочков А.В., Сабанов А.Г., Шелупанов А.А.
Концепция формирования метрик безопасности прикладного ПО для объектов КИИ 160

ЭЛЕКТРОТЕХНИКА

Полынцев Е.С., Кодорова И.Ю., Бартенев А.И., Аксенов А.А., Воробьев М.С., Кагадей В.А.
Силовая GaN-микросборка на основе AlN-DBC-платы с металлизированными переходными отверстиями 179

Бурцев Ю.А.
Развитие методов расчета переходных процессов на основе матричной экспоненты для цепей с нелинейными резисторами 189

Требования 199

Contents

ELECTRONICS, RADIO ENGINEERING AND COMMUNICATIONS

Vlasov S.V., Zhechev Y.S.
Methods for improving the quality factor and reducing the frequency dispersion of integrated inductor parameters: a review and formulation of research problems 7
Iuzvik D.A., Stepanov M.A.
A method for forming several minima of the electromagnetic field at required points in space 16
Murmansky M.S., Tomskikh A.E., Zhechev Y.S., Surovtsev R.S., Skornyakov I.A.
Crosstalk suppression methods in coupled transmission lines: a review 24
Zotov F.A., Merkulov A.A., Tsitovich A.A.
Design of a secondary power supply for a Ka-band AESA 39
Trinh X.B., Gazizov T.R.
Effect of optimal current grid approximation on the mass and monostatic radar cross-section of a wire grid parabolic reflector antenna 48
Skorik A.D.
Using monitoring of interference and ionospheric irregularities to improve the reliability of high-frequency communication 57
Kravtchenko A.N., Pokamestov D.A.
Encoding methods for random low-density parity-check codes 64
Bezrukova I.V., Semyonov E.V.
Modeling and features of the nonlinear response of small electronic devices using the intermodulation method of ultra-wideband sensing 74

CONTROL, COMPUTER SCIENCE AND INFORMATICS

Slyadnikov E.E., Turchanovsky I.Yu.
Mathematical model of non-equilibrium structural transitions in amorphous metal alloys under thermomechanical loading 85
Bogatov N.M., Volodin V.S., Grigoryan L.R., Kovalenko M.S.
Numerical solution of the Poisson equation in a silicon *n–i–p*-nanostructure taking into account surface states 93
Sapunov A.O., Nemirovich-Danchenko M.M.
Comparative analysis of the information modeling application in domestic and international practice of data lifecycle management 99
Babaev S.I., Kostrov B.V., Muratov E.R., Tarasov A.S., Tarasova V.Yu.
Keypoint-based algorithm for registration of heterogeneous images 107
Kolomiets V.V.
Automatic database indexing using reinforcement learning 116
Zakharov M.E., Chernyshov M.N., Dmitriev V.M., Zakharov K.E., Uliev L.M.
Methodology for nonlinear predictive control of oil and gas treatment complexes 129
Svintsov M.M., Myshlyaev L.P., Makarov G.V., Grachev V.V.
Mathematical models of control objects with coordinate-parametric couplings 135
Chukov K.I., Fayerman V.A.
Software package for modeling the optical path of a quantum communication system 140
Shtuka S.E., Borshevnikov A.E.
Modification of a neural network algorithm for highly reliable iris-based biometric authentication 148
Bochkov A.V., Sabanov A.G., Shelupanov A.A.
A concept for developing security metrics for application software in critical information infrastructure facilities 160

ELECTRICAL ENGINEERING

Polyntsev E.S., Kodorova I.Y., Bartenev A.I., Aksenov A.A., Vorobyov M.S., Kagadey V.A.
GaN-power micromodule based on an AlN-DBC-substrate with metallized vias 179
Burtsev Y.A.
Development of matrix-exponential-based methods for transient analysis of circuits with nonlinear resistors 189

Manuscript requirements 199

ЭЛЕКТРОНИКА,
РАДИОТЕХНИКА И СВЯЗЬ

В разделе статьи по следующим специальностям (технические науки)

2.2.13. Радиотехника, в том числе системы и устройства телевидения.
2.2.14. Антенны, СВЧ-устройства и их технологии.
2.2.16. Радиолокация и радионавигация.

УДК 621.318.4
Шифр специальности ВАК: 2.2.14

С.В. Власов, Е.С. Жечев

Методы улучшения добротности и снижения частотной дисперсии параметров интегральных катушек индуктивности: обзор и постановка исследовательских задач

Актуальность. Работа представляет собой обзор современных методов повышения добротности и снижения частотной дисперсии параметров интегральных катушек индуктивности; паразитные ёмкостные связи и омические потери ограничивают добротность и рабочую полосу частот интегральных индуктивностей, применяемых в согласующих цепях, фильтрах и резонаторах СВЧ-диапазона. ***Цель исследования.*** Систематизация современных подходов к проектированию планарных и объёмных интегральных индуктивностей, направленных на повышение добротности и снижение частотной дисперсии параметров, а также формулирование исследовательских гипотез в области конструктивно-топологической оптимизации. ***Методы.*** Аналитический обзор конструктивных решений, включая нерегулярную ширину проводников, фрактальные геометрии, многослойные структуры, методы экранирования подложки и способ соединения. ***Научная новизна.*** Комплексное рассмотрение факторов, определяющих добротность и частотную дисперсию интегральных индуктивностей, в совокупности с конструктивно-топологическими аспектами, ранее рассматриваемыми преимущественно раздельно. ***Результаты.*** Показано, что оптимизация геометрии витков и применение многоуровневой металлизации позволяют повысить добротность, при этом влияние способов соединения и взаимной ориентации многослойных структур остаётся недостаточно изученным; сформулированы гипотезы о возможности улучшения частотных характеристик за счёт комплексной оптимизации геометрии и топологии интегральных индуктивностей. ***Практическая значимость.*** Результаты обзора могут быть использованы при проектировании согласующих цепей, фильтров и резонаторов СВЧ-диапазона на этапе выбора геометрии и топологии интегральных катушек индуктивности.
Ключевые слова: интегральная индуктивность, добротность, частотная дисперсия параметров, собственный резонанс, паразитная емкость, топология СВЧ-схем, фрактальные структуры.
DOI: 10.21293/1818-0442-2026-29-2-7-15

Введение

Первые систематические исследования планарных интегральных катушек индуктивности связаны с работой Нгуена и Мейера [1], где было показано, что спиральные проводники, совместимые с кремниевой технологией, могут заменить дискретные катушки в LC-фильтрах, несмотря на невысокую на тот момент добротность. В последующие десятилетия развитие технологий металлизации и появление структур на основе полуизолирующих подложек (в том числе GaAs) позволило существенно повысить добротность и расширить рабочий диапазон интегральных индуктивностей.

Стремительное развитие систем беспроводной связи, переход к стандартам 5G/6G и освоение миллиметрового диапазона длин волн диктуют жесткие требования к компонентной базе СВЧ-устройств. В условиях роста рабочих частот и необходимости расширения полос пропускания критически важными становятся вопросы обеспечения частотной дисперсии характеристик интегральных элементов [2–4]. В современной радиоэлектронике важной характеристикой индуктивных элементов (в контексте широкополосных приложений) является частотная дисперсия их параметров, а именно эффективной индуктивности $L(\omega)$ и добротности $Q(\omega)$, в пределах рабочей полосы частот; минимизация этой дисперсии обеспечивает стабильную работу устройства в заданном диапазоне. Смещение частоты собственного резонанса вследствие технологических разбросов, паразитных емкостей и эффектов близости приводит к рассогласованию импедансов, сужению рабочей полосы и деградации амплитудно-частотных характеристик устройств [5].

Рост сложности современных интегральных технологий, связанный с применением многоуровневой металлизации, уменьшением топологических норм и увеличением плотности интеграции, неизбежно усиливает влияние паразитных электромагнитных эффектов [6, 7]. Это особенно критично для таких узлов, как малошумящие усилители, генераторы, управляемые напряжением, и согласующие цепи, где интегральные катушки индуктивности являются одними из ключевых частотно-задающих компонентов [8, 9].

Существует фундаментальное противоречие: требования к миниатюризации пассивных элементов вступают в конфликт с необходимостью сохранения высоких значений добротности и индуктивности. Попытки уменьшения площади кристалла в рамках стандартных технологий часто приводят к снижению собственной резонансной частоты и росту омических потерь [10, 11]. Работоспособность СВЧ-устройств напрямую определяется тем, насколько эффективно удается компенсировать эти паразитные эффекты [12].

Интегральные катушки индуктивности характеризуются рядом физических ограничений, влияющих на их частотный отклик [13, 14]. Значение индуктивности определяется геометрией проводников, числом витков и занимаемой площадью. Однако увеличение

числа витков для роста L сопровождается удлинением проводника и повышением плотности намотки. Это, в свою очередь, приводит к росту паразитной межвитковой емкости и емкости на подложку [15], а также к влиянию скин-эффекта и эффекта близости.

В ряде работ [16–18] показано, что ключевые параметры (добротность, индуктивность) жестко взаимосвязаны. Например, в [19] приведен сравнительный анализ различных технологий изготовления, подтверждающий, что максимальное улучшение характеристик возможно только при комплексном подходе. В диссертационных исследованиях [20] отмечается, что паразитная емкость является основным фактором, ограничивающим частотный диапазон, а потери в полупроводниковой подложке и тонких слоях металлизации снижают добротность.

Важным аспектом является влияние технологического процесса и материалов. Так, замена алюминиевой металлизации на медную позволяет повысить добротность в 1,5–3,6 раза [21], что критично, например, для высокочастотных фильтров. Для полупроводниковых структур на основе GaAs дополнительным фактором выступает выбор материала металлизации: ввиду ограниченной совместимости медных процессов с арсенид-галлиевой технологией на практике чаще применяется золотая металлизация, обладающая высокой проводимостью и стойкостью к окислению. Существенное влияние на добротность и паразитную ёмкость также оказывают свойства подложки: полуизолирующий GaAs, в отличие от низкоомного кремния, снижает потери на вихревые токи и связанные с ними диэлектрические потери в подложке [22].

Для анализа и проектирования таких структур широко используется метод эквивалентных схем. Классическая П-образная модель позволяет учитывать распределенные параметры R, L, C и рассчитывать частотную зависимость эффективной индуктивности через входной импеданс $Z_{вх}$ согласно выражению (1) [23]:

$$L_э(\omega) = \mathrm{Im}(Z_{вх})\omega^{-1} = \mathrm{Im}\left(Z_0 \frac{1+S_{11}}{1-S_{11}}\right)\omega^{-1}. \quad (1)$$

Однако, несмотря на обширную литературу по оптимизации геометрии (формы витков, сечения проводников), вопросы влияния способов соединения (топологии соединения витков, межслойной коммутации) и их взаимной ориентации на частотную дисперсию параметров в интегральном исполнении изучены недостаточно полно. Большинство исследований фокусируется на достижении номинала индуктивности, оставляя вопросы широкополосности и частотной дисперсии параметров на втором плане.

Целью данной статьи является обзор и систематизация существующих конструктивно-топологических методов уменьшения частотной дисперсии параметров и добротности интегральных катушек индуктивности, а также формулирование гипотез и определение перспективных направлений исследований в области конструктивно-топологической оптимизации межсоединений. Для достижения поставленной цели решаются следующие задачи:

– анализ влияния геометрии витков (включая фрактальные и нерегулярные структуры) на добротность и частотный диапазон;

– обзор методов снижения паразитных емкостей в многослойных структурах;

– выявление взаимосвязи между способом соединения (коммутации) индуктивных элементов и их частотными характеристиками.

Методы оптимизации геометрии и топологии

Выбор топологии планарной катушки индуктивности является первым этапом проектирования, определяющим компромисс между занимаемой площадью кристалла, индуктивностью и добротностью. В современной микроэлектронике наиболее распространены спиральные структуры: квадратные, шестиугольные, восьмиугольные и круглые.

Сравнительный анализ базовых геометрий, проведенный в работе [24], показывает, что форма витков существенно влияет на омические потери и распределение электромагнитного поля. Как показано на рис. 1, восьмиугольная топология демонстрирует наименьшие потери и более высокую добротность по сравнению с квадратной и шестиугольной конфигурациями, приближаясь по характеристикам к идеальной круглой спирали, но будучи более технологичной при сеточном проектировании.

Дальнейшее развитие геометрической оптимизации связано с переходом к объемным структурам. В исследовании [25] рассмотрен широкий спектр топологий, включая 3D-соленоиды, конические (сужающиеся) спирали и стековые структуры. Установлено, что для малых номиналов индуктивности (порядка 2 нГн) наименьшей частотной дисперсией обладает сужающаяся 3D-П-образная спираль, тогда как для больших номиналов (16 нГн) предпочтительна классическая сужающаяся 3D-спираль. Это объясняется снижением межвитковой паразитной емкости за счет пространственного разнесения витков.

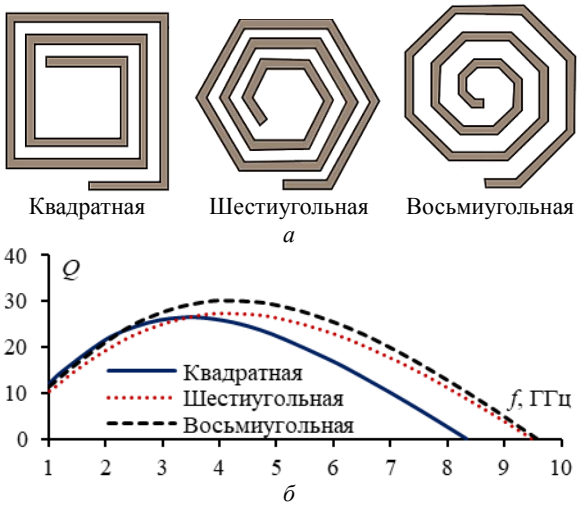


Рис. 1. Три типа планарных катушек индуктивности (*а*) и их характеристики (*б*) [24]

Одним из эффективных методов повышения добротности является использование проводников с переменной (нерегулярной) шириной. Физический смысл данного подхода заключается в подавлении эффекта вытеснения тока и скин-эффекта во внутренних витках спирали, где магнитное поле наиболее интенсивно.

В работе [26] предложена восьмиугольная спиральная катушка с монотонно уменьшающейся шириной проводника от внешних витков к внутренним (рис. 2). Экспериментальные данные подтверждают, что такая модификация позволяет увеличить максимальную добротность с 17,5 до 23,8 и повысить индуктивность в низкочастотной области с 6,4 до 8,3 нГн (для шестивитковой конфигурации).

Альтернативный подход к нарушению регулярности формы рассмотрен в [27], где анализируются одновитковые восьмиугольные индуктивности с асимметричными сторонами. Показано, что увеличение длины горизонтальных сегментов при сохранении вертикальных размеров способствует росту добротности на низких частотах (рис. 3). Однако уменьшение вертикального размера, как правило, ведет к деградации параметров, что требует тщательной оптимизации соотношения сторон.

Для интегральных преобразователей напряжения критическим параметром является не только добротность, но и активное сопротивление постоянному току, влияющее на КПД устройства. В [28] продемонстрировано, что модифицированная восьмиугольная форма, оптимизированная под площадь кристалла, позволяет снизить сопротивление с 1,56 до 1,48 Ом по сравнению с квадратным аналогом. Это обеспечивает прирост КПД преобразователя примерно на 5% (рис. 4).

Помимо геометрии проводников, существенный вклад в снижение добротности вносят вихревые токи, индуцируемые в полупроводниковой подложке. Для минимизации этого эффекта в работе [29] предложен метод структурирования подложки с использованием *n*-карманов в *p*-подложке (технология объемных КМОП).

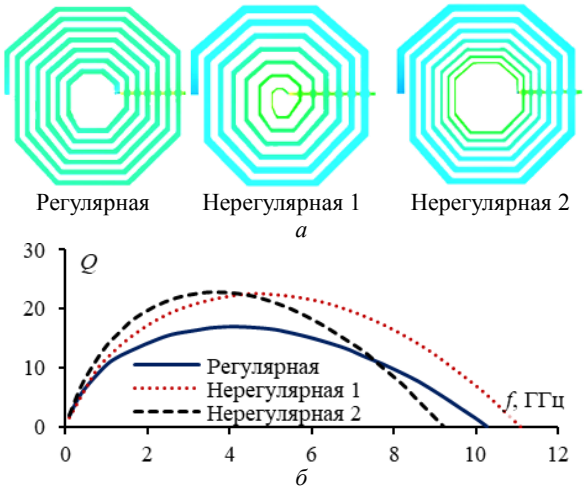


Рис. 2. Восьмиугольные катушка индуктивности с нерегулярной шириной проводников (*а*) и ее результаты (*б*) [26]

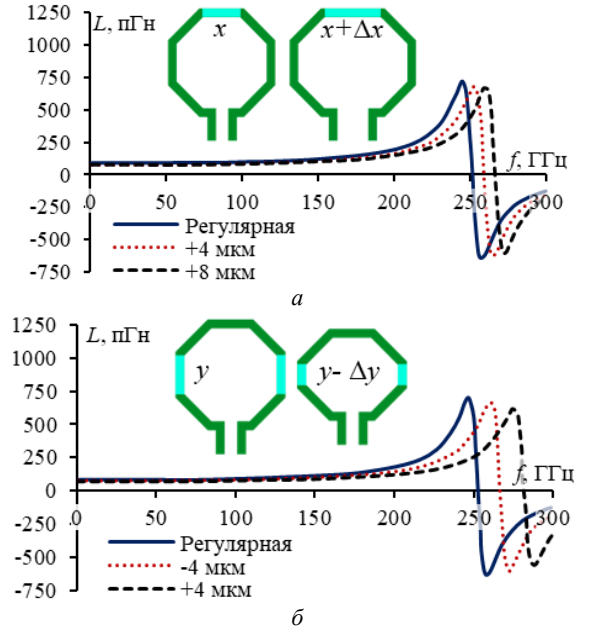


Рис. 3. График индуктивности для нерегулярной катушки индуктивности по горизонтали (*а*) и по вертикали (*б*) [27]

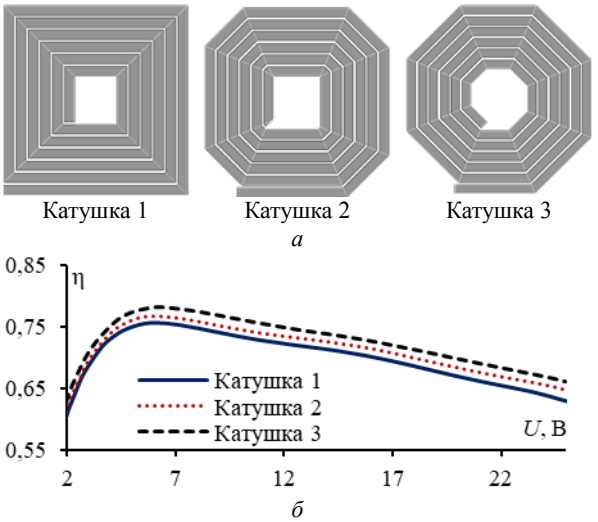


Рис. 4. Исследуемые катушки индуктивности различной формы (*а*) и их КПД (*б*) [27]

Сравнительный анализ двух топологий экранирования (рис. 5) показал следующие результаты:

– структура 1 (вертикальные полосы *n*-ячеек): обеспечивает рост добротности на 77% (с 3 до 5,5) и индуктивности на 8%;

– структура 2 (изолированные островки *n*-ячеек): дает прирост добротности на 16% и индуктивности на 4%.

Такой подход позволяет эффективно блокировать пути протекания вихревых токов без изменения топологии самой катушки.

Перспективным направлением миниатюризации при сохранении высоких значений индуктивности является использование фрактальных кривых. В работе [30] исследована планарная индуктивность на основе кривых Мура (рис. 6).

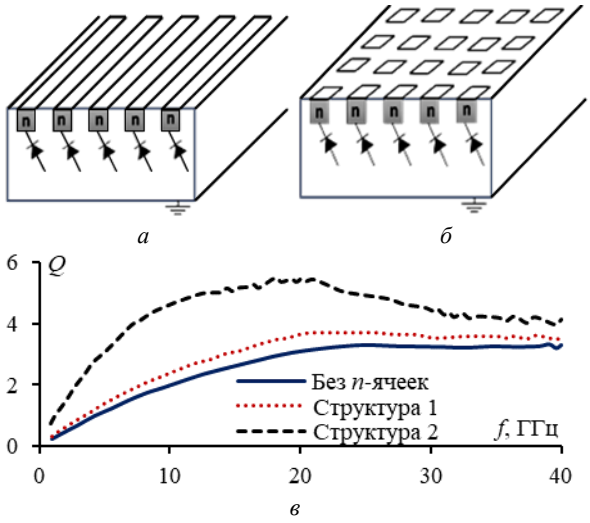


Рис. 5. Вид в разрезе структур 1 (*а*) и 2 (*б*) и их добротность (*в*) [29]

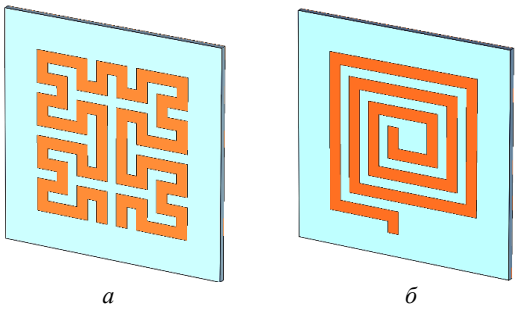


Рис. 6. Планарная катушка индуктивности с фрактальной (*а*) и спиральной структурой (*б*) [30]

Фрактальная геометрия позволяет существенно увеличить длину проводника на ограниченной площади, что обеспечивает высокую плотность индуктивности. Сравнение с традиционной спиральной катушкой (рис. 7) демонстрирует гибкость настройки индуктивности за счет изменения порядка фрактала (количества итераций узора).

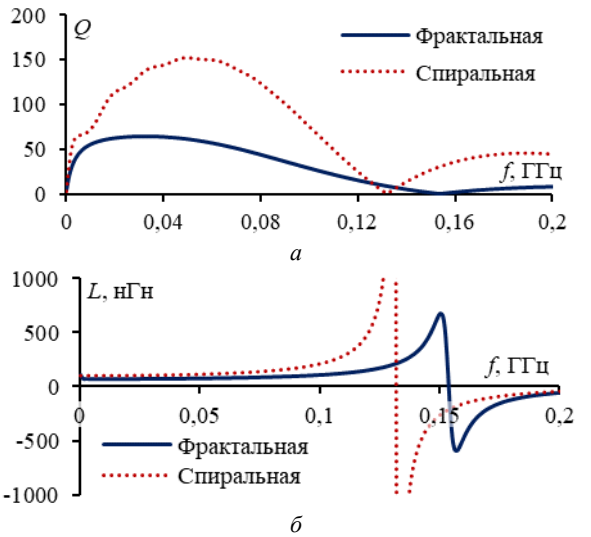


Рис. 7. Графики добротности (*а*) и индуктивности (*б*) для катушек индуктивности с фрактальной и спиральной топологиями

Однако следует отметить, что сложная форма проводника может приводить к росту паразитной межвитковой емкости и снижению частоты собственного резонанса, что требует дополнительного анализа при проектировании СВЧ-устройств.

Способы соединения и многослойные топологии

Анализ литературных источников показывает, что подавляющее большинство исследований фокусируется на оптимизации геометрии одиночных катушек. Однако в сложных СВЧ-устройствах (фильтры, согласующие цепи, дифференциальные усилители) индуктивности редко используются изолированно. Взаимное расположение и способ электрического соединения витков оказывают критическое влияние на результирующую индуктивность, добротность и, что не менее важно, на частотную дисперсию параметров [31–33].

Существует множество вариантов соединения индуктивных элементов, которые можно классифицировать по типу электрической связи (последовательное, параллельное) и магнитной связи (согласное, встречное включение). На рис. 8 представлены типовые способы соединения, включая варианты с сильной магнитной связью (трансформаторный тип).

Схемы 2 и 3 отличаются относительной ориентацией витков, однако при отсутствии магнитной связи между элементами ($M = 0$) это не влияет на результирующую индуктивность, которая в обоих случаях равна $L_1 + L_2$. Известно, что при наличии взаимной индукции M общая индуктивность системы L_{total} зависит от знака связи: $L_{\text{total}} = L_1 + L_2 \pm 2M$. Однако, как отмечается в [34–36], влияние способа соединения не ограничивается только изменением номинала L. Различные конфигурации по-разному перераспределяют потенциалы вдоль витков, что изменяет эффективную паразитную емкость C_{eff} и, следовательно, частоту собственного резонанса. Несмотря на значимость этого эффекта, системные исследования данного вопроса в контексте интегральных способов соединений катушек индуктивностей представлены фрагментарно [37].

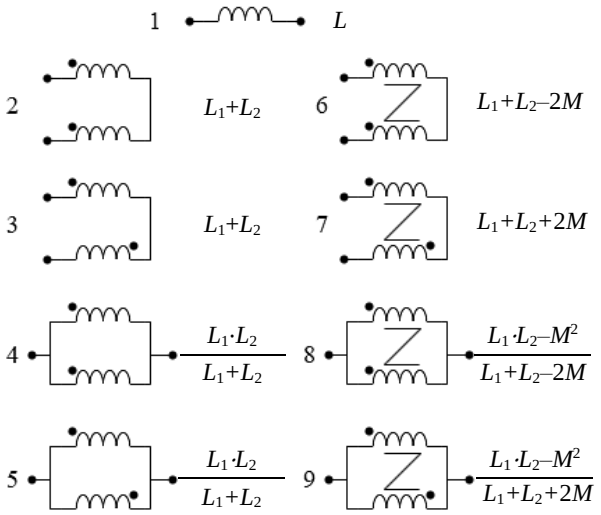


Рис. 8. Способы соединения катушек индуктивности

Управление взаимной индуктивностью открывает возможности для создания перестраиваемых компонентов. В работе [38] предложена топология на основе связанных линий, где коэффициент связи регулируется с помощью ключевых элементов (рис. 9). Такой подход позволяет изменять общую индуктивность без увеличения площади кристалла, реализуя функцию вариометра в планарном исполнении. Экспериментально подтверждено, что управление магнитной связью является эффективным методом коррекции частотной характеристики в широком диапазоне. Перестройка индуктивности реализована за счёт вторичного связанного контура, магнитная связь которого с основной катушкой управляется переключателем. В замкнутом состоянии переключателя (см. рис. 9, *а*, справа) во вторичном контуре наводится ток, изменяющий эффективную индуктивность первичной катушки за счёт взаимной индукции M; в разомкнутом состоянии (см. рис. 9, *а*, слева) цепь вторичного контура разорвана, и его вклад в результирующую индуктивность отсутствует, что соответствует состояниям «Включено» и «Выключено» на рис. 9, *б*.

В условиях дефицита площади на кристалле или печатной плате часто применяются многослойные катушки. Однако простое вертикальное размещение витков приводит к резкому росту межслойной паразитной емкости. В исследованиях [39, 40] проведен детальный анализ влияния последовательности соединения слоев на паразитную емкость четырехслойной структуры. Как показано на рис. 10 и в таблице, изменение порядка коммутации слоев (например, соединение слоя 1 со слоем 3, а затем 2 и 4) позволяет существенно снизить разность потенциалов между соседними витками.

Результаты показывают, что оптимизированный способ коммутации (чередующаяся намотка) снижает эквивалентную емкость до 77% по сравнению с последовательным соединением смежных слоев.

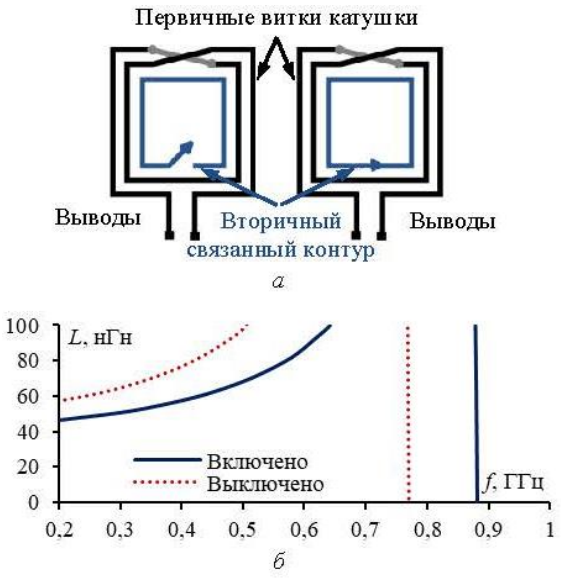


Рис. 9. Способ коммутации перестраиваемой катушки индуктивности (*а*) и ее индуктивность до и после переключения (*б*) [38]

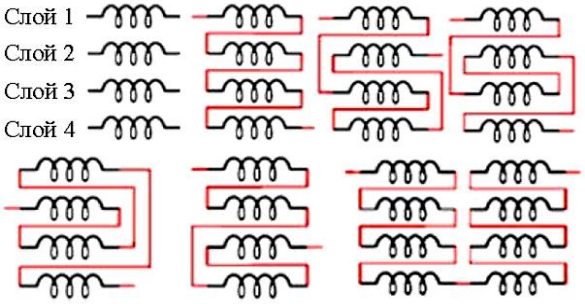


Рис. 10. Различные последовательные соединения слоев печатной платы [39, 40]

Сравнение эквивалентных емкостей для различных расположений витков внутри слоев

Намотка	Эквивалентная емкость
Типы соединений	Относительная емкость, % (1, 2, 3, 4)
(1,2,3,4)	100
(1,3,4,2)	142
(1,3,2,4)	135
(2,3,1,4)	138
(1,2,4,3)	110
Раздельные витки	77

Этот принцип, продемонстрированный на печатных платах, имеет прямой потенциал для применения в многоуровневых металлизациях интегральных катушек индуктивностей.

В работе [41] исследуется влияние геометрии подключения выводов на характеристики двухслойной спиральной катушки. Продемонстрировано, что оптимизация точек соединения витков (например, использование промежуточных отводов или перекрестное соединение слоев) позволяет модифицировать распределение электрического поля.

На рис. 11 видно, что различные способы соединения (внешнее, промежуточное, чередующееся подключение) при неизменной геометрии самой спирали приводят к заметному сдвигу резонансной частоты.

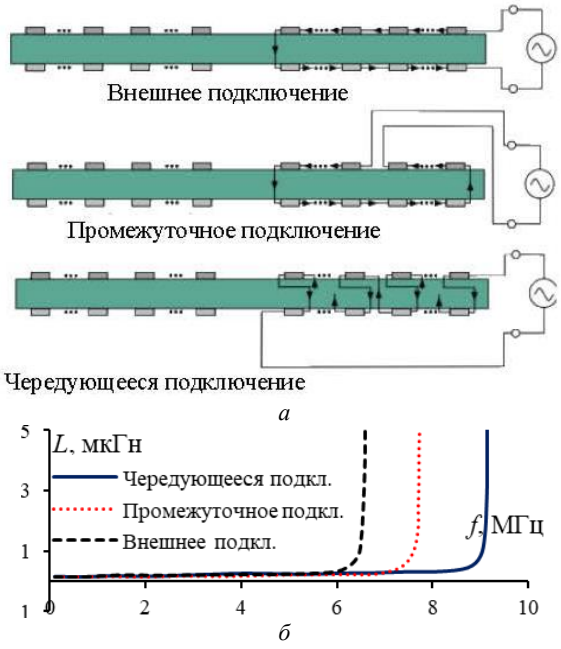


Рис. 11. Поперечное сечение и способы соединения катушки индуктивности

Это подтверждает тезис о том, что топология соединений является самостоятельным инструментом настройки частотной дисперсии параметров.

Постановка исследовательских задач

В качестве целевой технологии для дальнейших исследований рассматриваются катушки индуктивности на полупроводниковых подложках группы A3B5, в частности, на основе арсенида галлия. Данная технология остаётся наиболее востребованной при реализации СВЧ-монолитных интегральных схем благодаря высокой подвижности носителей заряда и полуизолирующим свойствам подложки, обеспечивающим низкие диэлектрические потери на высоких частотах, что делает целесообразным сосредоточить дальнейшие исследования именно на данном классе структур.

Проведенный обзор выявил существенный пробел в современной научно-технической литературе: отсутствие комплексного анализа влияния способов соединения на частотную дисперсию параметров интегральных индуктивностей. Большинство существующих решений рассматривают либо геометрию изолированной катушки, либо схемотехнику на уровне печатных плат.

На основании выявленных закономерностей и пробелов формулируются следующие научные гипотезы, определяющие направления дальнейших исследований:

– предполагается, что использование фрактальной геометрии (например, кривых Мура) в конструкции связанных катушек позволит повысить степень управляемости индуктивностью и расширить диапазон перестройки за счет более развитой поверхности взаимодействия полей;

– можно предположить, что применение проводников с оптимизированной нерегулярной шириной в трансформаторных включениях позволит найти оптимальный аналитический закон профилирования ширины для трансформаторных включений;

– комбинация подходов (разнесение слоев и профилирование ширины) должна привести к одновременному снижению паразитной емкости и омических потерь, что обеспечит максимальное расширение рабочей полосы частот;

– предполагается, что угловая ориентация является дополнительной степенью свободы, позволяющей точно настраивать коэффициент связи и минимизировать паразитные электромагнитные связи;

– предполагается, что существует оптимальная конфигурация коммутации, обеспечивающая минимальную частотную дисперсию параметров индуктивности за счет взаимной компенсации паразитных реактивностей.

Заключение

В работе проведен систематический обзор методов улучшения характеристик интегральных индуктивностей. Показано, что традиционные методы геометрической оптимизации (восьмиугольные спирали, экранирование подложки) приближаются к теоретическому пределу эффективности. В то же время конструктивно-топологические методы (оптимизация межслойных соединений, использование связанных структур с управляемой геометрией) обладают значительным нераскрытым потенциалом. Подтверждение сформулированных гипотез потребует проведения серии вычислительных и лабораторных экспериментов, результаты которых могут лечь в основу новых методик проектирования СВЧ-компонентов и устройств.

Литература

1. Nguyen N.M. Si IC-Compatible Inductors and LC Passive Filters / N.M. Nguyen, R.G. Meyer // IEEE Journal of Solid-State Circuits. – 1990. – Vol. 25, No. 4. – P. 1028–1031.

2. Оценка влияния отклонения конструктивных параметров интегральных LC-фильтров на их частотные характеристики / Ю.И. Егошин, А.А. Тюменцева, Т.С. Тимошенко, А.Н. Яковлев // Техника радиосвязи. – 2025. – № 1 (64). – С. 64–71.

3. Дубровин В.С. Оценка стабильности частотных характеристик управляемого фазовращателя // Вестник Астраханского гос. техн. ун-та. – Сер.: Управление, вычислительная техника и информатика. – 2015. – № 4. – С. 92–100.

4. Алимов А.П. Моделирование и проектирование избирательных усилителей второго порядка / А.П. Алимов, Б.Ф. Змий // Информатика: проблемы, методология, технологии: матер. XVI Междунар. науч.-метод. конф., Воронеж, 11–12 февраля 2016 г. / под ред. Н.А. Тюкачева. – Воронеж: Научно-исследовательские публикации, 2016. – С. 16–21.

5. Устройства приема и обработки сигналов в упражнениях и задачах: учеб. пособие / В.А. Аржанов, А.П. Науменко, А.И. Одинец, Т.В. Багаева. – Омск: Изд-во ОмГТУ, 2018. – 225 с.

6. A design methodology for RF/mmWave LNAs in 22 nm FD-SOI with cross-coupling-aware nested inductors and on-chip baluns / S. Drakakis, A. Michailidis, D. Tzagkas // Electronics. – 2026. – Vol. 15, No. 1. – Art. 25.

7. Пат. 4 867 703 США B2. Thin film structures with negative inductance and methods for fabricating inductors comprising the same / A. Widjaja, A. Sarangan (США). № 12/427,296; заявл. 21.04.2009; опубл. 07.06.2011. – 7 с.

8. Chen L. High inductance density in CMOS-compatible magnetically integrated 3D microinductors for radio-frequency applications / L. Chen, Z. Qiao, S. Liu, T. Noulis // Nature Communications. – 2025. – Vol. 16, No. 1. – P. 10072.

9. Ezhilarasu G. Integrated micro-capacitors and microinductors for next generation electronics / G. Ezhilarasu, S.S. Iyer // IEEE Nanotechnology Magazine. – 2024. – Vol. 18, No. 2. – P. 30–40.

10. Mi X. High-frequency applications / X. Mi, S. Ueda // Advanced Microwave Circuits and Systems / ed. V. Zhurbenko. – Hamburg: BoD – Books on Demand, 2010. – P. 249–268.

11. Shin K.R. Compact 5G n77 band pass filter with through silicon via (TSV) IPD technology / K.R. Shin, J. Arendell, K. Eilert // 2019 IEEE 20th Wireless and Microwave Technology Conference (WAMICON), Cocoa Beach, FL, USA, April 2019. – New York: IEEE, 2019. – P. 1–3.

12. Shaltout A.H. Layout optimization of planar inductors for high-efficiency integrated power converters / A.H. Shaltout, S. Gregori // Analog Integrated Circuits and Signal Processing. – 2020. – Vol. 102, No. 1. – P. 155–167.

13. Птицын Д.В. О влиянии параметров регулируемой катушки индуктивности на значение мощности, измеряемой в RLC-цепях / Д.В. Птицын, Е.В. Птицына, А.Б. Кувалдин // Энерго- и ресурсосбережение XXI век: сборник

матер. XII Междунар. науч.-практ. интернет-конф., Орёл, март–июнь 2014 г. – Орёл, 2014. – С. 61–64.

14. Computational analysis on self-resonance frequency of solenoid and planar inductor / M.S. Choi, S.J. Kim, Y.S. Lee, C.H. Cho, I.H. Seong, W.N. Jeong, S.J. You // Applied Science and Convergence Technology. – 2023. – Vol. 32, No. 2. – P. 54–57.

15. Агарков Н.Е. Влияние паразитной ёмкости согласующей цепи на вносимые потери антенного согласующего устройства // Техника радиосвязи. – 2025. – № 1 (64). – С. 40–55.

16. An overview of fully on-chip inductors / R. Ondica, M. Kovac, A. Hudec, R. Ravasz, D. Maljar, V. Stopjakova, D. Arbet // Radioengineering. – 2023. – Vol. 32, No. 1. – P. 11–22.

17. Spataro S. A Geometrically Scalable Lumped Model for Spiral Inductors in Radio Frequency GaN Technology on Silicon / S. Spataro, G. Sapone, M. Giuffrida, E. Ragonese // Electronics. – 2024. – Vol. 13, No. 13. – P. 2665.

18. Chen H.H. Analytic Design of On-Chip Spiral Inductor with Variable Line Width / H.H. Chen, Y.W. Hsu // Electronics. – 2022. – Vol. 11, No. 13. – P. 2029.

19. A review of engineering techniques for CMOS on-chip inductor design and quality factor enhancement from MHz-to-GHz frequency domains / S. Mariappan, J. Rajendran, N. Idros, A. A. Manaf, N. Kumar, A. Alghaihab, B.S. Yarman // IEEE Access. – 2025. – Vol. 13. – P. 140473–140499.

20. Борейко Д.А. Усовершенствованный подход к моделированию интегральных компонентов для LC-фильтров: спец-ть 05.13.10: дис. … канд. техн. наук. – Омск: Омский гос. техн. ун-т, 2017. – 187 с.

21. Ерохин В.В. Влияние характеристик технологического процесса на проектирование СВЧ LC-фильтров в СНК / В.В. Ерохин, С.А. Завьялов // Обмен опытом в области создания сверхширокополосных радиоэлектронных систем: матер. IX Всерос. науч.-техн. конф., Омск, 26–27 апреля 2022 г. – Омск: Омский гос. техн. ун-т, 2022. – С. 108–117.

22. Wu N. Microstructure Characterization and Interfacial Reactions between Au-Sn Solder and Different Back Metallization Systems of GaAs MMICs / N. Wu, Y. Hu, S. Sun // Materials. – 2020. – Vol. 13, No. 6. – Art. 1266.

23. Добуш И.М. Построение моделей пассивных элементов и автоматизированное проектирование СВЧ-монолитных усилителей с учетом влияния температуры: спец-ть 05.12.07 «Антенны, СВЧ-устройства и их технологии»: дис. … канд. техн. наук. – Томск, ТУСУР, 2012. – 199 с.

24. Effect of differently shaped spiral inductors on micro power integrated circuit / A. Namoune, R. Taleb, M.R. Benzidane, R. Melati, N. Mansour, Z. Benbouzid, H. Saidi // The Scientific Bulletin of Electrical Engineering Faculty. – 2024. – Vol. 24, No. 2. – P. 38–46.

25. Борейко Д.А. Исследование параметров катушек индуктивности, выполненных по LTCC-технологии / Д.А. Борейко, Т.С. Дьяченко, А.И. Тюменцев // Техника радиосвязи. – 2012. – № 17. – С. 73–83.

26. Tayenjam S. High-Q variable pitch spiral inductors for increased inductance density and figure-of-merit / S. Tayenjam, V. Vanukuru, K. Sundaram // IEEE Transactions on Electron Devices. – 2019. – Vol. 66, No. 4. – P. 4481–4485.

27. Chavva S. Experimental analysis of irregularly shaped octagonal on-chip inductors for improving area-efficiency in CMOS RFICs for millimeter wave applications / S. Chavva, I. Raja // Integration. – 2025. – Vol. 100. – P. 102259.

28. Shaltout A.H. Layout optimization of planar inductors for high-efficiency integrated power converters / A.H. Shaltout, S. Gregori // Analog Integrated Circuits and Signal Processing. – 2020. – Vol. 102, No. 1. – P. 155–167.

29. Chavva S. N-well patterning of P-type CMOS substrate for improving quality factor of on-chip inductors at millimeter wave frequencies / S. Chavva, I. Raja // 2025 38th International Conference on VLSI Design and 2024 23rd International Conference on Embedded Systems (VLSID), January 2025. – New York: IEEE, 2025. – P. 186–190.

30. Bjelica J.M. Modeling and performance analysis of planar fractal inductors / J.M. Bjelica, N.M. Djuric, S.M. Djuric // IEEE Transactions on Magnetics. – 2020. – Vol. 56, No. 11. – P. 1–8.

31. Alpaslan H. A new active device namely S-CCI and its applications: simulated floating inductor and quadrature oscillators / H. Alpaslan, E. Yuce, S. Minaei // IEEE Transactions on Circuits and Systems I: Regular Papers. – 2022. – Vol. 69, No. 9. – P. 3554–3564.

32. High-performance compact 3-D solenoids for RF applications / A. Cayron, C. Viallon, O. Bushueva, A. Ghannam, T. Parra // IEEE Microwave and Wireless Components Letters. – 2018. – Vol. 28, No. 6. – P. 479–481.

33. Advances and perspectives in magnetic-integrated inductors for RF ICs / W. Hao, X. Li, Z. Pan, R. Miao, Z. Yue, C. Yang, A.Z. Wang // IEEE Transactions on Materials for Electron Devices. – 2024. – Vol. 1. – P. 106–120.

34. An overview of fully integrated switching power converters based on switched-capacitor versus inductive approach and their advanced control aspects / G. Kiran Kumar, T. Kobaku, S. Sahoo, B. Subudhi, D. Elangovan, F. Blaabjerg // Energies. – 2021. – Vol. 14, No. 11. – P. 3250.

35. Wouters H. Integrated inductor-transformers for high-frequency converters: an overview / H. Wouters, W. Martinez // IEEE Transactions on Power Electronics. – 2025. – Vol. 40, No. 9. – P. 13157–13176.

36. Stability improvement of isolated multiple-output DC/DC converter using coupled inductors / E. Sanchis-Kilders, A. Ferreres, J.L. Gasent-Blesa, E. Maset, V. Esteve, J. Jordan, J.B. Ejea // IEEE Transactions on Aerospace and Electronic Systems. – 2016. – Vol. 52, No. 4. – P. 1644–1653.

37. Fabrication of QFN-packaged miniaturized GaAs-based bandpass filter with intertwined inductors and dendritic capacitor / J. Chen, Z.J. Wang, B.H. Zhu, E.S. Kim, N.Y. Kim // Materials. – 2020. – Vol. 13, No. 8. – P. 1932.

38. Experimental evaluation of coupled-line tunable inductors with switchable mutual coupling / Y. Kim, J. Lee, S. Kim, C. Park // Electronics. – 2025. – Vol. 14, No. 16. – P. 3228.

39. Kolahian P. A comprehensive review on planar magnetics and the structures to reduce the parasitic elements and improve efficiency / P. Kolahian, F. Grimm, M. Baghdadi // Energies. – 2023. – Vol. 16, No. 7. – P. 3254.

40. Shahabi Ghahfarokhi N. Minimising capacitive couplings and distributing copper losses in planar magnetic elements: Masters by Research thesis. – Brisbane: Queensland University of Technology, 2010.

41. Lope I. First self-resonant frequency of power inductors based on approximated corrected stray capacitances / I. Lope, C. Carretero, J. Acero // IET Power Electronics. – 2021. – Vol. 14. – P. 257–267.

Власов Сергей Владиславович
Аспирант, ассистент каф. сверхвысокочастотной и квантовой радиотехники (СВЧиКР)
Томского государственного университета систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0002-3848-1079
Тел.: +7 (382-2) 41-36-43
Эл. почта: sergei.v.vlasov@tusur.ru

Жечев Евгений Сергеевич
Канд. техн. наук, доцент каф. СВЧиКР ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0003-4469-7033
Тел.: +7 (382-2) 41-36-43
Эл. почта: zhechev75@gmail.com

Поступила в редакцию: 01.04.2026.
Принята к публикации: 27.08.2026.

Vlasov S.V., Zhechev Y.S.

Methods for improving the quality factor and reducing the frequency dispersion of integrated inductor parameters: a review and formulation of research problems

Relevance. This paper presents a review of current methods for improving the quality factor and reducing the frequency dispersion of integrated inductor parameters. Parasitic capacitive coupling and ohmic losses limit the quality factor and operating bandwidth of integrated inductors used in matching circuits, filters, and resonators of the microwave range. ***Objective.*** The study aims to systematize current approaches to the design of planar and three-dimensional integrated inductors aimed at increasing the quality factor and reducing the frequency dispersion of parameters, as well as formulating research hypotheses in the field of design-topological optimization. ***Methods.*** An analytical review of design solutions, including irregular conductor width, fractal geometries, multilayer structures, substrate shielding techniques, and connection method. ***Scientific novelty.*** A comprehensive consideration of the factors determining the quality factor and frequency dispersion of integrated inductors is presented, together with design-topological aspects that have previously been investigated mostly separately. ***Results.*** It is shown that optimizing the winding geometry and applying multilevel metallization improve the quality factor, while the influence of connection methods and the mutual orientation of multilayer structures remains insufficiently studied. Research hypotheses are formulated concerning the potential improvement of frequency characteristics through the comprehensive optimization of the geometry and topology of integrated inductors. ***Practical significance.*** The review results can be used in the design of microwave-band matching circuits, filters, and resonators at the stage of selecting the geometry and topology of integrated inductors.

Keywords: integrated inductor, quality factor, parameter frequency dispersion, self-resonance, parasitic capacitance, microwave circuit topology, fractal structures.
DOI: 10.21293/1818-0442-2026-29-2-7-15

References

1. Nguyen N.M., Meyer R.G. Si IC-Compatible Inductors and LC Passive Filters. *IEEE Journal of Solid-State Circuits*, 1990, vol. 25, no. 4, pp. 1028–1031.

2. Egoshin Yu.I., Tyumentseva A.A., Timoshenko T.S., Yakovlev A.N. [Assessment of the influence of design parameters of integral lc filters on their frequency characteristics]. *Radio communication technology*, 2025, no. 1 (64), pp. 64–71 (in Russ.).

3. Dubrovin V.S. [Assessment of the stability of frequency characteristics of controlled phase shifter]. *Vestnik of Astrakhan State Technical University. Series: management, computer science and informatics,* 2015, no. 4, pp. 92–100 (in Russ.).

4. Alimov A.P., Zmiy B.F. [Modelling and design of second-order selective amplifiers]. Informatika: problemy, metodologiya, tekhnologii. Materialy XVI Mezhdunarodnoy nauchno-metodicheskoy konferentsii [*Informatics: problems, methodology, technologies. Proc. of the XVI International scientific and methodological conference*], Voronezh, February 11–12, 2016. Voronezh, Nauchno-issledovatel'skie publikatsii Publ., 2016, pp. 16–21 (in Russ.).

5. Arzhanov V.A., Naumenko A.P., Odinets A.I., Bagaeva T.V. *Ustroystva priema i obrabotki signalov v uprazhneniyakh i zadachakh* [Signal reception and processing devices in exercises and problems]. Omsk, OmGTU Publ., 2018, 255 p. (in Russ.).

6. Drakakis S., Michailidis A., Tzagkas D., Pavlidis V.F., Noulis T. A design methodology for RF/mmWave LNAs in 22 nm FD-SOI with cross-coupling-aware nested inductors and on-chip baluns. *Electronics*, 2025, vol. 15, no. 1, art. 25.

7. Widjaja A., Sarangan A. *Thin film structures with negative inductance and methods for fabricating inductors comprising the same*. Patent USA, no. 7956715 B2, 2011.

8. Chen L., Qiao Z., Liu S., Noulis T. High inductance density in CMOS-compatible magnetically integrated 3D microinductors for radio-frequency applications. *Nature Communications*, 2025, vol. 16, no. 1, p. 10072.

9. Ezhilarasu G., Iyer S.S. Integrated micro-capacitors and micro-inductors for next generation electronics. *IEEE Nanotechnology Magazine*, 2024, vol. 18, no. 2, pp. 30–40.

10. Mi X., Ueda S. High-frequency applications. In: *Advanced Microwave Circuits and Systems*, ed. V. Zhurbenko. Hamburg, BoD – Books on Demand Publ., 2010, pp. 249–268.

11. Shin K.R., Arendell J., Eilert K. Compact 5G n77 band pass filter with through silicon via (TSV) IPD technology. 2019 IEEE 20th Wireless and Microwave Technology Conference (WAMICON), Cocoa Beach, FL, USA, April 2019. New York, IEEE Publ., 2019, pp. 1–3.

12. Shaltout A.H., Gregori S. Layout optimization of planar inductors for high-efficiency integrated power converters. *Analog Integrated Circuits and Signal Processing*, 2020, vol. 102, no. 1, pp. 155–167.

13. Ptitsyn D.V., Ptitsyna E.V., Kuvaldin A.B. [On the influence of parameters of a variable inductor on the power value measured in RLC circuits]. *Energo- i resursosberezheniye XXI vek. Sbornik materialov XII Mezhdunarodnoy nauchno-prakticheskoy internet-konferentsii* [Energy and resource saving XXI century. Proc. of the XII International scientific and practical internet conference], Oryol, March–June 2014. Oryol, 2014, pp. 61–64 (in Russ.).

14. Choi M.S., Kim S.J., Lee Y.S., Cho C.H., Seong I.H., Jeong W.N., You S.J. Computational analysis on self-resonance frequency of solenoid and planar inductor. *Applied Science and Convergence Technology*, 2023, vol. 32, no. 2, pp. 54–57.

15. Agarkov N.E. [Influence of parasitic capacitance of the matching circuit on the insertion losses of the antenna matching device]. *Radio communication technology*, 2025, no. 1 (64), p. 40–55 (in Russ.).

16. Ondica R., Kovac M., Hudec A., Ravasz R., Maljar D., Stopjakova V., Arbet D. An overview of fully on-chip inductors. *Radioengineering*, 2023, vol. 32, no. 1, pp. 11–22.

17. Spataro S., Sapone G., Giuffrida M., Ragonese E. A Geometrically Scalable Lumped Model for Spiral Inductors in Radio Frequency GaN Technology on Silicon. *Electronics*, 2024, vol. 13, no. 13, p. 2665.

18. Chen H.H., Hsu Y.W. Analytic Design of On-Chip Spiral Inductor with Variable Line Width. *Electronics*, 2022, vol. 11, no. 13, p. 2029.

19. Mariappan S., Rajendran J., Idros N., Manaf A.A., Kumar N., Alghaihab A., Yarman B.S. A review of engineering techniques for CMOS on-chip inductor design and quality factor enhancement from MHz-to-GHz frequency domains. *IEEE Access*, 2025, vol. 13, pp. 140473–140499.

20. Boreyko D.A. *Usovershenstvovannyy podkhod k modelirovaniyu integral'nykh komponentov dlya LC-fil'trov* [An improved approach to modeling integrated components for LC-filters]. Diss. Omsk, OmGTU, 2017, 187 p. (in Russ.).

21. Erokhin V.V., Zavyalov S.A. Vliyanie kharakteristik tekhnologicheskogo protsessa na proektirovanie SVCH LC-fil'trov v SNK [Influence of process characteristics on the design of microwave LC filters in SoC]. *Obmen opytom v oblasti sozdaniya sverkhshirokopolosnykh radioelektronnykh sistem. Materialy IX Vserossiyskoy nauchno-tekhnicheskoy konferentsii* [Experience exchange in the field of ultra-wideband radio-electronic systems. Proc. of the IX All-Russian scientific and technical conference], Omsk, April 26–27, 2022. Omsk, OmGTU Publ., 2022, pp. 108–117 (in Russ.).

22. Wu N., Hu Y., Sun S. Microstructure Characterization and Interfacial Reactions between Au-Sn Solder and Different Back Metallization Systems of GaAs MMICs. *Materials*, 2020, vol. 13, no. 6, art. 1266.

23. Dobush I.M. *Postroyeniye modeley passivnykh elementov i avtomatizirovannoye proyektirovaniye SVCH monolitnykh usiliteley s uchetom vliyaniya temperatury* [Building models of passive elements and computer-aided design of microwave monolithic amplifiers considering temperature effects]. Cand. Diss. Tomsk: TUSUR, 2012, 199 p. (in Russ.).

24. Namoune A., Taleb R., Benzidane M.R., Melati R., Mansour N., Benbouzid Z., Saidi H. Effect of differently shaped spiral inductors on micro power integrated circuit. *The Scientific Bulletin of Electrical Engineering Faculty*, 2024, vol. 24, no. 2, pp. 38–46.

25. Boreyko D.A., Dyachenko T.S., Tyumentsev A.I. *Issledovaniye parametrov katushek induktivnosti, vypolnennykh po LTCC tekhnologii* [Investigation of parameters of inductors made using LTCC technology]. *Radio communication technology*, 2012, no. 17, pp. 73–83 (in Russ.).

26. Tayenjam S., Vanukuru V., Sundaram K. High-Q variable pitch spiral inductors for increased inductance density and figure-of-merit. *IEEE Transactions on Electron Devices*, 2019, vol. 66, no. 10, pp. 4481–4485.

27. Chavva S., Raja I. Experimental analysis of irregularly shaped octagonal on-chip inductors for improving area-efficiency in CMOS RFICs for millimeter wave applications. *Integration*, 2025, vol. 100, p. 102259.

28. Shaltout A.H., Gregori S. Layout optimization of planar inductors for high-efficiency integrated power converters. *Analog Integrated Circuits and Signal Processing*, 2020, vol. 102, no. 1, pp. 155–167.

29. Chavva S., Raja I. N-well patterning of P-type CMOS substrate for improving quality factor of on-chip inductors at millimeter wave frequencies. *2025 38th International Conference on VLSI Design and 2024 23rd International Conference on Embedded Systems (VLSID)*, January 2025. New York, IEEE Publ., 2025, pp. 186–190.

30. Bjelica J.M., Djuric N.M., Djuric S.M. Modeling and performance analysis of planar fractal inductors. *IEEE Transactions on Magnetics*, 2020, vol. 56, no. 11, pp. 1–8.

31. Alpaslan H., Yuce E., Minaei S. A new active device namely S-CCI and its applications: simulated floating inductor and quadrature oscillators. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2022, vol. 69, no. 9, pp. 3554–3564.

32. Cayron A., Viallon C., Bushueva O., Ghannam A., Parra T. High-performance compact 3-D solenoids for RF applications. *IEEE Microwave and Wireless Components Letters*, 2018, vol. 28, no. 6, pp. 479–481.

33. Hao W., Li X., Pan Z., Miao R., Yue Z., Yang C., Wang A.Z. Advances and perspectives in magnetic-integrated inductors for RF ICs. *IEEE Transactions on Materials for Electron Devices*, 2024, vol. 1, pp. 106–120.

34. Kiran Kumar G., Kobaku T., Sahoo S., Subudhi B., Elangovan D., Blaabjerg F. An overview of fully integrated switching power converters based on switched-capacitor versus inductive approach and their advanced control aspects. *Energies*, 2021, vol. 14, no. 11, p. 3250.

35. Wouters H., Martinez W. Integrated inductor-transformers for high-frequency converters: an overview. *IEEE Transactions on Power Electronics*, 2025, vol. 40, no. 9, pp. 13157–13176.

36. Sanchis-Kilders E., Ferreres A., Gasent-Blesa J.L., Maset E., Esteve V., Jordan J., Ejea J.B. Stability improvement of isolated multiple-output DC/DC converter using coupled inductors. *IEEE Transactions on Aerospace and Electronic Systems*, 2016, vol. 52, no. 4, pp. 1644–1653.

37. Chen J., Wang Z.J., Zhu B.H., Kim E.S., Kim N.Y. Fabrication of QFN-packaged miniaturized GaAs-based bandpass filter with intertwined inductors and dendritic capacitor. *Materials*, 2020, vol. 13, no. 8, p. 1932.

38. Kim Y., Lee J., Kim S., Park C. Experimental evaluation of coupled-line tunable inductors with switchable mutual coupling. *Electronics*, 2025, vol. 14, no. 16, p. 3228.

39. Kolahian P., Grimm F., Baghdadi M. A comprehensive review on planar magnetics and the structures to reduce the parasitic elements and improve efficiency. *Energies*, 2023, vol. 16, no. 7, p. 3254.

40. Shahabi Ghahfarokhi N. Minimising capacitive couplings and distributing copper losses in planar magnetic elements. Masters by Research thesis. Brisbane, Queensland University of Technology, 2010.

41. Lope I., Carretero C., Acero J. First self-resonant frequency of power inductors based on approximated corrected stray capacitances. *IET Power Electronics*, 2021, vol. 14, pp. 257–267.

Sergey V. Vlasov
Postgraduate student, Assistant, Department of Microwave and Quantum Radio Engineering, Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0002-3848-1079
Phone: +7 (382-2) 41-36-43
Email: sergei.v.vlasov@tusur.ru

Yevgeniy S. Zhechev
Candidate of Engineering Sciences, Associate Professor, Department of Microwave and Quantum Radio Engineering, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0003-4469-7033
Phone: +7 (382-2) 41-36-43
Email: zhechev75@gmail.com

Received: 01.04.2026.
Accepted: 27.08.2026.

УДК 621.396.67
Шифр специальности ВАК: 2.2.14

Д.А. Юзвик, М.А. Степанов

Метод формирования нескольких минимумов электромагнитного поля в требуемых точках пространства

Актуальность. Рассмотрен метод формирования нескольких минимумов электромагнитного поля в требуемых областях пространства с целью снижения уровня помех для чувствительного оборудования. ***Цель исследования.*** Разработка метода расчета комплексных амплитуд излучаемых сигналов для формирования нескольких заданных областей пониженной амплитуды напряженности электрического поля при одновременной установке уровней поля в остальных точках приема с использованием регуляризации Тихонова. ***Методы.*** Система линейных алгебраических уравнений была уравновешена введением регуляризации Тихонова, обеспечивающей единственность решения. Также была поставлена оптимизационная задача по минимизации квадратичной невязки между требуемым и получаемым распределениями электрического поля. Проверка теоретических результатов выполнена путем математического моделирования. ***Научная новизна.*** Использована связь между комплексными амплитудами в точках приема и комплексными амплитудами излучаемых сигналов через матрицу S-параметров. Отмечено, что вычисление комплексных амплитуд излучаемых сигналов может приводить к недоопределенной системе линейных алгебраических уравнений, вследствие чего решения могут быть неустойчивыми: их может быть бесконечно много или решения может не быть вовсе. ***Результаты.*** Предложены три конфигурации электрического поля, для каждой из которых определены комплексные амплитуды излучаемых сигналов. Результаты математического моделирования подтвердили полученные теоретические положения. ***Практическая значимость.*** Полученные результаты могут быть использованы для формирования требуемых распределений электромагнитного поля с несколькими минимумами в заданных областях пространства.

Ключевые слова: антенные решетки, оптимизация, СЛАУ, электромагнитное поле, минимумы электрического поля.
DOI: 10.21293/1818-0442-2026-29-2-16-23

Введение

Фокусировка электромагнитного поля в требуемые точки пространства является одним из способов повышения отношения сигнал/шум (ОСШ) на апертуре приемной антенны абонентского устройства. Повышение ОСШ позволяет использовать высокие порядки квадратурной амплитудной модуляции, например, 4096-QAM. Использование высоких порядков QAM модуляции позволяет повышать скорость передачи информации в беспроводных сетях [1, 2]. В работах [3–5] были рассмотрены вопросы, касающиеся фокусировки линейных антенных решеток на конечное расстояние, изучены эффекты смещения максимума амплитуды напряженности электрического поля в сторону антенной решетки относительно точки фокуса.

В работе [6] рассмотрены вопросы, касающиеся формирования нескольких максимумов амплитуды напряженности электрического поля в требуемых областях пространства. Формирование максимумов амплитуды напряженности электрического поля обосновано тем, что могут быть заранее известны места расположения устройств-абонентов, особо актуально это может быть при обеспечении связью устройств умного дома, интернета вещей, так как местоположение устройств умного дома и интернета вещей, как правило, статично. Таким образом, возможно обеспечивать связью такие устройства и создавать минимум помех для остальных устройств. У этого способа есть другая сторона медали, например, в медицинских учреждениях располагается оборудование, которое может быть чувствительным к излишнему электромагнитному излучению от беспроводных телекоммуникационных сетей. При этом расположение абонентских устройств в подобных локациях предсказать проблематично. Поэтому разумно снижать амплитуду напряженности электрического поля в областях, где расположено чувствительное к излучению оборудование, а в остальных областях пространства обеспечивать сигнал на требуемом уровне. Целенаправленное формирование одного минимума амплитуды напряженности электрического поля на фокальной оси возможно за счет фазировки антенной решетки.

На практике чувствительного к электромагнитной обстановке оборудования в помещении может быть больше одного и может потребоваться создать несколько зон в пространстве, где обеспечиваются минимумы амплитуды напряженности электрического поля. В настоящей работе под минимумом амплитуды напряженности электрического поля понимается локальное понижение амплитуды напряженности электрического поля в окрестности заданной точки пространства относительно амплитуды в других точках пространства.

Стоит отметить, что предлагаемый в работе метод не заменяет классические методы обеспечения электромагнитной совместимости: экранировка аппаратуры, фильтрация или частотное разделение каналов и т.д. Снижение напряженности электрического поля в окрестности заданных точек может служить дополнительным способом снижения электро-

магнитных помех для чувствительного оборудования, расположенного в этих точках.

Исходя из этого, **целью работы** является разработка метода расчета комплексных амплитуд излучаемых сигналов для формирования нескольких заданных областей пониженной амплитуды напряженности электрического поля при одновременной установке уровней поля в остальных точках приема с использованием регуляризации Тихонова.

Математическая модель системы

В работе [6] был предложен способ представления антенной системы в виде многополюсника СВЧ, связывающий комплексные амплитуды излучаемых сигналов, комплексные амплитуды в точках приема через матрицу *S*-параметров. Под точкой приема подразумевается произвольная точка в объеме анализируемого пространства, где требуется управлять амплитудой напряженности электрического поля. Рассматриваемая модель подразумевает, что анализ происходит в свободном пространстве. Поляризация приемных и передающих антенн считается согласованной, так что поляризационными эффектами можно пренебречь. За основу математической модели возьмем этот способ и кратко его напомним [6].

Предлагаемый способ основан на следующем подходе. В рассматриваемой области пространства размещают N излучающих антенн и M точек приема, положение которых выбирается произвольным образом. Один из возможных вариантов расположения антенн и точек приема приведен на рис. 1.

Начало координат в области, содержащей излучающие антенны и точки приема, находится в геометрическом центре рассматриваемой области. Оси координат X и Y ориентированы параллельно антенным решеткам. Ось Z направлена вверх.

В показанном примере антенны установлены по границе области (вдоль стен помещения), тогда как точки приема находятся внутри нее. Для каждой точки приема заранее задают требуемое значение амплитуды напряженности электрического поля. Далее определяются комплексные коэффициенты передачи между каждой излучающей антенной и каждой точкой приема.

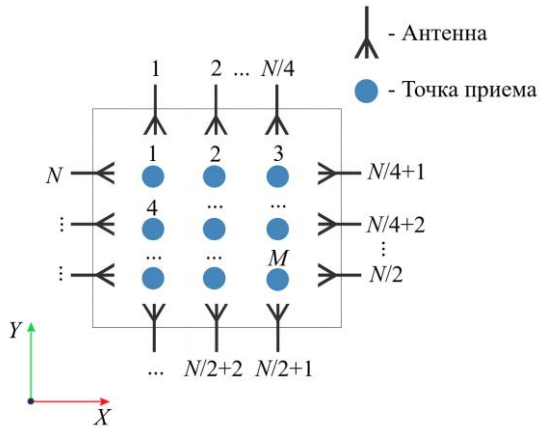


Рис. 1. Геометрическое расположение излучающих антенн и точек приема [6]

Путем решения системы линейных алгебраических уравнений (СЛАУ) определяются комплексные амплитуды сигналов, которые нужно подвести к каждой излучающей антенне для получения требуемого распределения поля в пространстве. Эффективность способа была продемонстрирована в работе [6].

Кратко напомним основную идею способа управления распределением электрического поля в заданном пространстве, показанного в работе [6].

В [6] приняты обозначения комплексных амплитуд излучаемых сигналов:

$$\dot{a}_n = a_n \cdot e^{j\cdot\phi_n}, \qquad (1)$$

где a_n – амплитуда сигнала, излучаемого из n-й точки; ϕ_n – начальная фаза сигнала, излучаемого из n-й точки; j – мнимая единица.

Комплексная амплитуда в точках приема [6]:

$$\dot{b}_m = b_m \cdot e^{j\cdot\psi_m}, \qquad (2)$$

где b_m – амплитуда сигнала в m-й точке; ψ_m – фаза сигнала в m-й точке.

Комплексные амплитуды сигналов в точках излучения можно представить в виде вектор-столбца $[\dot{\mathbf{A}}]$, а комплексные амплитуды сигналов в точках приема – вектор-столбца $[\dot{\mathbf{B}}]$.

Среда распространения электромагнитных волн от точек передачи к точкам приема, представленная на рис. 1, может рассматриваться как СВЧ-многополюсник (рис. 2) [6]. Количество входов многополюсника равно $N+M$, где N – количество излучающих антенн, M – количество точек приема. Очевидно, что входы с 1 по N соответствуют излучающим антеннам, входы с N+1 по $N+M$ соответствуют точкам приема.

Следовательно, задавая вектор-столбец выходных сигналов многополюсника $[\dot{\mathbf{Y}}]$, можно определять вектор-столбец входных сигналов $[\dot{\mathbf{X}}]$. Их связывает матрица рассеяния (*S*-параметров – $[\dot{\mathbf{S}}]$):

$$[\dot{\mathbf{Y}}] = [\dot{\mathbf{S}}] \cdot [\dot{\mathbf{X}}]. \qquad (3)$$

Зададим граничные условия. Поскольку управление распределением напряженности электрического поля должно осуществляться внутри объема помещения, а не на его границах, элементы вектор-столбца $[\dot{\mathbf{Y}}]$ от 0 до N принимаются равными нулю.

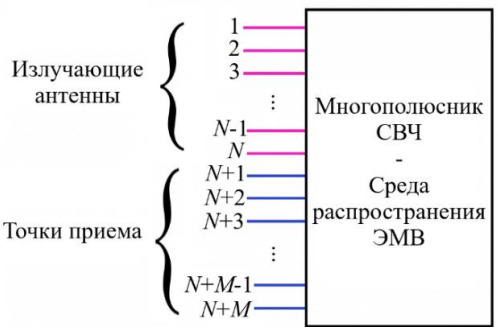


Рис. 2. Схематическое представление излучающих элементов, среды распространения и точек приема в виде многополюсника СВЧ [6]

Элементы с $N+1$ по $N+M$ при этом задают комплексные амплитуды сигналов ($\dot{b}_m$) в требуемых точках приема M.

$$\left[\dot{\mathbf{Y}}\right]=\left[\dot{y}_1=0;...;\dot{y}_N=0;\dot{y}_{N+1}=\dot{b}_1;...;\dot{y}_{N+M}=\dot{b}_M\right]^T. \quad (4)$$

Кроме того, предполагается, что элементы вектор-столбца входных сигналов $[\dot{\mathbf{X}}]$ с $N+1$ по $N+M$ равны нулю. Для упрощения принимается, что в рассматриваемых точках отражение отсутствует.

$$\left[\dot{\mathbf{X}}\right]=\left[\dot{x}_1=\dot{a}_1;...;\dot{x}_N=\dot{a}_N;\dot{x}_{N+1}=0;...;\dot{x}_{N+M}=0\right]^T. \quad (5)$$

Матрица S-параметров многополюсника имеет следующий вид [6]:

$$\left[\dot{\mathbf{S}}\right]=\begin{bmatrix} \dot{s}_{1,1} & .. & \dot{s}_{1,N} & \dot{s}_{1,N+1} & .. & \dot{s}_{1,N+M} \\ .. & .. & .. & .. & .. & .. \\ \dot{s}_{N,1} & .. & \dot{s}_{N,N} & \dot{s}_{N,N+1} & .. & \dot{s}_{N,N+M} \\ \dot{s}_{N+1,1} & .. & \dot{s}_{N+1,N} & \dot{s}_{N+1,N+1} & .. & \dot{s}_{N+1,N+M} \\ .. & .. & .. & .. & .. & .. \\ \dot{s}_{N+M,1} & .. & \dot{s}_{N+M,N} & \dot{s}_{N+M,N+1} & .. & \dot{s}_{N+M,N+M} \end{bmatrix}. \quad (6)$$

Диагональные элементы матрицы соответствуют коэффициентам отражения на каждом входе многополюсника. Внедиагональные элементы, в свою очередь, описывают коэффициенты передачи между различными входами многополюсника.

В свободном пространстве коэффициенты передачи определяются следующим выражением [6]:

$$\dot{s}_{i,k}=F_{i,k}\cdot F_{k,i}\cdot r_{i,k}^{-q}\cdot\exp\left(-j\cdot\beta\cdot r_{i,k}\right), \quad (7)$$

где $r_{i,k}$ – расстояние между точками, соответствующими i-му и k-му входам многополюсника; $\beta=\dfrac{2\pi}{\lambda}$ – коэффициент фазы; $F_{i,k}$ – описывает коэффициент усиления по полю антенны между i-м и k-м входами; множитель $r_{i,k}^{-q}$ учитывает уменьшение амплитуды напряженности при передаче от i-го входа к k-му входу. Для ближней зоны принимают $q=2$, для дальней зоны принимают $q=1$.

Для определения расстояния между входами многополюсника используется следующее выражение [6]:

$$r_{i,k}=\sqrt{\left(x_i-x_k\right)^2+\left(y_i-y_k\right)^2+\left(z_i-z_k\right)^2}, \quad (8)$$

где $\left(x_i,y_i,z_i\right)$ – декартовы координаты i-й точки.

Стоит отметить, что для реальной антенны в качестве ее координат необходимо использовать координаты ее фазового центра.

Совокупность входов многополюсника, которые являются излучающими антеннами, можно анализировать как антенную решетку. Соответственно, при вычислении коэффициентов передачи по выражению (7) необходимо определить, в какой зоне поля антенной решетки располагается вход многополюсника, для которого вычисляется коэффициент передачи, чтобы корректно выбрать значение q. Для определения границ ближней зоны используется выражение [7]

$$r_n\le\frac{2\cdot D^2}{\lambda}, \quad (9)$$

где D – размер апертуры антенной решетки.

Условие для определения q можно записать в следующем виде:

$$q=\begin{cases}2, & r_{i,k}\le\dfrac{2\cdot D^2}{\lambda},\\ 1, & r_{i,k}>\dfrac{2\cdot D^2}{\lambda}.\end{cases} \quad (10)$$

В [6] показано, что матрицу S-параметров (6) можно разбить на подматрицы

$$[\dot{\mathbf{S}}]=\begin{bmatrix}[\dot{\mathbf{S}}_{\mathrm{TT}}] & [\dot{\mathbf{S}}_{\mathrm{TR}}]\\ [\dot{\mathbf{S}}_{\mathrm{RT}}] & [\dot{\mathbf{S}}_{\mathrm{RR}}]\end{bmatrix}, \quad (11)$$

где $[\dot{\mathbf{S}}_{\mathrm{TT}}]$ описывает передачу между точками излучения, т.е. от одной антенны к другой и имеет размер $N\times N$; $[\dot{\mathbf{S}}_{\mathrm{TR}}]$ характеризует передачу от точек приема до точек излучения и имеет размер $N\times M$; $[\dot{\mathbf{S}}_{\mathrm{RT}}]$ – характеризует передачу от точек излучения до точек приема и имеет размер $M\times N$; $[\dot{\mathbf{S}}_{\mathrm{RR}}]$ характеризует передачу между точками приема и имеет размер $M\times M$.

На рис. 3 [6] схематично показано, какую роль играют отдельные подматрицы.

В идеальном случае, когда все приемные и передающие антенны согласованы, а взаимное влияние точек передачи друг с другом и точек приема друг с другом отсутствует, выражение (3) принимает вид:

$$[\dot{\mathbf{B}}]=[\dot{\mathbf{S}}_{\mathrm{RT}}]\cdot[\dot{\mathbf{A}}]. \quad (12)$$

Тогда количество уравнений сократится с $N+M$ до M.

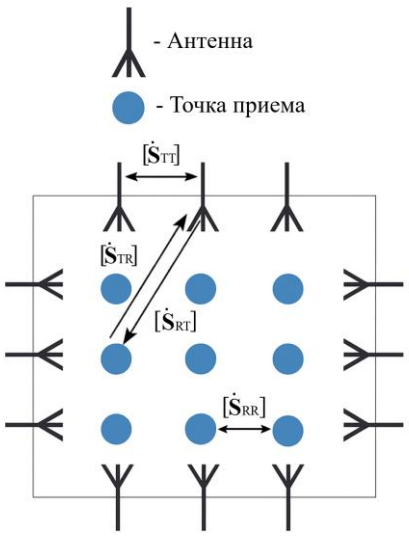


Рис. 3. Иллюстрация роли подматриц [6]

Формирование нескольких минимумов

Если для соответствующей точки приема задать условие $\dot{b}_m=0$, то минимум напряженности электрического поля окажется не одиночным, а будет

повторяться, распространяясь на множество соседних точек пространства. Чтобы минимум амплитуды напряженности электрического поля оставался одиночным, в задачу вводят дополнительные точки поддержки. В этих точках задаются такие значения амплитуды, которые отличаются от минимальных и позволяют ограничить область минимума. Соответственно, требуется задавать вектор-столбец $[\dot{\mathbf{B}}]$ так, чтобы в точках приема, где требуется обеспечить минимум, значение элемента $\dot{b}_m$ было равно нулю, а в заданной точке поддержки значение было больше нуля ($\dot{b}_m > 0$).

Установлено, что количество и расположение точек поддержки относительно точки минимума влияет на качество совпадения требуемого и формируемого распределения напряженности электрического поля. Вопрос определения количества и расположения точек поддержки выходит за рамки настоящей работы и подлежит отдельному изучению.

В работе [6] решалась задача по формированию максимумов амплитуды напряженности электрического поля в требуемых областях пространства. В них решалась система уравнений (12) при помощи метода Гаусса, но в таком случае при введении в систему дополнительных точек поддержки может возникать неустойчивость решения [8–10]. Помимо этого, если количество излучающих антенн *N* больше количества точек приема *M* в системе, то такая система будет являться недоопределенной. Такая недоопределенная система может иметь бесконечное количество решений или не иметь их вовсе. Поэтому для решения системы (12) вводим критериальную постановку задачи: найти вектор-столбец $[\dot{\mathbf{A}}]$, минимизирующий квадратичную невязку приближения требуемого распределения поля.

Чтобы иметь возможность управлять глубиной подавления в заданных точках минимума, нужно ввести дополнительные веса w_m в виде диагональной матрицы

$$[\mathbf{W}] = \mathrm{diag}(w_1, w_2, ..., w_m), \qquad (13)$$

чем больше значение w_m, тем выше приоритет обеспечения условий $\dot{b}_m \approx 0$ и $\dot{b}_m \approx 1$, в зависимости от того требуется обеспечить максимум или минимум амплитуды напряженности электрического поля в требуемой точке приема.

Тогда для определения вектор-столбца $[\dot{\mathbf{A}}]$ необходимо решить следующую задачу определения критерия минимизации невязки – Ψ:

$$\Psi([\dot{\mathbf{A}}]) = \left\| [\mathbf{W}] \cdot ([\dot{\mathbf{S}}_{\mathbf{RT}}] \cdot [\dot{\mathbf{A}}] - [\dot{\mathbf{B}}]) \right\|_2^2, \qquad (14)$$

где $\| \ \|_2$ – L2-норма, $\|x\|_2 = \sqrt{x_1^2 + ... x_n^2}$.

Фактически если ввести математическое определение невязки для введенной системы, то получим

$$[\dot{\boldsymbol{\varepsilon}}] = [\dot{\mathbf{S}}_{\mathrm{RT}}][\dot{\mathbf{A}}] - [\dot{\mathbf{B}}], \qquad (15)$$

в идеальном случае $[\dot{\boldsymbol{\varepsilon}}] = 0$.

Можно записать

$$\Psi([\dot{\mathbf{A}}]) = \left\| [\mathbf{W}] \cdot [\dot{\boldsymbol{\varepsilon}}] \right\|_2^2, \qquad (16)$$

таким образом, определен критерий минимизации невязки в (14) и (16).

Введение в систему дополнительных точек поддержки может привести к неустойчивости решений [8–10]. Задача становится плохо обусловленной. Поэтому в систему (14) введем метод регуляризации Тихонова [11], используемый для решения плохо обусловленных задач. Введение регуляризации Тихонова при $\mu > 0$ позволяет получить единственность решения. Тогда выражение (14) принимает следующий вид:

$$\Psi([\dot{\mathbf{A}}]) = \left\| [\mathbf{W}] \cdot ([\dot{\mathbf{S}}_{\mathrm{RT}}] \cdot [\dot{\mathbf{A}}] - [\dot{\mathbf{B}}]) \right\|_2^2 + \mu \left\| [\dot{\mathbf{A}}] \right\|_2^2, \qquad (17)$$

где μ – коэффициент регуляризации Тихонова.

Раскроем норму через эрмитово сопряжение. Известно [8–10], что: $\|x\|_2^2 = x^H \cdot x$, где H – эрмитово сопряжение.

Тогда первое слагаемое выражения (17) примет вид

$$\left\| [\mathbf{W}][\dot{\boldsymbol{\varepsilon}}] \right\|_2^2 = ([\mathbf{W}][\dot{\boldsymbol{\varepsilon}}])^H ([\mathbf{W}][\dot{\boldsymbol{\varepsilon}}]) = [\dot{\boldsymbol{\varepsilon}}]^H [\mathbf{W}]^H [\mathbf{W}][\dot{\boldsymbol{\varepsilon}}]. \qquad (18)$$

Поскольку матрица весов **[W]** является вещественной и диагональной, то получим

$$[\mathbf{W}]^H [\mathbf{W}] = [\mathbf{W}]^2 = \mathrm{diag}(w_1^2, ..., w_M^2). \qquad (19)$$

Тогда выражение (17) примет вид

$$\Psi([\dot{\mathbf{A}}]) = \left([\dot{\mathbf{S}}_{\mathrm{RT}}] \cdot [\dot{\mathbf{A}}] - [\dot{\mathbf{B}}]\right)^H \cdot [\mathbf{W}]^2 \times$$
$$\times \left([\dot{\mathbf{S}}_{\mathrm{RT}}] \cdot [\dot{\mathbf{A}}] - [\dot{\mathbf{B}}]\right) + \mu [\dot{\mathbf{A}}]^H [\dot{\mathbf{A}}]. \qquad (20)$$

Раскроем скобки в (20)

$$\Psi([\dot{\mathbf{A}}]) = [\dot{\mathbf{A}}]^H [\dot{\mathbf{S}}_{\mathrm{RT}}]^H [\mathbf{W}]^2 [\dot{\mathbf{S}}_{\mathrm{RT}}][\dot{\mathbf{A}}] -$$
$$-[\dot{\mathbf{A}}]^H [\dot{\mathbf{S}}_{\mathrm{RT}}]^H [\mathbf{W}]^2 [\dot{\mathbf{B}}] - [\dot{\mathbf{B}}]^H [\mathbf{W}]^2 [\dot{\mathbf{S}}_{\mathrm{RT}}][\dot{\mathbf{A}}] +$$
$$+[\dot{\mathbf{B}}]^H [\mathbf{W}]^2 [\dot{\mathbf{B}}] + \mu [\dot{\mathbf{A}}]^H [\dot{\mathbf{A}}]. \qquad (21)$$

Согласно [12] считается, что требуется найти комплексный вектор-столбец $[\dot{\mathbf{A}}]$, при котором выражение $\Psi([\dot{\mathbf{A}}])$ не уменьшается при малых изменениях действительной и мнимой части, т.е. находится стационарная точка – глобальный минимум функции

$$\frac{\partial \Psi([\dot{\mathbf{A}}])}{\partial [\dot{\mathbf{A}}]^*} = 0, \qquad (22)$$

где $[\dot{\mathbf{A}}]^*$ – комплексно-сопряженный вектор-столбец $[\dot{\mathbf{A}}]$.

Требуется выражение (21) продифференцировать согласно (22). В результате получаем выражение

$$\frac{\partial \Psi([\dot{\mathbf{A}}])}{\partial [\dot{\mathbf{A}}]^*} = [\dot{\mathbf{S}}_{\mathrm{RT}}]^H [\mathbf{W}]^2 [\dot{\mathbf{S}}_{\mathrm{RT}}][\dot{\mathbf{A}}] -$$
$$-[\dot{\mathbf{S}}_{\mathrm{RT}}]^H [\mathbf{W}]^2 [\dot{\mathbf{B}}] + \mu [\dot{\mathbf{A}}]. \qquad (23)$$

В качестве условия минимума запишем

$$[\dot{\mathbf{S}}_{\mathrm{RT}}]^H [\mathbf{W}]^2 [\dot{\mathbf{S}}_{\mathrm{RT}}][\dot{\mathbf{A}}] - [\dot{\mathbf{S}}_{\mathrm{RT}}]^H [\mathbf{W}]^2 [\dot{\mathbf{B}}] + \mu [\dot{\mathbf{A}}] = 0. \qquad (24)$$

Тогда

$$[\dot{\mathbf{S}}_{\mathrm{RT}}]^H[\mathbf{W}]^2[\dot{\mathbf{S}}_{\mathrm{RT}}][\dot{\mathbf{A}}]+\mu[\dot{\mathbf{A}}]=[\dot{\mathbf{S}}_{\mathrm{RT}}]^H[\mathbf{W}]^2[\dot{\mathbf{B}}]. \quad (25)$$

Преобразуем (25)

$$\left([\dot{\mathbf{S}}_{\mathrm{RT}}]^H[\mathbf{W}]^2[\dot{\mathbf{S}}_{\mathrm{RT}}]+\mu[\mathbf{I}]\right)[\dot{\mathbf{A}}]=[\dot{\mathbf{S}}_{\mathrm{RT}}]^H[\mathbf{W}]^2[\dot{\mathbf{B}}], \quad (26)$$

где $[\mathbf{I}]$ – единичная матрица.

Таким образом, решая систему (26), определяется вектор-столбец излучаемых сигналов $[\dot{\mathbf{A}}]$ для обеспечения требуемого распределения электрического поля в заданной области пространства.

Из (26) можно выразить $[\dot{\mathbf{A}}]$:

$$[\dot{\mathbf{A}}]=\left([\dot{\mathbf{S}}_{\mathrm{RT}}]^H[\mathbf{W}]^2[\dot{\mathbf{S}}_{\mathrm{RT}}]+\mu[\mathbf{I}]\right)^{-1}[\dot{\mathbf{S}}_{\mathrm{RT}}]^H[\mathbf{W}]^2[\dot{\mathbf{B}}]. \quad (27)$$

Проверим полученный математический аппарат с помощью математического моделирования.

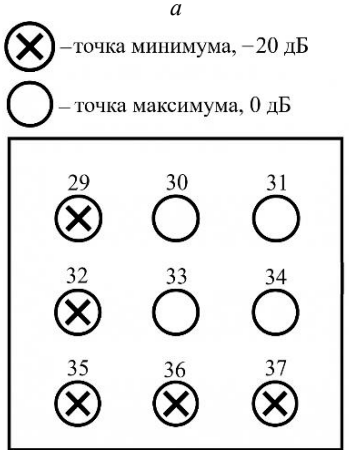
Математическое моделирование

Рассмотрим антенную систему, состоящую из 28 излучателей, которая расположена по периметру области (по 7 антенн на каждой из сторон). Данная система рассчитана на работу с 9 точками приема. Рабочая частота составляет 2,5 ГГц, расстояние между излучателями составляет половину длины волны λ/2. Диаграмма направленности одиночного излучателя принята равной изотропной, все антенны полагаются идеально согласованными. Выбор исходных данных эксперимента обусловлен с аналогичными выборами в работах [13–18].

На рис. 4 представлены требуемые конфигурации распределения напряженности электрического поля в пространстве.

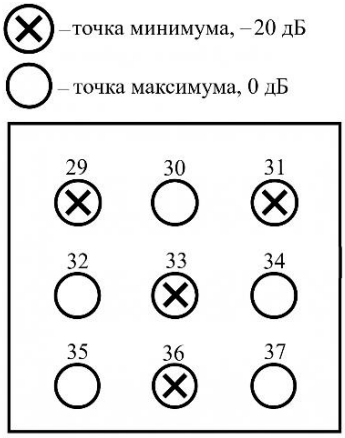


а



б

Рис. 4 (начало)



в

Рис. 4 (окончание). Требуемые конфигурации распределения напряженности электрического поля в пространстве: *а* – минимумы в углах пространства; *б* – антибуква «L»; *в* – антибуква «Y»

В табл. 1 приведены координаты излучателей.

Таблица 1

Координаты излучателей

№	*X*, м	*Y*, м	*Z*, м
1	–0,1800	–0,2100	0
2	–0,1200	–0,2100	0
3	–0,0600	–0,2100	0
4	0	–0,2100	0
5	0,0600	–0,2100	0
6	0,1200	–0,2100	0
7	0,1800	–0,2100	0
8	0,2100	–0,1800	0
9	0,2100	–0,1200	0
10	0,2100	–0,0600	0
11	0,2100	0	0
12	0,2100	0,0600	0
13	0,2100	0,1200	0
14	0,2100	0,1800	0
15	0,1800	0,2100	0
16	0,1200	0,2100	0
17	0,0600	0,2100	0
18	0	0,2100	0
19	–0,0600	0,2100	0
20	–0,1200	0,2100	0
21	–0,1800	0,2100	0
22	–0,2100	0,1800	0
23	–0,2100	0,1200	0
24	–0,2100	0,0600	0
25	–0,2100	0	0
26	–0,2100	–0,0600	0
27	–0,2100	–0,1200	0
28	–0,2100	–0,1800	0

В табл. 2 представлены амплитуды и начальные фазы излучаемых сигналов, рассчитанные по выражению (27) для формирования заданного распределения электрического поля в пространстве.

На рис. 5 представлены результаты математического моделирования распределения электрического поля в пространстве.

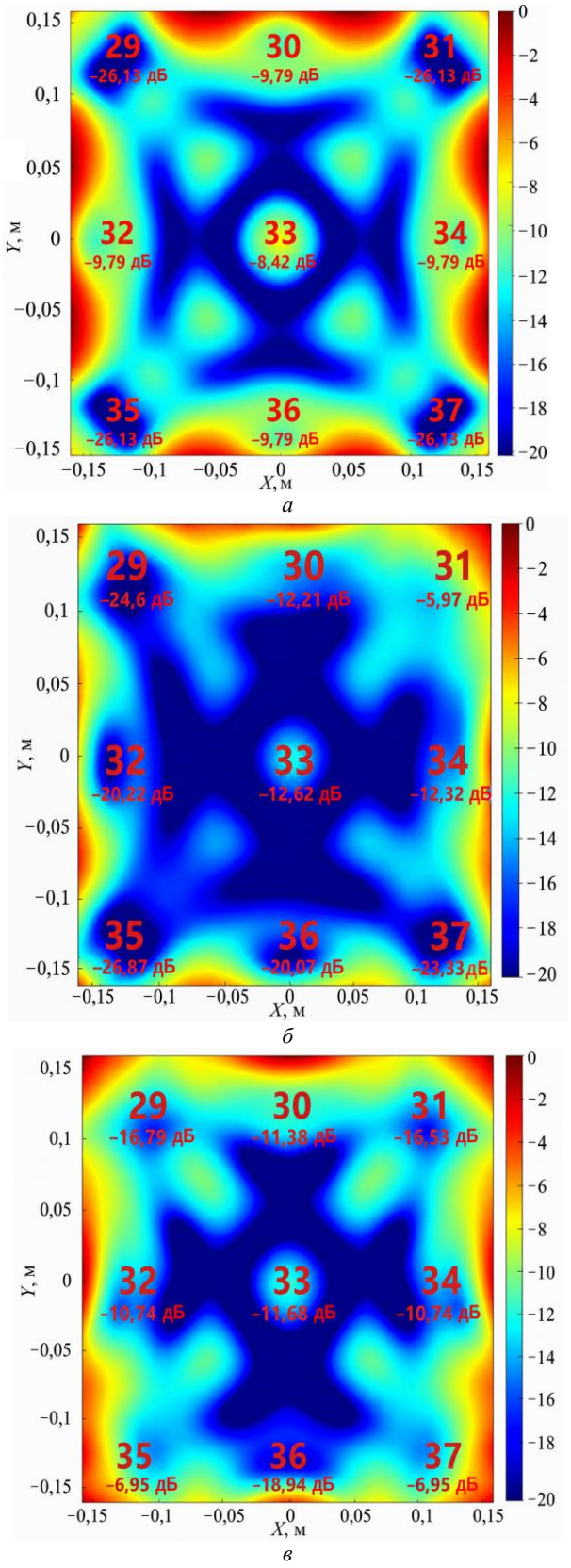
Таблица 2

Амплитуды и начальные фазы излучаемых сигналов

№	Минимумы в углах пространства		Антибуква «L»		Антибуква «Y»	
	Амплитуда, В/м	Фаза, град	Амплитуда, В/м	Фаза, град	Амплитуда, В/м	Фаза, град
1	6,7	179	6,7	–166	10	–142
2	5,7	–66	4,9	2	3,7	–42
3	14,9	–24	7	8	2,9	–14
4	16,1	–61	8,5	–40	8,5	–55
5	14,9	–24	6,7	–14	2,9	–14
6	5,7	–66	2,6	–51	3,7	–42
7	6,6	179	8,3	–149	10	–142
8	6,6	179	5,8	–155	7,8	–136
9	5,7	–66	5,6	–3	7,2	–20
10	14,9	–24	8,2	–9	4,2	–6
11	16,1	–61	11,6	–44	11,1	–49
12	14,9	–24	5,3	–30	6,9	–30
13	5,7	–66	5,56	–48	3,4	–70
14	6,6	179	11,1	–134	9,1	–140
15	6,6	179	11,1	–134	7,7	–141
16	5,7	–66	5,56	–48	3,6	–42
17	14,9	–24	5,3	–30	5	–28
18	16,1	–61	11,6	–44	13	–46
19	14,9	–24	8,2	–9	5	–28
20	5,7	–66	5,6	–3	3,6	–42
21	6,6	179	5,8	–155	7,7	–141
22	6,6	179	8,3	–149	9,1	–140
23	5,7	–66	2,6	–51	3,4	–70
24	14,9	–24	6,7	–14	6,9	–30
25	16,1	–61	8,5	–40	11,1	–49
26	14,9	–24	7	8	4,2	–6
27	5,7	–66	4,8	2	7,2	–20
28	6,6	179	6,7	–166	7,8	–136

Как можно наблюдать из результатов, показанных на рис. 5 и в табл. 3, предложенный метод определения амплитуд и начальных фаз сигналов позволяет обеспечить минимум амплитуды напряженности электрического поля в нескольких требуемых точках пространства, а также обеспечить разницу между точкой минимума и максимума в худшем случае в 5,41 дБ, а в лучшем случае – 17,71 дБ, что говорит о потенциальной возможности применения предложенного метода на практике.

Также стоит прокомментировать разницу по нормированным значениям амплитуды напряженности электрического поля между целевыми значениями и результатами моделирования. Стоит отметить, что разница по целевым показателям в лучшем случае составила 6 дБ, в худшем – 9,3 дБ. Эта разница определяется плохой обусловленностью задачи: имеется множество решений для выражения (27) для разных значений коэффициента μ, характеризующихся одинаковой или близкой величиной квадратичной невязки. Часть требований (значений напряженности электрического поля в разных точках) оказывается взаимно ограничивающей, и качество полученного решения определяется минимально возможной величиной квадратичной невязки между целевым и результирующим распределением поля в пространстве. Кроме того, некоторые получаемые решения хоть и характеризуется низкой величиной невязки и обеспечивают распределение напряженности электрического поля, близкое к заданному, обладают физически нереализуемыми значениями комплексных амплитуд излучаемых сигналов.



а

б

в

Рис. 5. Распределения электрического поля:
а – минимумы в углах пространства,
б – антибуква «L», *в* – антибуква «Y»

Таблица 3

Значения напряженности электрического поля в точках приема

№ точки приема	Минимумы в углах пространства, дБ	Антибуква «L», дБ	Антибуква «Y», дБ
29	**–26,13** (точка минимума)	**–24,6** (точка минимума)	**–16,79** (точка минимума)
30	–9,79	–12,21	–11,38
31	**–26,13** (точка минимума)	–5,97	**–16,53** (точка минимума)
32	–9,79	**–20,22** (точка минимума)	–10,74
33	–8,42	–12,62	–11,68
34	–9,79	–12,32	–10,74
35	**–26,13** (точка минимума)	**–26,87** (точка минимума)	–6,95
36	–9,79	**–20,07** (точка минимума)	**–18,94 (**точка минимума)
37	**–26,13** (точка минимума)	**–23,33** (точка минимума)	–6,95

Стоит отметить, что рассматриваемый квадратичный критерий не является единственно возможным. При развитии данного направления исследования целесообразно рассмотреть иные способы минимизации невязки.

Заключение

В работе рассмотрен подход, позволяющий определить комплексные амплитуды излучаемых сигналов, при которых в заранее выбранных областях пространства формируются несколько минимумов амплитуды напряженности электрического поля. Связью между комплексными амплитудами излучаемых сигналов и комплексными амплитудами в точках приема является матрица S-параметров. Вычисление комплексных амплитуд излучаемых сигналов в работе предложено решать при помощи решения СЛАУ. Полученная в работе СЛАУ может являться недоопределенной системой, что может давать бесконечное количество решений или неустойчивость решений, поэтому задача была сведена к минимизации квадратичной невязки между требуемым и получаемым распределениями электрического поля в пространстве. Также в систему была введена регуляризация Тихонова, которая обеспечивает единственность решения у решаемой СЛАУ. Результатом решения задачи оптимизации при введенной квадратичной мере невязки является решение новой СЛАУ.

Были предложены три конфигурации электрического поля, для каждой из конфигураций были вычислены амплитуды и начальные фазы сигналов. Математическое моделирование распределения электрического поля с использованием полученных амплитуд и начальных фаз сигналов показало работоспособность предложенного метода.

Литература

1. Voudoukis N. Performance analysis, characteristics, and simulation of digital QAM // European Journal of Electrical Engineering and Computer Science. – 2017. – Vol. 1, No. 1. – P. 1–7.

2. Svensson A. An introduction to adaptive QAM modulation schemes for known and predicted channels // Proceedings of the IEEE. – 2007. – Vol. 95, No. 12. – P. 2322–2336.

3. Hansen R. Focal region characteristics of focused array antennas // IEEE Transactions on Antennas and Propagation. – 2003. – Vol. 33, No. 12. – P. 1328–1337.

4. Iuzvik D. Formation of the maximum amplitude of the electric field strength using two one-dimensional linear antenna arrays located orthogonal to each other / D. Iuzvik, M. Stepanov // Journal of Electromagnetic Waves and Applications. – 2024. – Vol. 38, No. 18. – P. 2106–2123.

5. Iuzvik D. Ensuring the maximum amplitude of the electric field strength in given coordinates using a linear antenna array focused on a finite distance / D. Iuzvik, M. Stepanov // Journal of Electromagnetic Waves and Applications. – 2024. – Vol. 38, No. 4. – P. 411–427.

6. Iuzvik D. Focusing of the electromagnetic field in several given areas of space / D. Iuzvik, M. Stepanov // Progress in Electromagnetics Research M. – 2022. – Vol. 113. – P. 11–22.

7. Skolnik M.I. Radar Handbook. – 3rd ed. – New York, NY, USA: McGraw-Hill, 2008. – 1328 p.

8. Strang G. Linear Algebra and Its Applications. – 4th ed. – Belmont: Thomson Brooks/Cole, 2006. – 487 p.

9. Гантмахер Ф.Р. Теория матриц: учеб. пособие. – 5-е изд. – М.: ФИЗМАТЛИТ, 2010. – 560 с.

10. Ben-Israel A. Generalized Inverses: Theory and Applications / A. Ben-Israel, T.N.E. Greville. – New York, NY, USA: Springer, 2003. – 420 p.

11. Tikhonov A.N. Solutions of Ill-Posed Problems / A.N. Tikhonov, V.Y. Arsenin. – New York: Winston, 1977. – 258 p.

12. Brandwood D.H. A complex gradient operator and its application in adaptive array theory // IEE Proceedings F (Communications, Radar and Signal Processing). – 1983. – Vol. 130, No. 1. – P. 11–16.

13. Energy-efficient near-field beamforming: A review on practical channel models / H. Ni et al. // Energies. – 2025. – Vol. 18, No. 11. – Art. no. 2966.

14. Gaussian approach for the synthesis of periodic and aperiodic antenna arrays: Method review and design guidelines / G. Buttazzoni et al. // Sensors. – 2021. – Vol. 21, No. 7. – Art. no. 2343.

15. Yang S.J. An accurate near-field focusing of array antenna based on near-field active element pattern and infinitesimal dipole modeling / S.J. Yang, Y.D. Kim // IEEE Access. – 2021. – Vol. 9. – P. 143771–143781.

16. Fault diagnosis of realistic arrays from a reduced number of phaseless near-field measurements / R. Palmeri et al. // IEEE Transactions on Antennas and Propagation. – 2023. – Vol. 71, No. 9. – P. 7206–7219.

17. A review on different techniques of mutual coupling reduction between elements of any MIMO antenna. – Part 2: Metamaterials and many more / A. Kumar, A.Q. Ansari, B.K. Kanaujia, J. Kishor, L. Matekovits // Radio Science. – 2021. – Vol. 56, No. 3. – P. 1–22.

18. Mandal S. Mutual coupling reduction in a patch antenna array based on planar frequency selective surface structure / S. Mandal, C.K. Ghosh // Radio Science. – 2022. – Vol. 57. – Art. no. e2021RS007392.

Юзвик Денис Андреевич
Канд. техн. наук, инженер каф. РПиРПУ
Новосибирского гос. технического университета (НГТУ)
Карла Маркса пр-т, 20, г. Новосибирск, Россия, 630073
ORCID: 0000-0002-8365-1639
Тел.: +7 (383-3) 46-15-46
Эл. почта: zetanicestar@gmail.com

Степанов Максим Андреевич
Д-р техн. наук, зав. каф. РПиРПУ НГТУ
Карла Маркса пр-т, 20, г. Новосибирск, Россия, 630073
ORCID: 0000-0001-6095-1172
Тел.: +7 (383-3) 46-15-46
Эл. почта: m.stepanov@corp.nstu.ru

Поступила в редакцию: 19.03.2026.
Принята к публикации: 26.06.2026.

Iuzvik D.A., Stepanov M.A.

A method for forming several minima of the electromagnetic field at required points in space

Relevance. The article considers a method for forming several electromagnetic field minima in the required spatial regions in order to reduce the level of interference for sensitive equipment. ***Purpose of the study***. Developing a method for calculating the complex amplitudes of emitted signals to form several specified regions of reduced electric field amplitude while simultaneously setting field levels at other reception points using Tikhonov regularization. ***Methods***. A system of linear algebraic equations was balanced by introducing Tikhonov regularization, ensuring the uniqueness of the solution. An optimization problem was also formulated to minimize the quadratic residual between the required and obtained electric field distributions. The theoretical results were verified using mathematical modeling. ***Scientific novelty***. The relationship between complex amplitudes at reception points and complex amplitudes of emitted signals was exploited via the S^{-} parameter matrix. It is noted that calculating the complex amplitudes of emitted may lead to an underdetermined system of linear algebraic equations, resulting in unstable solutions: there may be an infinite number of them, or there may be no solution at all. ***Results***. Three electric field configurations are proposed, and for each of them the complex amplitudes of the emitted signals are determined. The results of mathematical modeling confirmed the obtained theoretical findings. ***Practical significance***. The obtained results can be used to generate the required electromagnetic field distributions with several minima in specified regions of space.

Keywords: antenna arrays, optimization, system of linear equations, electromagnetic field, electric field minima.
DOI: 10.21293/1818-0442-2026-29-2-16-23

References

1. Voudoukis N. Performance analysis, characteristics, and simulation of digital QAM. European Journal of Electrical Engineering and Computer Science, 2017, vol. 1, no. 1, pp. 1–7.

2. Svensson A. An introduction to adaptive QAM modulation schemes for known and predicted channels. Proceedings of the IEEE, 2007, vol. 95, no. 12, pp. 2322–2336.

3. Hansen R. Focal region characteristics of focused array antennas. IEEE Transactions on Antennas and Propagation, 2003, vol. 33, no. 12, pp. 1328–1337.

4. Iuzvik D., Stepanov M. Formation of the maximum amplitude of the electric field strength using two one-dimensional linear antenna arrays located orthogonal to each other. Journal of Electromagnetic Waves and Applications, 2024, vol. 38, no. 18, pp. 2106–2123.

5. Iuzvik D., Stepanov M. Ensuring the maximum amplitude of the electric field strength in given coordinates using a linear antenna array focused on a finite distance. Journal of Electromagnetic Waves and Applications, 2024, vol. 38, no. 4, pp. 411–427.

6. Iuzvik D., Stepanov M. Focusing of the electromagnetic field in several given areas of space. Progress in Electromagnetics Research M, 2022, vol. 113, pp. 11–22.

7. Skolnik M.I. Radar handbook. 3rd ed. New York, McGraw-Hill, 2008, 1328 p.

8. Strang G. Linear algebra and its applications. 4th ed. Belmont, Thomson Brooks/Cole, 2006, 487 p.

9. Gantmakher F.R. *Teoriya matrits* [The theory of matrices], textbook, M., FIZMATLIT, 2010, 560 p. (in Russ.).

10. Ben-Israel A., Greville T.N.E. Generalized inverses: theory and applications. 2nd ed. New York, Springer, 2003, 420 p.

11. Tikhonov A.N., Arsenin V.Y. Solutions of ill-posed problems. Washington, V.H. Winston & Sons, 1977, 258 p.

12. Brandwood D.H. A complex gradient operator and its application in adaptive array theory. IEE Proceedings F – Communications, Radar and Signal Processing, 1983, vol. 130, no. 1, pp. 11–16.

13. Ni H. et al. Energy-efficient near-field beamforming: a review on practical channel models. Energies, 2025, vol. 18, no. 11, art. no. 2966.

14. Buttazzoni G. et al. Gaussian approach for the synthesis of periodic and aperiodic antenna arrays: method review and design guidelines. Sensors, 2021, vol. 21, no. 7, art. no. 2343.

15. Yang S.J., Kim Y.D. An accurate near-field focusing of array antenna based on near-field active element pattern and infinitesimal dipole modeling. IEEE Access, 2021, vol. 9, pp. 143771–143781.

16. Palmeri R. et al. Fault diagnosis of realistic arrays from a reduced number of phaseless near-field measurements. IEEE Transactions on Antennas and Propagation, 2023, vol. 71, no. 9, pp. 7206–7219.

17. Kumar A., Ansari A.Q., Kanaujia B.K., Kishor J., Matekovits L. A review on different techniques of mutual coupling reduction between elements of any MIMO antenna. Part 2: Metamaterials and many more. Radio Science, 2021, vol. 56, no. 3, pp. 1–22.

18. Mandal S., Ghosh C.K. Mutual coupling reduction in a patch antenna array based on planar frequency selective surface structure. Radio Science, 2022, vol. 57, art. no. e2021RS007392.

Denis A. Iuzvik
Candidate of Technical Sciences, Engineer,
Department of Radio receiving and Radio transmitting devices,
Novosibirsk State Technical University
20, Karl Marx av., Novosibirsk, Russia, 630073
ORCID: 0000-0002-8365-1639
Phone: +7 (383-3) 46-15-46
Email: zetanicestar@gmail.com

Maksim A. Stepanov
Doctor of Technical Sciences, Head of Department
of Radio receiving and Radio transmitting devices,
Novosibirsk State Technical University
20, Karl Marx av., Novosibirsk, Russia, 630073
ORCID: 0000-0001-6095-1172
Phone: +7 (383-3) 46-15-46
Email: m.stepanov@corp.nstu.ru

Received: 19.03.2026.
Accepted: 26.06.2026.

УДК 621.396.41
Шифр специальности ВАК: 2.2.14

М.С. Мурманский, А.Е. Томских, Е.С. Жечев, Р.С. Суровцев, И.А. Скорняков

Способы подавления перекрестных помех в связанных линиях передачи: обзор

Актуальность. С развитием высокоскоростной электроники и увеличением плотности монтажа печатных плат проблема перекрестных помех становится одним из главных препятствий для обеспечения качества передачи сигнала. ***Цель исследования.*** Систематизация и обзор современных конструктивных способов к минимизации перекрестных помех в микрополосковых линиях передачи для выявления наиболее эффективных стратегий проектирования. ***Методы.*** В работе применены методы анализа научной литературы, сравнительное сопоставление физических механизмов подавления помех, а также оценка количественных характеристик различных топологических решений. ***Научная новизна.*** Предложена обобщенная классификация способов защиты от перекрестных помех, основанная на физическом механизме подавления (компенсация разности коэффициентов связи, отвод энергии в опорный проводник, локальная коррекция погонных параметров), а также проведен анализ эффективности гибридных и высокотехнологичных структур. ***Результаты.*** Выявлено, что классические правила трассировки теряют эффективность на частотах выше 10 ГГц, в то время как инновационные решения на базе периодических структур и метаповерхностей обеспечивают подавление наводок до минус 50 дБ. Установлены ключевые закономерности влияния дефектов опорного слоя и пассивного проводника на волновое сопротивление и фазовые скорости мод, а также определены границы применимости каждого из рассмотренных методов. ***Практическая значимость.*** Результаты обзора служат техническим руководством для инженеров-разработчиков СВЧ-аппаратуры, позволяя обоснованно выбирать конкретную топологию межсоединений в зависимости от требуемого уровня помехозащищенности, доступной площади платы и сложности производственного цикла.
Ключевые слова: перекрестные помехи, связанные линии передачи, печатные платы, целостность сигнала, защитные проводники, дефектные структуры земли, метаматериалы, метаповерхности.
DOI: 10.21293/1818-0442-2026-29-2-24-38

Введение

Современная тенденция развития инфокоммуникационных систем характеризуется освоением новых частотных диапазонов и переходом к сверхвысоким скоростям передачи данных. Реализация этих задач требует минимизации массогабаритных характеристик устройств при одновременном повышении их функциональной плотности, что обусловливает высокую плотность монтажа компонентов на современных печатных платах (ПП). Различные типы линий передачи, в частности, микрополосковые линии (МПЛ), выступают в качестве элементов межсоединений, связывающих функциональные узлы ПП. Однако уменьшение расстояний между активным и пассивным проводником в связанной МПЛ неизбежно приводит к электромагнитному взаимодействию, из-за которого возникают перекрестные помехи [1].

Феномен перекрестных помех определяется как нежелательное наведение сигнала с активного проводника связанной линии на смежный пассивный проводник вследствие распределенной взаимной емкостной и индуктивной связи между ними [2, 3]. Данное явление вызывает существенное ухудшение целостности сигналов: искажение волнового фронта, флуктуации задержки распространения, сужение раскрыва глазковой диаграммы и рост уровня детерминированного и случайного шумов [4].

В теории линий передачи принято дифференцировать перекрестные наводки на ближнем конце (ПНБК, от англ. Near-End CrossTalk – NEXT) и на дальнем конце (ПНДК, от англ. Far-End CrossTalk – FEXT) пассивного (по отношению к источнику помех) проводника. Амплитуда ПНБК (NEXT) в значительной степени определяется условиями согласования в портах и не зависит от длины связанной МПЛ [5]. Уровень помехи FEXT, напротив, демонстрирует линейную зависимость от длины линии [6]. Физическая природа возникновения FEXT обусловлена модовой дисперсией – разностью фазовых скоростей распространения четной и нечетной мод. Она возникает из-за неоднородности диэлектрического заполнения пространства вокруг проводников (сочетания материала подложки и воздушной среды), что приводит к различию эффективных диэлектрических проницаемостей для разных мод [7]. На рис. 1 представлена эквивалентная схема включения связанной МПЛ с перекрестными помехами, где R – сопротивление нагрузок, E – источник ЭДС; l – длина связанной МПЛ; *1–4* – узлы связанной МПЛ; *1–3* – узлы на ближнем конце; *2–4* – на дальнем конце.

Для количественной оценки помеховых воздействий используются амплитуды напряжений на ближнем (V_{NEXT}) и дальнем (V_{FEXT}) концах пассивного проводника.

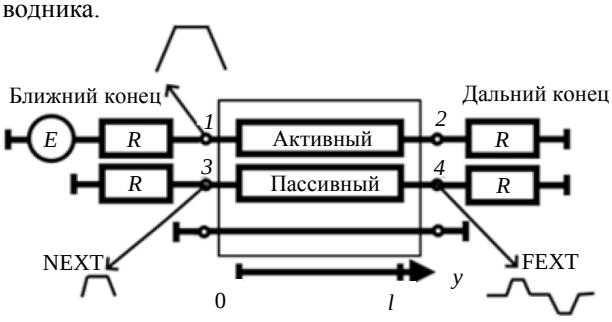


Рис. 1. Эквивалентная схема включения связанной МПЛ с условным отображением перекрёстных помех

Математические выражения для их расчета, ставшие фундаментальными после классической работы [7], позволяют связать указанные величины с временем нарастания сигнала, длиной области связи МПЛ и коэффициентами электромагнитной связи. В [8] показано, что итоговая форма наводок определяется совместным влиянием коэффициентов k_C и k_L: в случае для V_{NEXT} емкостная и индуктивная составляющие складываются, а для V_{FEXT} вычитаются, формируя специфический импульс FEXT, амплитуда которого пропорциональна длине пробега волны и скорости нарастания фронта [7, 9–11]:

$$V_{\text{NEXT}} = \frac{1}{4}(k_C + k_L)V_{\text{вх}}, \quad (1)$$

$$V_{\text{FEXT}} = \frac{1}{2}(k_C - k_L)\left(\frac{l}{\nu}\right)\left(\frac{V_{\text{вх}}}{t_{\text{н}}}\right), \quad (2)$$

где $V_{\text{вх}}$ – амплитуда сигнала на входе активного проводника, l – длина связанной МПЛ, ν – скорость распространения сигнала, $t_{\text{н}}$ – время нарастания сигнала, а k_C и k_L – коэффициенты ёмкостной и индуктивной связи, которые вычисляются как [12, 13]

$$k_C = \frac{|C_{12}|}{\sqrt{C_{11}C_{22}}}, \quad (3)$$

$$k_L = \frac{L_{12}}{\sqrt{L_{11}L_{22}}}, \quad (4)$$

где $L_{11}, L_{12}, L_{22}, C_{11}, C_{12}, C_{22}$ – элементы матриц погонных коэффициентов электромагнитной (**L**) и электростатической (**C**) индукции. В данном случае элементы L_{22} и C_{22} добавлены для учета межлинейной асимметрии [14, 15].

С точки зрения теории сверхвысоких частот (СВЧ) в квази-ТЕМ-приближении условие полной теоретической компенсации V_{FEXT} в связанных МПЛ без потерь соответствует достижению равенства коэффициентов индуктивной и емкостной связи ($k_L = k_C$). В микрополосковых структурах, характеризующихся неоднородным диэлектрическим заполнением, данное условие нарушается ($k_L \neq k_C$), что эквивалентно возникновению разности фазовых скоростей четной и нечетной мод.

В работах [8, 16] доказано, что при прохождении цифрового сигнала через связанные МПЛ в неоднородной среде происходит физическое расщепление исходного импульса на модовые составляющие. Поскольку четная и нечетная моды распространяются с разной скоростью, на временном интервале это проявляется в виде искажения формы сигнала и становится причиной появления перекрестных помех. Итоговая форма FEXT определяется не только разностью коэффициентов связи, но и накопленным временным рассогласованием между модами.

К традиционным и применяемым на практике способам уменьшения перекрестных помех относят [9]:

1. Соблюдение правила «разноса проводников» $s = 3–5w$ (минимальное расстояние между соседними проводниками должно составлять 3–5 ширины проводника w).

2. Введение «защитного» проводника между активным и пассивным проводниками линии [9, 11].

3. Применение диэлектрического заполнения платы за счёт нанесения на её поверхность диэлектрического покрытия оптимальной толщины [17].

Правило $s = 3 – 5w$ уменьшает k_C и k_L за счёт пространственного разноса проводников, что приводит к пропорциональному снижению как V_{NEXT}, так и V_{FEXT}. Принцип диэлектрического заполнения схож с правилом пространственного разноса.

Заземлённый проводник (обычно с переходными отверстиями на опорный слой через каждые λ/20 или чаще) «отбирает» значительную часть энергии с активного проводника, отводя его в опорный слой. В результате лишь малая доля энергии достигает пассивного проводника, что обеспечивает эффективное ослабление как V_{NEXT}, так и V_{FEXT}.

В работе [18] сравнены три конфигурации линий передачи: с применением правила $3w$, защитного проводника и диэлектрического заполнения. Показано, что для коротких линий эффективны линии с защитным проводником, тогда как для длинных – линии с диэлектрическим заполнением.

Несмотря на широкое применение традиционных решений, известно множество других эффективных способов подавления перекрестных помех. Поэтому целью данной работы является систематизация и сравнительный анализ способов подавления перекрестных помех в связанных линиях передачи.

Анализ публикационной активности в области подавления перекрестных помех

Для оценки актуальности проблемы перекрестных помех проведён анализ публикационной активности членов международного сообщества IEEE за период 2020–2025 гг. Поиск осуществлялся по ключевым словам (и их комбинациям): crosstalk, transmission line, microstrip line, PCB, coupled lines в IEEE Xplore. На рис. 2 приведены гистограммы числа опубликованных работ в журналах и трудах конференций IEEE.

Из анализа следует, что ежегодно публикуется не менее 10 работ в высокорейтинговых журналах IEEE и существенно больше – в трудах конференций. Общий объём публикаций за период составил 362 статьи, из которых 234 – в трудах конференций, а 128 – в журналах.

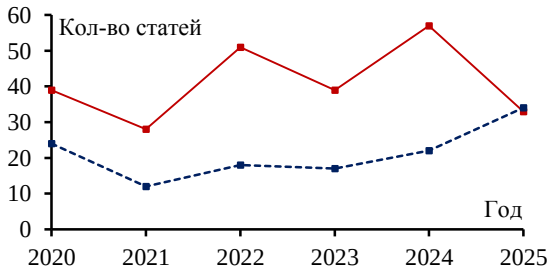


Рис. 2. Количество работ за 5 лет в издательстве IEEE: конференции (–), журналы (--)

Среди журналов лидерами по количеству опубликованных работ являются:

– IEEE Transactions on Electromagnetic Compatibility – 25 работ;

– IEEE Transactions on Components, Packaging and Manufacturing Technology – 11 работ;

– IEEE Transactions on Signal and Power Integrity – 9 работ;

– IEEE Transactions on Microwave Theory and Techniques – 8 работ.

По количеству докладов в трудах конференций наиболее релевантными являются:

– Asia-Pacific International Symposium on Electromagnetic Compatibility – 9 работ;

– IEEE International Symposium on Electromagnetic Compatibility & Signal/Power Integrity – 7 работ;

– IEEE Joint International Symposium on Electromagnetic Compatibility, Signal & Power Integrity: EMC Japan – 4 работы;

– International Conference on Electronic Packaging Technology – 3 работы.

Приведённые результаты показывают устойчивый интерес научного сообщества к проблеме перекрестных помех. Это говорит об актуальности исследования способов минимизации перекрестных помех в межсоединениях ПП для обеспечения целостности сигналов высокоскоростных устройств.

По результатам проведенного анализа можно выделить шесть основных групп способов подавления перекрестных помех (рис. 3). Часть способов (например, встречно-штыревые структуры) направлена на увеличение ёмкостной связи между проводниками для компенсации преобладающей индуктивной связи, что приводит к минимизации V_{FEXT} или даже к полной её компенсации. «Защитные» проводники и переходные структуры направлены на минимизацию наведения энергии на пассивный проводник за счёт отвода энергии в опорный проводник. Дефектные структуры в опорном слое и вырезы в пассивном проводнике позволяют локально корректировать коэффициенты связи, компенсируя обе компоненты, как V_{NEXT}, так и V_{FEXT}. Рассмотрим каждый из способов более детально.

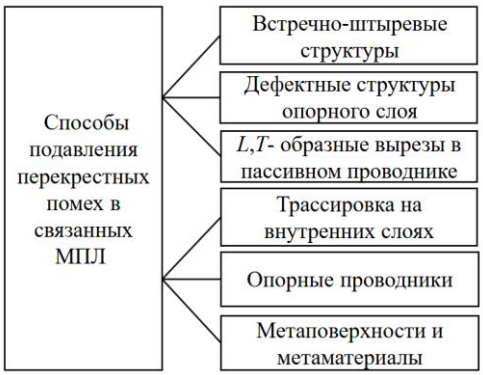


Рис. 3. Структурная схема конструктивных способов подавления перекрестных помех в связанных МПЛ

Трассировка МПЛ на внутренних слоях ПП

Одним из простых и эффективных способов подавления V_{FEXT} является трассировка связанных МПЛ на внутренних слоях ПП, где проводники полностью окружены диэлектриком и двумя опорными плоскостями. В отличие от линий на внешних слоях, полосковые линии в однородном диэлектрике теоретически не имеют перекрестных помех (за счёт равенства фазовых компонент сигнала). Однако на практике V_{FEXT} всегда присутствует и может достигать значительных уровней из-за неоднородности самого диэлектрика.

Например, в работе [19] предложена модель полосковой линии, состоящая из 3 слоёв:

1. Ядро – слой основы с фиксированной толщиной.
2. Препрег – армированный стеклотканью слой для ламинации.
3. Смола – тонкий слой чистой эпоксидной смолы, образующийся между слоями «ядро» и «препрег» при прессовании.

На рис. 4 показана модель поперечного сечения связанной МПЛ на внутреннем слое ПП.

Разница в толщине слоёв (ядра и препрега) приводит к асимметрии диэлектрической среды и росту V_{FEXT}. Даже незначительная разница между геометрическими и электрофизическими параметрами слоёв (особенно из-за разного содержания стекловолокна и смолы) вызывает дисперсию мод и значительный рост V_{FEXT}. Слой смолы часто имеет относительную диэлетрическую проницаемость, отличающуюся от слоёв «ядро» «и препрега» (обычно она выше из-за отсутствия стекла), что усиливает этот эффект.

В работе [20] показано, что даже в однородном диэлектрике отражения от несогласованных нагрузок приводят к появлению FEXT, амплитуда и полярность которой зависят от коэффициентов отражения на ближнем и дальнем концах. А также предложено эмпирическое правило определения полярности V_{FEXT} в многослойных структурах с учётом свойств ядра и препрега: если $h_{препрег} > h_{ядро}$ и $\varepsilon_{rпрепрег} < \varepsilon_{rядро}$, то полярность V_{FEXT} будет отрицательной, а если $\varepsilon_{rпрепрег} > \varepsilon_{rядро}$, то положительной. Основная рекомендация авторов – стремиться к равенству толщин ($h_{препрег} = h_{ядро}$) и близости их ε_r.

В работе [21] для более точного учёта реальной структуры ПП (особенно с армированным стекловолокном) предложена пятислойная модель поперечного сечения линии. Эта детализация позволяет точнее извлекать эффективные диэлектрические проницаемости и прогнозировать V_{FEXT} с высокой точностью.

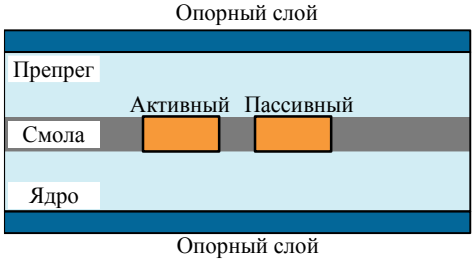


Рис. 4. Модель поперечного сечения связанной МПЛ на внутреннем слое ПП

Практические рекомендации авторов во многом совпадают с выводами из [17, 19]: контроль содержания смолы, равенство толщин диэлектрических

слоев, выбор материалов с близкими ε_r и обязательное включение смолы в модель поперечного сечения.

В [22] рассмотрен технологический аспект: даже при одинаковой конфигурации поперечного сечения образцов различие V_{FEXT} для разных ПП может превышать 10%. Это напрямую связано с разбросом эффективной диэлектрической проницаемости слоя «препрег», вызванным неравномерным распределением стекловолокна и смолы. Предложена концепция критического процента содержания смолы, при котором отклонение V_{FEXT} минимально.

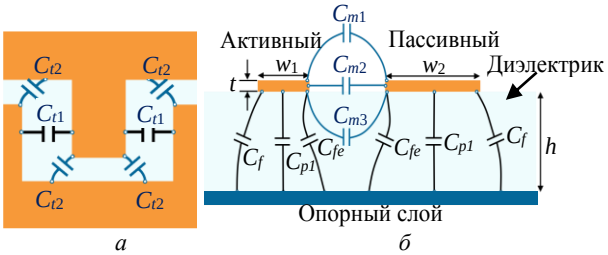
В работе [23] Т.Р. Газизова показано, что полной компенсации перекрестных помех в двухслойном диэлектрике можно достичь изменением геометрических параметров поперечного сечения. Помимо работ, посвященных уравниванию толщин диэлектрических слоев, известен патент [24], в котором компенсация как V_{NEXT}, так и V_{FEXT} достигается за счет «перекрещивания слоев». Дифференциальная пара трассируется на двух смежных слоях так, что сначала положительный проводник верхнего слоя располагается строго над положительным проводником нижнего, а отрицательный – над отрицательным. На половине длины линии проводники нижнего слоя меняются местами, и положительный проводник верхнего слоя оказывается над отрицательным проводником нижнего (и наоборот). Амплитуды помех, наводимые до и после перестановки, имеют противоположную полярность и взаимно компенсируются.

Таким образом, в работах [19–24] выделены следующие ключевые причины, влияющие на V_{FEXT} в полосковых линиях:

1. Несогласованность линии с сопротивлениями на концах как самостоятельная причина появления V_{FEXT} даже в однородном диэлектрике.
2. Различие h и ε_r разных слоёв стека ПП.
3. Влияние «паразитного» слоя смолы на эффективной диэлектрической проницаемости.
4. Влияние содержания смолы и стекловолокна слоя «препрег» на V_{FEXT} при производстве.
5. Необходимость моделей многослойных ПП для точного моделирования.

Применение встречно-штыревых структур

Одним из эффективных способов подавления V_{FEXT} в МПЛ является применение встречно-штыревых структур. Он основан на искусственном увеличении ёмкостной связи между проводниками для компенсации преобладающей в МПЛ индуктивной связи. Вид сверху встречно-штыревой структуры связанной линии и распределение собственных и взаимных ёмкостей на компоненты показан на рис. 5.



а *б*

Рис. 5. Вид сверху на встречно-штыревую структуру (*а*) и компоненты взаимных и собственных ёмкостей (*б*)

В обычной связанной МПЛ C_{11} и C_{21} выражаются как сумма компонент

$$C_{11} = C_{fe} + C_{p1} + C_f, \quad (5)$$

$$C_{21} = C_{m1} + C_{m2} + C_{m3}, \quad (6)$$

где C_f, C_{fe}, C_{m1}, C_{m3} – краевые емкости, а C_{p1}, C_{p2}, C_{m2} – взаимные.

В связанной встречно-штыревой структуре возникают дополнительные взаимные ёмкости: краевая ёмкость между штырём и проводником (C_{t2}) и ёмкость между соседними штырями C_{t1} [25]. В результате (6) преобразуется к виду

$$C_{21} = C_{m1} + C_{m3} + C_{t1} + C_{t2}. \quad (7)$$

Выражение (7) учитывает как базовые компоненты ёмкостной связи, так и дополнительные краевые и межштыревые ёмкости, что дает рост k_C. Важно отметить, что собственная ёмкость C_{11} также возрастает из-за увеличения эффективной ширины проводников, однако при оптимальном выборе параметров рост C_{21} оказывается доминирующим.

Для демонстрации этого введем выражения для вычисления коэффициента FEXT из (2):

$$k_{FEXT} = \frac{1}{2\nu}(k_C - k_L), \quad (8)$$

$$\nu = \frac{\nu_c}{\sqrt{\varepsilon_{r\text{эфф}}}}, \quad (9)$$

$$\varepsilon_{r\text{эфф}} = (0{,}64\varepsilon_r + 0{,}36), \quad (10)$$

где ν – скорость распространения сигнала; $\varepsilon_{r\text{эфф}}$ – эффективная диэлектрическая проницаемость; ν_c – скорость света.

Из (8) видно, что при неизменной индуктивной связи (индуктивность практически не изменяется при введении штырей) увеличивается ёмкостная связь, что уменьшает k_{FEXT}.

В табл. 1 представлены коэффициенты k_C, k_L и k_{FEXT}, вычисленные на основе данных из [25].

Из табл. 1 видно, что добавление штырей дало рост k_C в 3,42 раза при неизменном k_L. В результате k_{FEXT} снижается на 34%. Отметим, что рост k_C также увеличит V_{NEXT}. Также надо учитывать, что введение штырей локально изменяет волновое сопротивление, что приводит к появлению отражений.

Дополнительно на рис. 6 приведены частотные зависимости V_{FEXT} в диапазоне 6–15 ГГц. Видно, что встречно-штыревая структура даёт уменьшение V_{FEXT} на 10–20 дБ в широкой полосе частот по сравнению с обычной связанной МПЛ. Это подтверждает эффективность данного способа для подавления V_{FEXT}.

Примечательно, что оценка роста k_C во встречно-штыревой структуре возможна в аналитическом виде без затратного электродинамического моделирования. Для этого разработаны аналитические модели в замкнутой форме для вычисления погонных параметров структуры с учётом асимметрии её проводников [26].

Таблица 1

Влияние встречно-штыревой структуры на коэффициенты k_C, k_L и k_{FEXT}

Структура	k_C	k_L	$k_{FEXT} \cdot 10^{-10}$
Без штырей	0,012	0,075	28,69
Со штырями	0,041	0,075	19,04

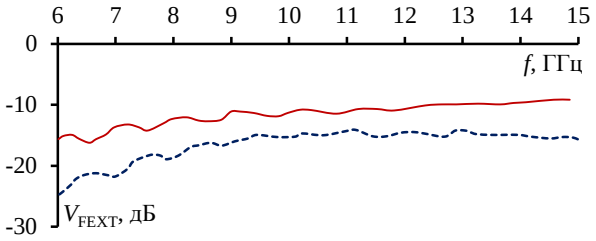


Рис. 6. Частотная зависимость V_{FEXT} без штырей (–) и со штырями (--)

Одним из направлений развития способа является оценка возможности выбора оптимальной геометрии штырей и их применение в различных топологиях. В [27] встречно-штыревые выступы реализованы непосредственно в контактных площадках и проводниках соединителя интерфейса HDMI 2.1. Это позволило эффективно увеличить эквивалентное расстояние между дифференциальными парами без их физического разнесения, что важно в ограниченном пространстве соединителя, а также уменьшить V_{FEXT} на 8–15 дБ в полосе частот до 12 ГГц.

Исследовано также влияние формы штырей на емкостную связь. К примеру, в работе [28] показано, что треугольная форма штырей способствует контролю k_C и уменьшает V_{FEXT} на частотах резонанса по сравнению с прямоугольной формой [25]. Известен также подход с применением машинного обучения [29]: встречно-штыревые структуры располагаются в многопроводных линиях передачи не с одинаковым периодом, а псевдослучайно, что позволяет управлять разностью $k_C - k_L$.

Встречно-штыревые структуры исследуют не только в одиночном режиме передачи сигнала, но и в дифференциальном. Так, в [30] предложен подход, позволяющий повысить k_C при плотной трассировке дифференциальных пар. Известен и комбинированный подход [31], в котором трассировка дифференциальных пар на внутреннем слое многослойной ПП сочетается со встречно-штыревой структурой на переходных участках. Такая трассировка позволяет обеспечить подавление V_{NEXT} более чем на 30 дБ при сохранении уровня V_{FEXT} в отличие от простой связанной МПЛ в одиночном режиме передачи сигнала.

Несмотря на эффективность встречно-штыревых структур для подавления V_{FEXT}, способ обладает рядом недостатков:

1. Повышение V_{NEXT} за счёт роста общей ёмкостной связи.

2. Усложнение технологического производства.

3. Неоднородность структуры. На высоких частотах возможно уменьшение импеданса.

Важно помнить, что V_{FEXT}, согласно выражению (1), определяется суммой коэффициентов емкостной и индуктивной связи, и любое увеличение k_C для компенсации V_{FEXT} неизбежно приводит к росту помехи V_{NEXT}.

Дефектные структуры опорного слоя

Другим конструктивным способом для подавления перекрестных помех является применение дефектных структур опорного слоя (ДСОС). Способ заключается в целенаправленном нарушении целостности опорного слоя для управления путём протекания возвратного тока (увеличения его длины) [32]. На рис. 7 представлен пример связанной МПЛ с периодической ДСОС.

В МПЛ со сплошным опорным проводником возвратные токи сосредоточены непосредственно под сигнальными проводниками из-за наличия между ним и опорным слоем ёмкостной и индуктивной связи. Введение периодических дефектов в опорном проводнике увеличивает эффективную длину пути протекания возвратного тока. ДСОС также локально влияет на собственные и взаимные погонные параметры связанной МПЛ. В табл. 2 сведены количественные оценки влияния ДСОС на коэффициенты k_C, k_L и k_{FEXT}, полученные на основе данных из [32].

Из табл. 2 видно, что введение ДСОС в МПЛ дало уменьшение k_C и k_L в 4 и 5 раз соответственно. В результате этого k_{FEXT} уменьшился, однако это уменьшение оказалось незначительным. Частотные зависимости V_{FEXT} в МПЛ с ДСОС и без неё показаны на рис. 8. Видно, что ДСОС позволяет подавить V_{FEXT} на 5 дБ в широком диапазоне частот.

Известно множество исследований, посвящённых оптимизации формы, размера и расположения ДСОС в связанных МПЛ.

В работе [33] проведено сравнение ДСОС различных форм: круглой, квадратной и прямоугольной.

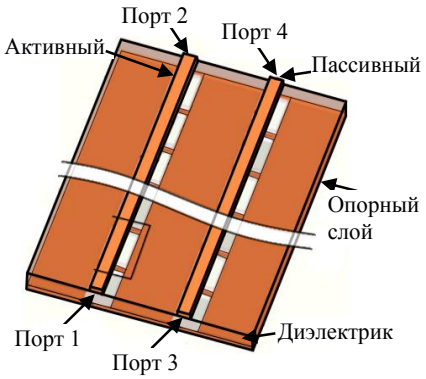


Рис. 7. Пример ДСОС в связанной МПЛ

Таблица 2

Влияние ДСОС на коэффициенты k_C, k_L и k_{FEXT}

Структура	k_C	k_L	$k_{FEXT} \cdot 10^{-10}$
Без ДСЗ	0,016	0,075	20,2
С ДСЗ	0,004	0,015	17,3

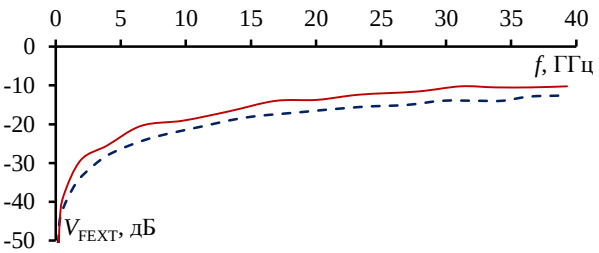


Рис. 8. Частотные зависимости V_{FEXT} без ДСЗ (–) и с ней (--)

Авторы показали, что круглые ДСОС более эффективны для целей подавления помех благодаря большей площади дефекта при той же ширине, что обеспечивает локальное снижение электромагнитной связи между проводниками и более выраженное увеличение пути возвратного тока. В [34] предложена ДСОС в виде периодических штыревых вырезов в опорной плоскости МПЛ. Это даёт экономию площади на ПП в условиях плотной трассировки и ограничений по габаритам устройств, сохраняя при этом низкий уровень излучаемых помех и эффективное подавление V_{FEXT}. На рис. 9 представлен вид сверху для периодической структуры с ДСОС.

Особый интерес представляет применение ДСОС в дифференциальных линиях [35]. В отличие от решений, описанных выше (с полными вырезами в опорном слое под двумя проводниками, как в [32–34]), здесь предложено периодическое удаление металлизации опорного слоя до ширины печатного проводника и удаление избыточной металлизации на расстоянии более $3w$ от края проводника. Это позволяет уменьшить соотношение сигнал/шум для синфазной моды, минимизировать преобразование дифференциального сигнала в синфазный и повысить качество передачи сигнала в дифференциальных парах.

Кроме формы, критически важен и размер ДСОС. В исследовании [36] выполнено сравнение структур разного размера и выявлена закономерность: увеличение площади дефекта приводит к росту пиковой напряженности электрического поля в области связи, что, напротив, может негативно влиять на V_{NEXT} и V_{FEXT} (из-за усиления локальных резонансов или неконтролируемого изменения импеданса). Следовательно, существует оптимальный размер ДСЗ, зависящий от рабочей частоты и параметров подложки.

Таким образом, несмотря на доказанную эффективность применения ДСОС для подавления перекрестных помех (вплоть до 25 дБ для V_{NEXT} и V_{FEXT} в зависимости от конкретной конфигурации), способ имеет ряд ограничений и нерешённых задач. Среди них можно выделить:

1. Отсутствие унифицированной методики проектирования связанных МПЛ с ДСОС (требуется эмпирический подбор формы, размера и расположения для каждой конкретной топологии и частотного диапазона).

2. Применение ДСОС может приводить к изменению волнового сопротивления, появлению паразитных резонансов, росту потерь на излучение или к отражениям.

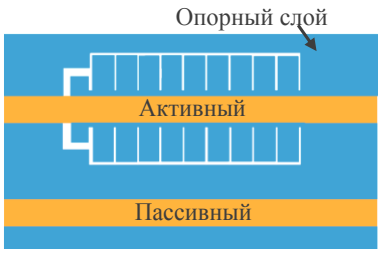


Рис. 9. Периодическая структура ДСОС (вид сверху)

Вырезы в пассивном проводнике

Альтернативным подходом к подавлению V_{FEXT} является создание вырезов (дефектов) не в опорном слое (как в случае ДСОС), а непосредственно в пассивном проводнике. Вырезы также локально изменяют путь протекания тока и электромагнитное поле вокруг проводника, что приводит к локальной коррекции k_C и k_L и уменьшению k_{FEXT}. На рис. 10 показан пример реализации *LL*-образных вырезов в проводниках связанной МПЛ [37].

Структура с *LL*-образными вырезами исследована в [37]. Результаты оценки влияния вырезов в МПЛ на k_C, k_L и k_{FEXT} на основе параметров из [37] приведены в табл. 3.

Как видно из табл. 3, введение *LL*-образных вырезов увеличивает k_C в 2,51 раза и несущественно уменьшает k_L. Это даёт уменьшение k_{FEXT} в 2,2 раза.

Результаты измерения частотной зависимости V_{FEXT} в диапазоне до 18 ГГц показаны на рис. 11 [37].

Из рис. 11 следует, что применение *LL*-образных вырезов уменьшает V_{FEXT} в широкой полосе частот (с пиковым подавлением до минус 45 дБ), в то время как в МПЛ без вырезов V_{FEXT} достигает минус 10 дБ.

Помимо *LL*-образных вырезов, интерес представляют конфигурации с *TT*-образными (*T* и инвертированная *T*) и *TL*-образными вырезами, а также периодическими спиральными резонаторами. В исследовании [38] предложен *TT*-образный вырез, где в пассивном проводнике вытравливается массив комбинаций *T*- и инвертированной *T*-форм.

Это обеспечивает высокое подавление как V_{NEXT}, так и V_{FEXT} за счёт локального изменения пути протекания тока и создания компенсирующих полей.

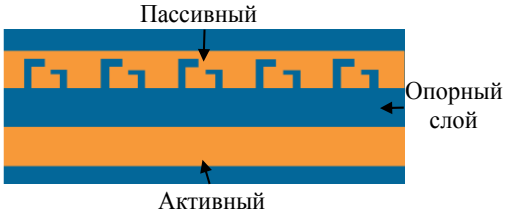


Рис. 10. Использование *LL*-образных вырезов в пассивном проводнике (вид сверху)

Таблица 3

Влияние *LL*-вырезов на k_C, k_L и k_{FEXT}

Структура	k_C	k_L	$k_{FEXT} \cdot 10^{-10}$
Без вырезов	0,027	0,079	14,75
С вырезами	0,068	0,076	6,68

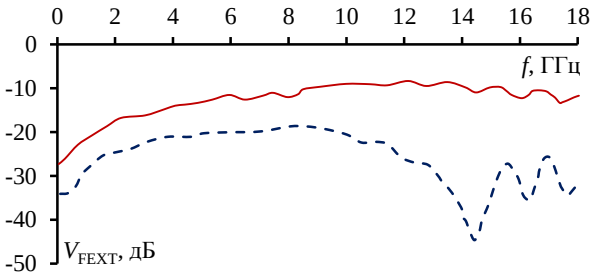


Рис. 11. Частотная зависимость V_{FEXT} без *LL*-вырезов (–) и с ними (--)

Другой принцип описан в [39], где вырезы выполнены в форме квадратных катушек индуктивности, что даёт эффективное уменьшение обоих видов помехи. В [40] исследован вырез в форме меандра.

В работе [41] оценивалась эффективность *TL*-образных вырезов. Примечательно, что этот тип вырезов более эффективен по сравнению с другими, благодаря выравниванию k_C и k_L, а также минимальным паразитным эффектам.

Также исследуется эффективность гибридных решений. В [42] показано одновременное применение встречно-штыревых структур и дефектов в проводнике, что позволило достичь ослабления V_{FEXT} до минус 25 дБ в широкой полосе частот (в отличие от обычной связанной МПЛ). Особенно выделяется способ, предложенный в [43]: вырез в проводнике в виде двух встречно направленных периодических спиральных катушек индуктивности. На рис. 12 представлен пример такой МПЛ.

Таким образом, разнообразие форм вырезов в проводниках позволяет обеспечить высокое подавление V_{NEXT} и V_{FEXT} (часто на 20–40 дБ ниже, чем в обычной связанной МПЛ) в широком диапазоне частот. Гибридные решения обеспечивают рост ослабления. К недостаткам данного способа можно отнести технологические ограничения и ограниченную применимость в интегральных схемах и высокоскоростных интерфейсах с $w \leq 100$–150 мкм.

Опорные проводники

Эффективным конструкторским способом, часто применяемым на практике, для подавления V_{NEXT} и V_{FEXT} является введение опорного проводника между активным и пассивным проводниками связанной МПЛ. Типовая конфигурация такой МПЛ показана на рис. 13.

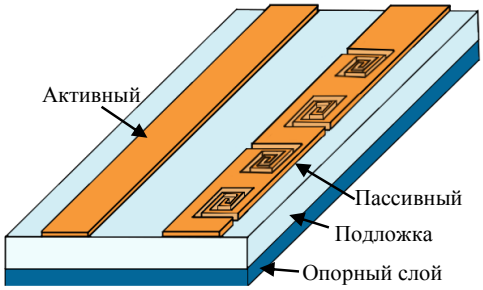


Рис. 12. Связанная МПЛ с вырезами в виде встречно направленных катушек индуктивности

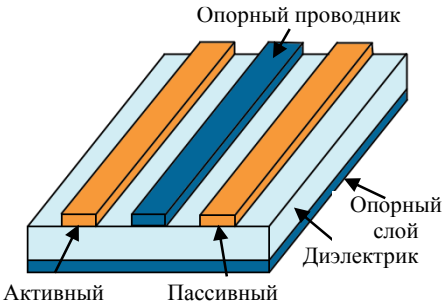


Рис. 13. Связанная МПЛ с опорным проводником

Способ основан на том, что опорный проводник «отбирает» существенную часть электромагнитной энергии. В результате этого большая часть энергии отводится в землю, а не наводится на пассивный проводник, что приводит к заметному снижению V_{NEXT} и V_{FEXT}.

Одним из примечательных исследований в этой области является работа [44], где предложено использовать опорный проводник в форме меандра вместо обычного прямолинейного проводника. Такая конфигурация увеличивает ёмкостную связь между активным и опорным проводником, сохраняя при этом индуктивность на относительно низком уровне, что эффективно для компенсации V_{FEXT}.

В табл. 4 сведены k_C, k_L и k_{FEXT} для различных конфигураций защитного проводника, полученные на основе данных из [44].

Из табл. 4 видно, что прямой опорный проводник в 2 раза увеличивает k_C при сохранении k_L. Добавление периодических переходных отверстий на опорный слой снижает k_L в 2 раза, что эффективно подавляет V_{NEXT}. Проводник в форме меандра даёт максимальное увеличение k_C при сохранении k_L на уровне обычной линии, что приводит к наилучшей компенсации V_{FEXT}.

На рис. 14 представлен временной отклик V_{FEXT} [44].

Из рис. 14 видно, что структура с меандровой линией обеспечивает наименьшую пиковую V_{FEXT} (в 3–5 раз ниже, чем без нее), а также более короткую длительность помехи.

Примечательно исследование, где рассмотрено совместное применение заземленного проводника со встречно-штыревой структурой [45]. Пример топологии такой структуры показан на рис. 15.

Данное решение приводит к одновременному увеличению как собственных, так и взаимных составляющих матриц **C** и **L**, что позволяет эффективно компенсировать разность скоростей мод и снижать V_{FEXT} в диапазоне до 40 ГГц и выше.

Таблица 4

Значения k_C, k_L и k_{FEXT} для различных форм опорного проводника

Структура	k_C	k_L	$k_{FEXT} \cdot 10^{-12}$
Без опорного проводника	0,0075	0,036	–2,85
С опорным проводником без заземления	0,015	0,036	–2,1
С опорным проводником с заземлением	0,003	0,017	–1,4
Меандровая форма опорного проводника	0,0248	0,035	–1,2

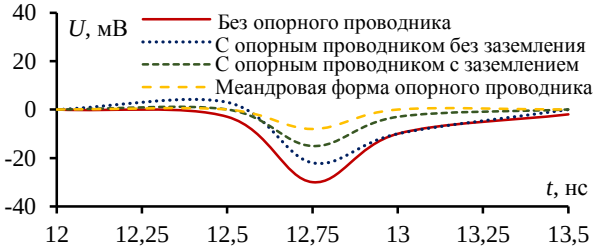


Рис. 14. Временной отклик дальней перекрестной помехи для различных форм опорных проводников

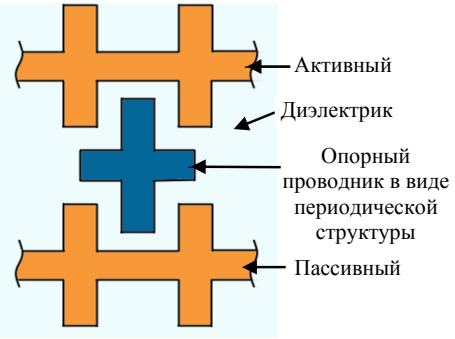


Рис. 15. Топология гибридного решения на основе встречно-штыревой структуры с опорными периодическими структурами

Схожая идея развивается в [46], где гибридная структура включает уже три метода одновременно: встречно-штыревую структуру, опорный проводник в виде периодической структуры, трассировку проводников в виде меандра. В [47] акцент сделан на анализе различных граничных условий (холостой ход и короткое замыкание) на концах опорного проводника в виде прямой линии и периодической структуры. Исследование показывает, что оптимальные характеристики достигаются при использовании периодических структур.

В [48] предложена гибридная структура – опорный проводник в форме меандра. Такой способ обеспечивает уменьшение уровня перекрестных помех более чем в 2 раза по сравнению с обычными меандровыми линиями, при этом заземление опорного проводника достаточно выполнить только на концах (без необходимости переходных отверстий, как в более ранних работах [10]).

В [49] рассмотрено применение нерегулярного защитного проводника с изменяющейся шириной (от 762 до 256 мкм) и зазором (от 254 до 508 мкм). Введение нерегулярности позволяет адаптировать структуру к конкретной частотной характеристике и добиться высокого подавления V_{NEXT} за счёт оптимизации локальной связи. Также исследуется применение защитных трасс в дифференциальных линиях [50]. Разработаны практические рекомендации по их применению:

1. Защитные проводники обязательно должны быть заземлены (через переходные отверстия) для эффективного отвода энергии.

2. Если заземление через каждые λ/20 невозможно, лучше полностью удалить защитный проводник и компенсировать амплитуды помех увеличением расстояния между проводниками МПЛ.

Метаматериалы

Одним из новых и перспективных решений для борьбы с перекрестными помехами является использование метаматериалов и метаповерхностей. Эти искусственные структуры вводятся в поперечное сечение МПЛ в виде дополнительных слоёв или элементов между проводниками. Метаматериалы позволяют поглощать, перенаправлять или подавлять паразитные электромагнитные поля, а метаповерхности – выравнивать фазовые скорости составляющих сигнала.

Примером применения метаматериалов для подавления помех является работа [51]. Авторы предлагают встраивать вспененный материал с низкой ε_r (близкой к 1) между слоями стандартного FR-4. Поперечное сечение такой структуры показано на рис. 16.

Толщины слоёв пены и нижнего слоя FR-4 подбираются так, чтобы k_C стал близок к k_L. В результате k_{FEXT} становится близким к нулю. Это достигается за счёт искусственного изменения $\varepsilon_{r\text{эфф}}$ в области связи, что компенсирует естественную неоднородность МПЛ. Механизм подавления помех в данной структуре заключается в локальном снижении эффективной диэлектрической проницаемости в области концентрации электрического поля связи. Использование материала с $\varepsilon_r \approx 1$ позволяет приблизить параметры распространения мод к условиям в однородной среде. Это минимизирует модовую дисперсию и позволяет достичь практически нулевой V_{FEXT} в широкой полосе частот. Количественные оценки эффективности применения метода, полученные на основе параметров из [51], приведены в табл. 5.

Видно, что слой пены позволяет уменьшить k_{FEXT} до 0. Это обеспечивает практически полную компенсацию V_{FEXT} в рабочей полосе частот при должном согласовании. На рис. 17 представлена частотная зависимость V_{FEXT} для структур с метаматериалом и без него. Видно, что структура с метаматериалом обеспечивает V_{FEXT} ниже –20 дБ в широкой полосе (с пиковым подавлением до –48 дБ), в то время как для обычной МПЛ V_{FEXT} находится вблизи –10 дБ.

В работе [52] предложена структура на основе дефектной линии передачи в виде резонатора, интегрированного в МПЛ.

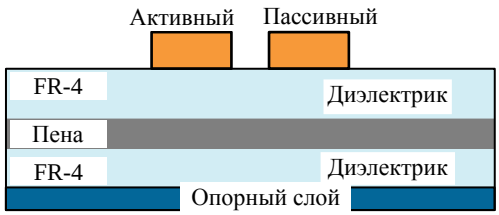


Рис. 16. Поперечное сечение связанной МПЛ с добавлением пены

Таблица 5

Влияние метаматериала на значения k_C, k_L и k_{FEXT}

Структура	k_C	k_L	$k_{FEXT}\cdot10^{-10}$
Без пены	0,082	0,175	–0,1
С пеной	0,147	0,147	0

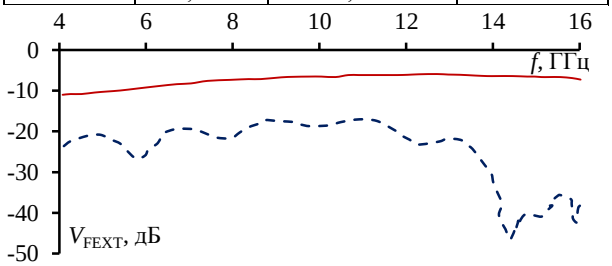


Рис. 17. Частотная зависимость V_{FEXT} для структур с метаматериалом (--) и без него (–)

На рис. 18 показана реализация прототипа с резонаторными элементами. Данная конфигурация позволяет создать локальные «запрещённые» зоны распространения для паразитных полей, что приводит к значительному снижению V_{FEXT} (до –30 дБ) в широком диапазоне частот без заметного влияния на полезный сигнал.

Ещё одним примером метаповерхности является периодическая структура в плоскости земли с прямоугольными элементами в многослойных ПП [53]. Суть способа заключается во введении периодических прямоугольных вырезов и штырей в пассивном проводнике. Это позволяет искусственно приблизить k_C к k_L и снизить k_{FEXT} в 3 раза. На рис. 19 показан пример данной структуры.

Аналогичный принцип используется в [54], где в опорной плоскости создаются вырезы в виде встречных *C*-форм. Такая конфигурация формирует резонансные ячейки, эффективно подавляющие поверхностные волны и уменьшающие V_{FEXT} в СВЧ-диапазоне. Примером метаповерхностей также являются круглые резонаторы [55]. Размещение периодических резонаторов позволяет сохранять значение индуктивной связи и тем самым подавлять V_{FEXT}.

Таким образом, применение метаматериалов и метаповерхностей для подавления перекрестных помех позволяет модифицировать емкостную и индуктивную связь и обеспечивать полное подавление V_{FEXT}.

Многопроводные линии передачи

В ПП встречаются не только пары связанных МПЛ, но и целые группы из нескольких МПЛ. Перекрестные помехи в них особенно опасны, так как помеха может последовательно наводиться с одной линии на другую, из-за чего необходимо контролировать сигнал в каждой линии [56]. Существует несколько подходов к решению этой проблемы.

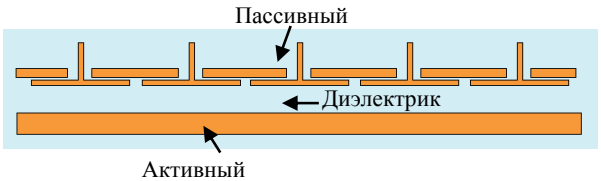


Рис. 18. Прототип дефектной линии передачи с метаматериальным резонатором

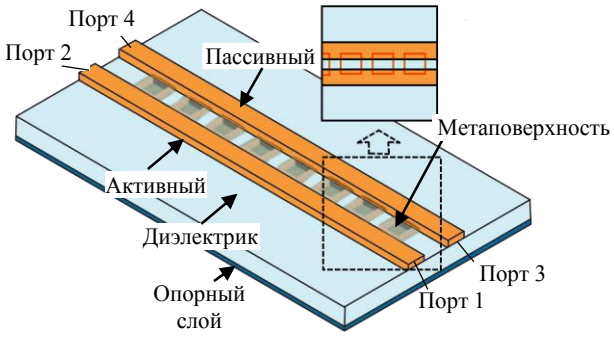


Рис. 19. Применение метаповерхности в виде периодических прямоугольников в межслойном опорном слое под связанными МПЛ

Например, в работах [57, 58] авторы предлагают использовать трансформаторы. В способе для четырехпроводных линий [57] обмотки трансформатора соединяются последовательно и встречно. Это позволяет разделить сигналы на «суммарный» (четная мода) и «разностный» (нечетная мода). Такие моды распространяются по проводам независимо и не наводят помех друг на друга. Для линий с большим числом проводников [58] в схему добавляют матричный коммутатор. Он распределяет сигналы по обмоткам так, чтобы в линии передачи возникали ортогональные электромагнитные поля, которые также не взаимодействуют между собой.

В работе [59] описан способ подавления FEXT и NEXT, в котором помеха на выходе конкретной линии рассматривается как сумма влияний всех остальных проводников, характеризуемых специальными коэффициентами – факторами перекрестной помехи. Зная эти коэффициенты, можно управлять фазами сигналов на входе так, чтобы помехи от разных линий складывались в противофазе и взаимно компенсировались.

Другой математический метод описан в [60] – использование сингулярного разложения. Матрицу передачи линии приводят к диагональному виду: ненулевые элементы остаются только на главной диагонали, а внедиагональные элементы, описывающие паразитные связи (помехи), обращаются в нуль.

Также перекрестные помехи можно подавить с помощью кодирования сигналов. В работе [61] используется преобразователь мод на входе линии. Он распределяет сигналы по проводникам в определенных комбинациях, в результате чего в линии распространяются волны, математически независимые друг от друга. На выходе установлен аналогичный преобразователь, выполняющий обратную операцию: он восстанавливает полезный сигнал и подавляет помехи.

Другие обзорные работы

Известен обзор [62], в котором рассмотрены различные способы подавления перекрестных помех. Методы с защитными проводниками остаются одними из самых надёжных для подавления V_{NEXT} и V_{FEXT}. Паяльная маска как дополнительное диэлектрическое покрытие заметно снижает V_{FEXT} (за счёт увеличения $\varepsilon_{r\text{эфф}}$ и выравнивания k_C и k_L), но приводит к удорожанию производства ПП.

В обзоре [63] авторы уделяют внимание современным способам подавления FEXT и NEXT. Структурные методы с вырезами в проводниках, а также метаматериалы демонстрируют лучшую эффективность по сравнению с классическим способом ($s = 3 – 5w$).

Сравнение рассмотренных способов

Выполнено сравнение представленных в работе способов (табл. 6). Оно показывает, что эффективными способами подавления V_{FEXT} в широкой полосе частот являются метаматериалы и метаповерхности, а также гибридные комбинации, например, сочетание встречно-штыревых структур с дифференциальной трассировкой. Однако заземленные проводники остаются одними из простых практических решений, хотя и требуют значительной площади на ПП.

Таблица 6

Сравнение способов подавления перекрестных помех

Способ	Принцип работы	Подавление	Недостатки	Источник
Трассировка на внутренних слоях ПП	Обеспечение равенства параметров диэлектрических слоёв структуры	Уменьшение V_{FEXT} в 2–5 раз	Невозможность полной минимизации V_{FEXT} в условиях неоднородной среды; зависимость от параметров диэлектриков	[15–21]
Применение встречно-штыревых структур	Увеличение ёмкостной связи для минимизации разницы $k_C - k_L$	Уменьшение V_{FEXT} на 10–30 дБ в широкой полосе частот	Повышение V_{NEXT} из-за увеличения емкостной связи	[22–28]
Дефектные структуры земли	Нарушение целостности опорного слоя для увеличения длины пути протекания возвратных токов и локальное изменение ёмкостной и индуктивной связей	Уменьшение V_{NEXT} / V_{FEXT} на 10–25 дБ (для FEXT эффективность выше)	Рост излучаемых помех; возникновение резонансов при увеличении размера	[29–33]
Вырезы в пассивном проводнике	Увеличение длины пути протекания тока в проводнике и локальное изменение ёмкостной и индуктивной связей	Уменьшение V_{FEXT} на 15–40 дБ	Рост потерь и отражений; изменение импеданса; рост сложности травления; ограниченная применимость в интегральных схемах и линиях малой ширины	[34–40]
Опорные проводники	«Отбор» энергии опорным проводником и ее отвод в опорную плоскость	Уменьшение V_{NEXT}/ V_{FEXT} на 20–50 дБ	Значительная площадь на ПП; необходимость переходных отверстий	[41–47]
Метаматериалы и метаповерхности	Искусственное изменение ε_r или μ_r; компенсация разности фазовых скоростей мод	Уменьшение V_{FEXT} на 20–50 дБ	Увеличение толщины платы; высокая сложность и стоимость изготовления	[48–52]
Многопроводные линии передачи	Использование модового преобразования, управление фазами сигналов-возбудителей, кодирование и декодирование сигналов	Полное подавление V_{FEXT} (в идеализированных моделях)	Высокая схемотехническая сложность; необходимость установки дополнительных корректирующих блоков	[54–58]

Дефектные структуры хорошо подходят для условий плотной компоновки, однако их эффективность зависит от геометрии и частоты.

Переход на внутренние слои с контролем диэлектрика является способом снижения V_{FEXT} в многослойных ПП в условиях неоднородной структуры.

Заключение

Выбор способа подавления перекрестных помех перестал быть тривиальной задачей: он напрямую определяется рабочими частотами, плотностью трассировки и технологическими возможностями производства.

Традиционные подходы, такие как правило 3w или использование защитных проводников, остаются основой для систем с рабочими частотами до 10 ГГц. Однако на более высоких частотах их потенциал исчерпывается из-за жестких требований к габаритам. Им на смену приходят способы, основанные на искусственном выравнивании коэффициентов связи ($k_L \approx k_C$), что позволяет добиться подавления V_{FEXT} до уровня минус 50 дБ.

Перспективным направлением является использование метаматериалов и метаповерхностей, которые дают возможность локально управлять электрофизическими параметрами среды. Это позволяет уменьшить модовую дисперсию за счет выравнивания фазовых скоростей мод.

В целом развитие современных высокоскоростных интерфейсов диктует переход от простых эмпирических правил трассировки к строгому электродинамическому синтезу. Только глубокий контроль параметров распространения мод на этапе проектирования позволит обеспечить целостность сигналов в инфокоммуникационных системах.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (проект № FEWM-2026-0013).

Литература

1. Wu T.-L. Overview of signal integrity and EMC design technologies on PCB: Fundamentals and latest progress / T.-L. Wu, F. Buesink, F. Canavero // IEEE Transactions on Electromagnetic Compatibility. – 2013. – Vol. 55, No. 4. – P. 624–638.

2. Far-End Crosstalk Mitigation for Transmission Lines in DDR5 Using Glass-Weave Coating Structure / X.-B. Yu et al. // 2021 IEEE USNC-URSI Radio Science Meeting (Joint with AP-S Symposium). – Singapore, Singapore: IEEE, 2021. – P. 23–24.

3. Bogatin E. Signal and Power Integrity – Simplified. – New Jersey, U.S.: Prentice Hall, 2010. – 757 p.

4. Worst Eye Performance Analysis for Advanced Package Die-to-Die Interconnects / H.-C. Kuo, P.-C. Pan, L.-C. Hung, M.-F. Jhong, C.-C. Wang // 2025 IEEE 29th Workshop on Signal and Power Integrity (SPI). – Gaeta, Italy: IEEE, 2025. – P. 1–4.

5. Study on the Impact of Imbalance Between Transmission Lines on Crosstalk: A Novel Perspective of Displacement Current / X. Lu, L. Song, H. Xu, D. Su // Chinese Journal of Electronics. – 2025. – Vol. 34, No. 1. – P. 114–124.

6. A 32 Gb/s 0.41 pJ/bit Single-Ended Transmitter with TX-Based-Only Adaptive Crosstalk Cancellation for Ultra-Short-Reach Wireline Applications / Y. Shen, Z. Su, C. Liu, X. Ma, L. Wang, B. Li // Applied Sciences. – 2026. – Vol. 16, No. 11. – Art. no. 5672.

7. Krage M.K. Characteristics of coupled microstrip transmission lines-I: Coupled-mode formulation of inhomogeneous lines / M.K. Krage, G.I. Haddad // IEEE Transactions on Microwave Theory and Techniques. – 1970. – Vol. 18, No. 4. – P. 217–222.

8. DeFalco J.A. Reflection and crosstalk in logic circuit interconnections // IEEE Spectrum. – 1970. – Vol. 7, No. 7. – P. 44–50.

9. Кечиев Л.Н. Проектирование печатных плат для цифровой быстродействующей аппаратуры. – М.: ООО «Группа ИДТ», 2007. – 615 с.

10. Mbairi F.D. High-Frequency Transmission Lines Crosstalk Reduction Using Spacing Rules / F.D. Mbairi, W.P. Siebert, H. Hesselbom // IEEE Transactions on Components and Packaging Technologies. – 2008. – Vol. 31, No. 3. – P. 601–610.

11. Novak I. Measurement by vector-network analyzer and simulation of crosstalk reduction on printed circuit boards with additional center traces / I. Novak, B. Eged, L. Hatvani // 1993 IEEE Instrumentation and Measurement Technology Conference. – Irvine, CA, USA: IEEE, 1993. – P. 269–274.

12. Сычев А.Н. Системы параметров одинаковых связанных линий с неуравновешенной электромагнитной связью / А.Н. Сычев, С.М. Стручков // Доклады ТУСУР. – 2014. – № 1 (31). – С. 39–50.

13. Сычев А.Н. Анализ и синтез несимметричных связанных линий в однородной диэлектрической среде // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2019. – Т. 22, № 1. – С. 11–19.

14. Сычев А.Н. Параметры несимметричных связанных линий с неоднородным диэлектриком / А.Н. Сычев, Н.Ю. Рудый // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2018. – Т. 21, № 4-1. – С. 7–15.

15. Сычев А.Н. Синтез модальных и погонных параметров несимметричных связанных линий с неоднородным диэлектриком // Журнал радиоэлектроники. – 2024. – № 6. – С. 1–41.

16. Gilb J.P. Pulse distortion on multilayer coupled microstrip lines / J.P. Gilb, C.A. Balanis // IEEE Transactions on Microwave Theory and Techniques. – 1989. – Vol. 37, No. 10. – P. 1620–1628.

17. Surovtsev R.S. Optimization of protective varnish thickness to minimize crosstalk in multiconductor bus of spacecraft PCB / R.S. Surovtsev, T.R. Gazizov // Proc. of IEEE Moscow Workshop on Electronic and Networking Technologies. – Moscow, Russia: IEEE, 2018. – P. 1–6.

18. Сравнение эффективности методов уменьшения перекрестных связей в линиях передачи печатных плат / И.А. Скорняков, С. Карри, Р.С. Суровцев, А.Е. Мандель // Доклады ТУСУР. – 2025. – Т. 28, № 2. – С. 22–28.

19. Far-End Crosstalk Modeling and Prediction for Stripline with Inhomogeneous Dielectric Layers (IDLs) / Y. Liu et al. // IEEE Transactions on Signal and Power Integrity. – 2022. – Vol. 1. – P. 93–103.

20. A Comprehensive and Practical Way to Look at Crosstalk for Transmission Lines with Mismatched Terminals / S. Yong, K. Cai, B. Sen et al. // 2018 IEEE Symposium on Electromagnetic Compatibility, Signal Integrity and Power Integrity (EMC, SI & PI). – Long Beach, CA, USA: IEEE, 2018. – P. 538–543.

21. Prepreg and Core Dielectric Permittivity (ϵ_r) Extraction for Fabricated Striplines' Far-End Crosstalk Modeling / S. Yong et al. // IEEE Transactions on Electromagnetic Compatibility. – 2022. – Vol. 64, No. 1. – P. 209–218.

22. Extraction of DKglass and DKresin for PCB Striplines and Analysis of Far-End Crosstalk / J. Park, Y. Guo, X. Ye, D. Kim // IEEE Transactions on Signal and Power Integrity. – 2025. – Vol. 4. – P. 149–155.

23. Gazizov T.R. Far-end crosstalk reduction in double-layered dielectric interconnects // IEEE Transactions on Electromagnetic Compatibility. – 2001. – Vol. 43, No. 4. – P. 566–572.

24. Пат. 2022/0311114 A1 США, МПК H 01 P 3/08. Dual-stripline with crosstalk cancellation / A. Sutono (США), X. Ye (США); заявитель Intel Corporation (США). – № 17/214,111; заявл. 26.03.2021; опубл. 29.09.2022. – 25 с.

25. High-Performance Interconnects with Reduced Far-End Crosstalk for High-Speed ICs and Communication Systems / J. Ge, R. Floyd, A. Khan, G. Wang // IEEE Transactions on Components, Packaging and Manufacturing Technology. – 2023. – Vol. 13, No. 7. – P. 1013–1020.

26. Jiang W. Equation-based solutions to coupled, asymmetrical, lossy, and nonuniform microstrip lines for tab-routing applications / W. Jiang, X.-D. Cai, B. Sen, G. Wang // IEEE Transactions on Electromagnetic Compatibility. – 2019. – Vol. 61, No. 2. – P. 548–557.

27. Design and Analysis of HDMI 2.1 Connector for Crosstalk Reduction using Tabs and Inverse Tabs / D. Lho et al. // 2021 IEEE 30th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS). – San Jose, CA, USA: IEEE, 2021. – P. 1–3.

28. Analysis of Differential Stripline Routing Approaches within a PCB Via Field for Crosstalk Mitigation / S. Datta et al. // 2023 IEEE 32nd Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS). – Milpitas, CA, USA: IEEE, 2023. – P. 1–3.

29. Homogeneous, lossy and Asymmetrical Strip Line with tabbed design for High-Speed Interface application / A. Patel, G. Pandey, G. Verma et al. // ITM Web of Conferences. – 2023. – Vol. 54. – Art. no. 02007.

30. Analysis Method for Curved Coupled Stripline With Triangular Tabs Within Pin Field of High-Speed Channels / Y. Zhang, X-D. Cai et al. // 2025 IEEE International Symposium on Electromagnetic Compatibility, Signal & Power Integrity (EMC+SIPI). – Raleigh, NC, USA: IEEE, 2025. – P. 338–342.

31. Reinforcement Learning-Based Design of Interconnects with Reduced Far-End Crosstalk / C.-J. Mai, Ch.-S. Mao, et al. // IEEE Transactions on Electromagnetic Compatibility. – 2025. – Vol. 67, No. 4. – P. 1370–1374.

32. A low loss microstrip line on thin flexible substrate film by defected ground structure / A. Tsuchiya, Yu. Yamaguchi et al. // IEICE Communications Express. – 2023. – Vol. 12, No. 11. – P. 582–585.

33. Sindhadevi M. Crosstalk Reduction Using Defective Ground Plane Structures in RF Printed Circuit Boards / M. Sindhadevi, K. Malathi, A. Henridass et al. // Arabian

Journal for Science and Engineering. – 2014. – Vol. 39. – P. 1107–1116.

34. A SSPP-Integrated Defected Ground Structure with Low EMI / X. Luo, K. Yang, L. Li, H. Liu // IEEE Letters on Electromagnetic Compatibility Practice and Applications. – 2025. – Vol. 8, No. 1. – P. 7–10.

35. Balan T.S. Suppression of Common and Differential Component of Crosstalk Coupling onto the Differential Line and Field analysis of the Crosstalk / T.S. Balan, B. Devadas, G.K. Srivastava // 2018 IEEE Symposium on Electromagnetic Compatibility, Signal Integrity and Power Integrity (EMC, SI & PI). – Long Beach, CA, USA: IEEE, 2018. – P. 260–264.

36. Kazerooni M. Unit Length Parameters, Transition Sharpness and Level of Radiation in Defected Microstrip Structure (DMS) and Defected Ground Structure (DGS) Interconnections / M. Kazerooni, A. Cheldavi, M. Khalil // Progress in Electromagnetics Research M. – 2009. – Vol. 10. – P. 93–102.

37. Zhang Y. UWB Far-End Crosstalk Mitigation with «LL» Shaped Defected Tabbed Routing Structures / Y. Zhang, G. Wang // IEEE Transactions on Electromagnetic Compatibility. – 2025. – Vol. 67, No. 1. – P. 247–253.

38. Crosstalk Reduction in Coupled Microstrip Lines using TT-shaped DMS Approach / L.K. Baghel, S. Kumar, S.A. Sis, J. Kim // 2023 Joint Asia-Pacific International Symposium on Electromagnetic Compatibility and International Conference on ElectroMagnetic Interference & Compatibility (APEMC/INCEMIC). – Bengaluru, India: IEEE, 2023. – P. 1–4.

39. Equivalent Circuit Model of Crosstalk Reduction Parallel Transmission Lines with Defected Microstrip Structures / X. Liu, Y. Li, Y. Zhao et al. // 2018 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC). – Xuzhou, China: IEEE, 2018. – P. 1–2.

40. Gobinath A. Reduction of near-end and far-end crosstalk in microwave and millimetre wave band of parallel transmission lines using meander shaped DMS / A. Gobinath, N.S. Kumar, P. Rajeswari // Journal of Physics: Conference Series. – 2023. – Vol. 2466, No. 1. – Art. no. 012016.

41. Li C. Crosstalk reduction between microstrip lines using TL-shaped defected microstrip structure / C. Li, Y. Wang, X. Li // Engineering Reports. – 2023. – Vol. 5, No. 1. – Art. no. e12559.

42. Yokesh V. On the Design of a Stepped Mixed Signal Transmission Line with Defected Microstrip Structure for High-Frequency PCB Traces // Journal of High-Frequency Communication Technologies. – 2024. – Vol. 2, No. 1. – P. 120–129.

43. Near end and far end crosstalk reduction in high-speed signaling channel with periodical spiral resonator defected microstriplines in a high performance printed circuit board / G. Arumugam, S.-K. Natarajan, R. Packianathan, M.-S. Karoui // IEICE Electronics Express. – 2025. – Vol. 22, No. 1. – Art. no. 20240268.

44. A Serpentine Guard Trace to Reduce the Far-End Crosstalk Voltage and the Crosstalk Induced Timing Jitter of Parallel Microstrip Lines / K. Lee, H.-B. Lee, H.-K. Jung et al. // IEEE Transactions on Advanced Packaging. – 2008. – Vol. 31, No. 4. – P. 809–817.

45. Dai X. Reduction of UWB Far-End Crosstalk in Microwave and Millimeter-Wave Band of Parallel Periodically Loaded Transmission Lines with Discontinuous Structured Guard Lines / X. Dai, W. Feng, W. Che // IEEE Transactions on Plasma Science. – 2020. – Vol. 48, No. 7. – P. 2372–2383.

46. Dai X. Investigation and Improvements of UWB Microstrip Delay Lines / X. Dai, W. Feng, W. Che // IEEE Transactions on Components, Packaging and Manufacturing Technology. – 2021. – Vol. 11, No. 8. – P. 1292–1300.

47. Wu C.H. An Application of the Subwavelength Periodic Microstrip Guard Trace in High-Speed Circuits / C.H. Wu, P. Ma, G. Zhou et al. // IEEE Access. – 2022. – Vol. 10. – P. 42640–42655.

48. Shiue G.-H. Guard Trace Design for Improvement on Transient Waveforms and Eye Diagrams of Serpentine Delay Lines / G.-H. Shiue, C.-Y. Chao, R.-B. Wu // IEEE Transactions on Advanced Packaging. – 2010. – Vol. 33, No. 4. – P. 1051–1060.

49. Tani L. Minimizing crosstalk on printed circuit board using non uniform guard traces / L. Tani, N. El Ouazzani // 2016 International Conference on Information Technology for Organizations Development (IT4OD). – Fez, Morocco: IEEE, 2016. – P. 1–4.

50. Mbairi F.D. On The Problem of Using Guard Traces for High Frequency Differential Lines Crosstalk Reduction / F.D. Mbairi, W.P. Siebert, H. Hesselbom // IEEE Transactions on Components and Packaging Technologies. – 2007. – Vol. 30, No. 1. – P. 67–74.

51. Raj A. Investigations on Crosstalk Reduction in Coupled Microstrip Transmission Line / A. Raj, N. Gupta // IEEE Electromagnetic Compatibility Magazine. – 2024. – Vol. 13, No. 3. – P. 66–73.

52. Abbasi H.R. Cross□talk reduction in metamaterial and non-uniform transmission lines / H.R. Abbasi, J.A. Rashed-Mohassel, M.N. Moghadasi // IET Microwaves, Antennas & Propagation. – 2024. – Vol. 18, No. 6. – P. 459–466.

53. Zhang Y. High Signal Integrity Interconnects Embedded with Metasurface for Far-End Crosstalk Reduction and High-Speed Data Transfer / Y. Zhang, G. Wang // IEEE Transactions on Signal and Power Integrity. – 2024. – Vol. 3. – P. 212–218.

54. Azhagumurugan R. A novel metamaterial structure to reduce far-end crosstalk in printed circuit boards / R. Azhagumurugan, P. Indumathi // International Journal of Engineering & Technology. – 2014. – Vol. 6, No. 2. – P. 588–591.

55. Munir A. Performance enhancement of coupled microstrip lines using split ring resonators / A. Munir, L.R.G. Soeka // 2014 International Conference on Electrical Engineering and Computer Science (ICEECS). – Kuta, Bali, Indonesia: IEEE, 2014. – P. 385–388.

56. Газизов Р.Р. Локализация максимумов сигнала в многопроводных линиях передачи печатных плат с помощью системы TALGAT / Р.Р. Газизов, А.М. Заболоцкий, П.Е. Орлов // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2015. – № 4 (38). – С. 147–150.

57. Пат. 2579915 РФ, МПК H 04 B 3/32. Способ подавления перекрестных помех в многопроводных линиях связи / В.Н. Федоров [и др.]. – № 2014116590/07; заявл. 25.04.2014; опубл. 10.04.2016, Бюл. № 30. – 12 с.

58. Пат. 89293 РФ, МПК H 02 H 3/00. Способ подавления перекрестных помех в многопроводных линиях связи / В.Н. Федоров [и др.]. – № 2009126800/22; заявл. 15.07.2009; опубл. 27.11.2009, Бюл. № 33. – 22 с.

59. Crosstalk Suppression Method at the End of Multiconductor Transmission Lines Based on Crosstalk Factor Theory / B. Yin, D. Cheng, R. Wang et al. // IEEE Transactions on Components, Packaging and Manufacturing Technology. – 2026. – Vol. 16, No. 4. – P. 861–871.

60. Simulation for a Crosstalk Canceling Algorithm in Lossless Multiconductor Transmission Lines / J. Zhou et al. // IEEE Letters on Electromagnetic Compatibility Practice and Applications. – 2021. – Vol. 3, No. 2. – P. 52–55.

61. Broydé F. A new method for the reduction of crosstalk and echo in multiconductor interconnections / F. Broydé, E. Clavelier // IEEE Transactions on Circuits and Systems I: Regular Papers. – 2005. – Vol. 52, No. 2. – P. 405–416.

62. Coskun S. Manufacturing and design practices: a review of crosstalk reduction in electronic circuit / S. Coskun,

M. Yildiz, T. Sonmezocak // International Journal of Electronics and Telecommunications. – 2025. – Vol. 71, No. 1. – P. 113–121.

63. Briones M.A. Transmission Lines in Modern Communication Systems: A Systematic Review / M.A. Briones, A.S. Fung, E.R. Arboleda // International Journal of Research Publication and Reviews. – 2024. – Vol. 5, No. 7. – P. 724–735.

Мурманский Михаил Семенович
Аспирант каф. сверхвысокочастотной и квантовой радиотехники (СВЧиКР) Томского гос. ун-та систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0002-2072-6891
Тел.: +7-923-443-90-10
Эл. почта: mikhail.s.murmanskii@tusur.ru

Томских Алёна Евгеньевна
Студентка каф. радиотехнических систем (РТС) ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0009-0006-1599-2892
Тел.: +7-950-434-66-82
Эл. почта: timskikhalyona@gmail.com

Жечев Евгений Сергеевич
Канд. техн. наук, доцент каф. СВЧиКР ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0003-4469-7033
Тел.: +7-913-102-34-89
Эл. почта: evgenii.zhechev@tusur.ru

Суровцев Роман Сергеевич
Д-р техн. наук, проф. каф. СВЧиКР ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0001-6609-0449
Тел.: +7 (382-2) 41-34-39
Эл. почта: surovtsevrs@gmail.com

Скорняков Иван Анатольевич
Аспирант каф. телевидения и управления (ТУ) ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0003-4464-5380
Тел.: +7 (382-2) 41-34-39
Эл. почта: ivan_sk94@mail.ru

Поступила в редакцию: 21.04.2026.
Принята к публикации: 14.09.2026.

Murmansky M.S., Tomskikh A.E., Zhechev Y.S., Surovtsev R.S., Skornyakov I.A.

Crosstalk suppression methods in coupled transmission lines: a review

Relevance. The rapid development of high-speed electronics and increasing printed circuit board (PCB) packaging density have made crosstalk one of the major obstacles to maintaining signal transmission quality. ***Purpose of the study.*** The study aims to systematize and review modern structural approaches to minimizing crosstalk in microstrip transmission lines in order to identify the most effective design strategies. ***Methods.*** The study employs methods of scientific literature analysis, comparative analysis of the physical mechanisms underlying crosstalk suppression, and quantitative evaluation of the characteristics of various topological solutions. ***Scientific novelty.*** A generalized classification of crosstalk mitigation techniques is proposed based on the underlying physical suppression mechanism, including compensation of differences in coupling coefficients, energy diversion to a reference conductor, and local correction of per-unit-length parameters. The effectiveness of hybrid and advanced high-technology structures is also analyzed. ***Results.*** It is demonstrated that classical routing guidelines become less effective at frequencies above 10 GHz, whereas innovative solutions based on periodic structures and metasurfaces can provide crosstalk suppression of up to −50 dB. Key regularities governing the effects of defects in the reference layer and passive conductors on characteristic impedance and modal phase velocities are identified, and the applicability limits of each of the reviewed methods are established. ***Practical significance.*** The review results serve as a technical guide for microwave hardware design engineers, enabling an informed selection of a specific interconnect topology depending on the required level of electromagnetic interference immunity, available PCB area, and manufacturing process complexity.

Funding: This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation (Project No. FEWM-2026-0013).

Keywords: crosstalk, coupled transmission lines, printed circuit boards, signal integrity, protective conductor, defected ground structures, metamaterials, metasurfaces.

DOI: 10.21293/1818-0442-2026-29-2-24-38

References

1. Wu T.-L., Buesink F., Canavero F. Overview of signal integrity and EMC design technologies on PCB: Fundamentals and latest progress. *IEEE Transactions on Electromagnetic Compatibility*, 2013, vol. 55, no. 4, pp. 624–638.

2. X.-B. Yu et al. Far-End Crosstalk Mitigation for Transmission Lines in DDR5 Using Glass-Weave Coating Structure. 2021 IEEE USNC-URSI Radio Science Meeting (Joint with AP-S Symposium), Singapore, Singapore, IEEE, 2021, pp. 23–24.

3. Bogatin E. Signal and Power Integrity – Simplified, New Jersey, U.S., Prentice Hall, 2010, 757 p.

4. Kuo H.-C., Pan P.-C., Hung L.-C., Jhong M.-F., Wang C.-C. Worst Eye Performance Analysis for Advanced Package Die-to-Die Interconnects. 2025 IEEE 29th Workshop on Signal and Power Integrity (SPI), Gaeta, Italy, IEEE, 2025, pp. 1–4.

5. Lu X., Song L., Xu H., Su D. Study on the Impact of Imbalance Between Transmission Lines on Crosstalk: A Novel Perspective of Displacement Current. *Chinese Journal of Electronics*, 2025, vol. 34, no. 1, pp. 114–124.

6. Shen Y., Su Z., Liu C., Ma X., Wang L., Li B. A 32 Gb/s 0.41 pJ/bit Single-Ended Transmitter with TX-Based-Only Adaptive Crosstalk Cancellation for Ultra-Short-Reach Wireline Applications. *Applied Sciences*, 2026, vol. 16, no. 11, art. no. 5672.

7. Krage M.K., Haddad G.I. Characteristics of coupled microstrip transmission lines-I: Coupled-mode formulation of inhomogeneous lines. *IEEE Transactions on Microwave Theory and Techniques*, 1970, vol. 18, no. 4, pp. 217–222.

8. DeFalco J.A. Reflection and crosstalk in logic circuit interconnections. *IEEE Spectrum*, 1970, vol. 7, no. 7, pp. 44–50.

9. Kechiev L.N. *Proektirovanie pechatnykh plat dlya tsifrovoy bystrodeystvuyushchey apparatury* [Design of printed circuit boards for digital high-speed equipment], Moscow, DT Group LLC, 2007, 615 p (in Russ.).

10. Mbairi F.D., Siebert W.P., Hesselbom H. High-Frequency Transmission Lines Crosstalk Reduction Using Spacing Rules. *IEEE Transactions on Components and Packaging Technologies*, 2008, vol. 31, no. 3, pp. 601–610.

11. Novak I., Eged B., Hatvani L. Measurement by vector-network analyzer and simulation of crosstalk reduction on printed circuit boards with additional center traces. 1993 IEEE

Instrumentation and Measurement Technology Conference, Irvine, CA, USA, IEEE, 1993, pp. 269–274.

12. Sychev A.N., Struchkov S.M. [Parameter systems of identical coupled lines with unbalanced electromagnetic coupling] *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2014, no. 1(31), pp. 39–50 (in Russ.).

13. Sychev A.N. [Analysis and synthesis of asymmetric coupled lines in a homogeneous dielectric medium]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2019, vol. 22, no. 1, pp. 11–19 (in Russ.).

14. Sychev A.N., Rudyy N.Yu. [Parameters of asymmetric coupled lines with inhomogeneous dielectrics]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2018, vol. 21, no. 4-1, pp. 7–15 (in Russ.).

15. Sychev A.N. [Synthesis of modal and distributed parameters of asymmetric coupled lines with inhomogeneous dielectrics]. *Journal of Radio Electronics*, 2024, no. 6, pp. 1–41 (in Russ.).

16. Gilb J.P., Balanis C.A. Pulse distortion on multilayer coupled microstrip lines. *IEEE Transactions on Microwave Theory and Techniques*, 1989, vol. 37, no. 10, pp. 1620–1628.

17. Surovtsev R.S., Gazizov T.R. Optimization of protective varnish thickness to minimize crosstalk in multiconductor bus of spacecraft PCB. Proc. of IEEE Moscow Workshop on Electronic and Networking Technologies (MWENT-2018), Moscow, Russia, IEEE, 2018, pp. 1–6.

18. Skornyakov I.A., Karri S., Surovtsev R.S. Mandel A.E. [Comparison of the effectiveness of cross-coupling reduction methods in PCB transmission lines]. *Doklady TUSUR*, 2025, vol. 28, no. 2, pp. 22–28 (in Russ.).

19. Liu Y. et al. Far-End Crosstalk Modeling and Prediction for Stripline with Inhomogeneous Dielectric Layers (IDLs). *IEEE Transactions on Signal and Power Integrity*, 2022, vol. 1, pp. 93–103.

20. Yong S., Cai K., Sen B. [et al.] A Comprehensive and Practical Way to Look at Crosstalk for Transmission Lines with Mismatched Terminals. 2018 IEEE Symposium on Electromagnetic Compatibility, Signal Integrity and Power Integrity (EMC, SI & PI), Long Beach, CA, USA, IEEE, 2018, pp. 538–543.

21. Yong S. [et al.] Prepreg and Core Dielectric Permittivity (ϵ_r) Extraction for Fabricated Striplines' Far-End Crosstalk Modeling. *IEEE Transactions on Electromagnetic Compatibility*, 2022, vol. 64, no. 1, pp. 209–218.

22. Park J., Guo Y., Ye X., Kim D. Extraction of DKglass and DKresin for PCB Striplines and Analysis of Far-End Crosstalk. *IEEE Transactions on Signal and Power Integrity*, 2025, vol. 4, pp. 149–155.

23. Gazizov T.R. Far-end crosstalk reduction in double-layered dielectric interconnects. *IEEE Transactions on Electromagnetic Compatibility*, 2001, vol. 43, no. 4, pp. 566–572.

24. Sutono A., Ye X. Dual-stripline with crosstalk cancellation. Pat. US 2022/0311114 A1.; Intel Corporation, no. 17/214,111, 2022.

25. Ge J., Floyd R., Khan A., Wang G. High-Performance Interconnects with Reduced Far-End Crosstalk for High-Speed ICs and Communication Systems. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2023. vol. 13, no. 7, pp. 1013–1020.

26. Jiang W., Jiang W., Cai X.-D., Sen B., Wang G. Equation-based solutions to coupled, asymmetrical, lossy, and nonuniform microstrip lines for tab-routing applications. *IEEE Transactions on Electromagnetic Compatibility*, 2019, vol. 61, no. 2, pp. 548–557.

27. Lho D. et al. Design and Analysis of HDMI 2.1 Connector for Crosstalk Reduction using Tabs and Inverse Tabs. 2021 IEEE 30th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS), San Jose, CA, USA, IEEE, 2021, pp. 1–3.

28. Datta S. et al. Analysis of Differential Stripline Routing Approaches within a PCB Via Field for Crosstalk Mitigation. 2023 IEEE 32nd Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS), Milpitas, CA, USA, IEEE, 2023, pp. 1–3.

29. Patel A., Pandey G., Verma G. [et al.] Homogeneous, lossy and Asymmetrical Strip Line with tabbed design for High-Speed Interface application. *ITM Web of Conferences*, 2023, vol. 54, art. no. 02007.

30. Zhang Y., Cai X-D., [et al.] Analysis Method for Curved Coupled Stripline with Triangular Tabs Within Pin Field of High-Speed Channels. 2025 IEEE International Symposium on Electromagnetic Compatibility, Signal & Power Integrity (EMC+SIPI), Raleigh, NC, USA, IEEE, 2025, pp. 338–342.

31. Mai C.-J. [et al.] Reinforcement Learning-Based Design of Interconnects with Reduced Far-End Crosstalk. *IEEE Transactions on Electromagnetic Compatibility*, 2025, vol. 67, no. 4, pp. 1370–1374.

32. Tsuchiya A. [et al.] A low loss microstrip line on thin flexible substrate film by defected ground structure. *IEICE Communications Express*, 2023, vol. 12, no. 11, pp. 582–585.

33. Sindhadevi M., Malathi K., Henridass A. Crosstalk Reduction Using Defective Ground Plane Structures in RF Printed Circuit Boards. *Arabian Journal for Science and Engineering*, 2014, vol. 39, pp. 1107–1116.

34. Luo X., Yang K., Li L., Liu H. A SSPP-Integrated Defected Ground Structure with Low EMI. *IEEE Letters on Electromagnetic Compatibility Practice and Applications*, 2025, vol. 8, no. 1, pp. 7–10.

35. Balan T.S., Devadas B., Srivastava G.K. Suppression of Common and Differential Component of Crosstalk Coupling onto the Differential Line and Field analysis of the Crosstalk. 2018 IEEE Symposium on Electromagnetic Compatibility, Signal Integrity and Power Integrity (EMC, SI & PI), Long Beach, CA, USA, IEEE, 2018, pp. 260–264.

36. Kazerooni M., Cheldavi A., Khalil M. Unit Length Parameters, Transition Sharpness and Level of Radiation in Defected Microstrip Structure (DMS) and Defected Ground Structure (DGS) Interconnections. *Progress in Electromagnetics Research M*, 2009, vol. 10, pp. 93–102.

37. Zhang Y., Wang G. UWB Far-End Crosstalk Mitigation With «LL» Shaped Defected Tabbed Routing Structures. *IEEE Transactions on Electromagnetic Compatibility*, 2025, vol. 67, no. 1, pp. 247–253.

38. Baghel L.K., Kumar S., Sis S.A., Kim J. Crosstalk Reduction in Coupled Microstrip Lines using TT-shaped DMS Approach. 2023 Joint Asia-Pacific International Symposium on Electromagnetic Compatibility and International Conference on ElectroMagnetic Interference & Compatibility (APEMC/INCEMIC), Bengaluru, India, IEEE, 2023, pp. 1–4.

39. Liu X., Li Y., Zhao Y. [et al.] Equivalent Circuit Model of Crosstalk Reduction Parallel Transmission Lines with Defected Microstrip Structures. 2018 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC), Xuzhou, China, IEEE, 2018, pp. 1–2.

40. Gobinath A., Kumar N.S., Rajeswari P. Reduction of near-end and far-end crosstalk in microwave and millimetre wave band of parallel transmission lines using meander shaped DMS. *Journal of Physics: Conference Series*, 2023, vol. 2466, no. 1, pp. 012016.

41. Li C., Wang Y., Li X. Crosstalk reduction between microstrip lines using TL-shaped defected microstrip structure. *Engineering Reports*, 2023, vol. 5, no. 1, pp. e12559.

42. Yokesh V. On the Design of a Stepped Mixed Signal Transmission Line with Defected Microstrip Structure for High-Frequency PCB Traces. *Journal of High-Frequency Communication Technologies*, 2024, vol. 2, no. 1, pp. 120–129.

43. Arumugam G., Natarajan S.-K., Packianathan R., Karoui M.-S. Near end and far end crosstalk reduction in high-speed signaling channel with periodical spiral resonator defected microstriplines in a high performance printed circuit board. *IEICE Electronics Express*, 2025, vol. 22, no. 1, pp. 20240268.

44. Lee K., Lee H.-B., Jung H.-K. [et al.] A Serpentine Guard Trace to Reduce the Far-End Crosstalk Voltage and the Crosstalk Induced Timing Jitter of Parallel Microstrip Lines. *IEEE Transactions on Advanced Packaging*, 2008, vol. 31, no. 4, pp. 809–817.

45. Dai X., Feng W., Che W. Reduction of UWB Far-End Crosstalk in Microwave and Millimeter-Wave Band of Parallel Periodically Loaded Transmission Lines with Discontinuous Structured Guard Lines. *IEEE Transactions on Plasma Science*, 2020, vol. 48, no. 7, pp. 2372–2383.

46. Dai X., Feng W., Che W. Investigation and Improvements of UWB Microstrip Delay Lines. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2021. vol. 11, no. 8, pp. 1292–1300.

47. Wu C.H., P. Ma, G. Zhou, [et al.] An Application of the Subwavelength Periodic Microstrip Guard Trace in High-Speed Circuits. *IEEE Access*, 2022, vol. 10, pp. 42640–42655.

48. Shiue G.-H., Chao C.-Y., Wu R.-B. Guard Trace Design for Improvement on Transient Waveforms and Eye Diagrams of Serpentine Delay Lines. *IEEE Transactions on Advanced Packaging*, 2010, vol. 33, no. 4, pp. 1051–1060.

49. Tani L., El Ouazzani N. Minimizing crosstalk on printed circuit board using non uniform guard traces. 2016 International Conference on Information Technology for Organizations Development (IT4OD), Fez, Morocco: IEEE, 2016, pp. 1–4.

50. Mbairi F.D., Siebert W.P., Hesselbom H. On The Problem of Using Guard Traces for High Frequency Differential Lines Crosstalk Reduction. *IEEE Transactions on Components and Packaging Technologies*, 2007, vol. 30, no. 1, pp. 67–74.

51. Raj A., Gupta N. Investigations on Crosstalk Reduction in Coupled Microstrip Transmission Line. *IEEE Electromagnetic Compatibility Magazine*, 2024, vol. 13, no. 3, pp. 66–73.

52. Abbasi H.R., Rashed- Mohassel J.A., Moghadasi M.N. Cross- talk reduction in metamaterial and non- uniform transmission lines. *IET Microwaves, Antennas & Propagation*, 2024, vol. 18, no. 6, pp. 459–466.

53. Zhang Y., Wang G. High Signal Integrity Interconnects Embedded with Metasurface for Far-End Crosstalk Reduction and High-Speed Data Transfer. *IEEE Transactions on Signal and Power Integrity*, 2024, vol. 3, pp. 212–218.

54. Azhagumurugan R., Indumathi P. A novel metamaterial structure to reduce far-end crosstalk in printed circuit boards. *International Journal of Engineering & Technology*, 2014, vol. 6, no. 2, pp. 588–591.

55. Munir A., Soeka L.R.G. Performance enhancement of coupled microstrip lines using split ring resonators. 2014 International Conference on Electrical Engineering and Computer Science (ICEECS), Kuta, Bali, Indonesia, IEEE, 2014, pp. 385–388.

56. Gazizov R. R., Zabolotskiy A. M., Orlov P. E. Signal maximum localization in multiconductor transmission lines of printed circuit boards using TALGAT system. *Doklady TUSUR*, 2015, no. 4(38), pp. 147–150 (in Russ.).

57. Fedorov V.N. [et al.]. Pat. 2579915 RF. *Sposob podavleniya perekrestnykh pomekh v mnogoprovodnykh liniyakh svyazi* [Method for suppressing crosstalk in multiconductor communication lines], no. 2014116590/07, 2016 (in Russ.).

58. Fedorov V.N. [et al.]. Pat. 89293 RF. *Sposob podavleniya perekrestnykh pomekh v mnogoprovodnykh liniyakh svyazi* [Method for suppressing crosstalk in multiconductor communication lines], no. 2009126800/22, 2009 (in Russ.).

59. Yin B., Cheng D., Wang R. [et al.] Crosstalk Suppression Method at the End of Multiconductor Transmission Lines Based on Crosstalk Factor Theory. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2026, vol. 16, no. 4, pp. 861–871.

60. Zhou J. [et al.] Simulation for a Crosstalk Canceling Algorithm in Lossless Multiconductor Transmission Lines. *IEEE Letters on Electromagnetic Compatibility Practice and Applications*, 2021, vol. 3, no. 2, pp. 52–55.

61. Broydé F., Clavelier E. A new method for the reduction of crosstalk and echo in multiconductor interconnections. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2005, vol. 52, no. 2, pp. 405–416.

62. Coskun S., Yildiz M., Sonmezocak T. Manufacturing and design practices: a review of crosstalk reduction in electronic circuit. *International Journal of Electronics and Telecommunications*, 2025, vol. 71, pp. 113–121.

63. Briones M.A., Fung A.S., Arboleda E.R. Transmission Lines in Modern Communication Systems: A Systematic Review. *International Journal of Research Publication and Reviews*, 2024, vol. 5, no. 7, pp. 724–735.

Mikhail S. Murmansky
Postgraduate student, Department of Microwave and Quantum Radio Engineering, Tomsk State University
of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0002-2072-6891
Phone: +7-923-443-90-10
Email: mikhail.s.murmanskii@tusur.ru

Alena E. Tomskikh
Student, Department of Radio Engineering Systems, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0009-0006-1599-2892
Phone: +7-950-434-66-82
Email: timskikhalyona@gmail.com

Yevgeniy S. Zhechev
Candidate of Technical Sciences, Associate Professor, Department of Microwave and Quantum Radio Engineering, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0003-4469-7033
Phone: +7-913-102-34-89
Email: evgenii.zhechev@tusur.ru

Roman S. Surovtsev
Doctor of Engineering Sciences, Professor, Department of Microwave and Quantum Radio Engineering, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0001-6609-0449
Phone: +7 (382-2) 41-34-39
Email: surovtsevrs@gmail.com

Ivan A. Skornyakov
Post-graduate student, Department of Television and Control, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0003-4464-5380
Phone: +7 (382-2) 41-34-39
Email: ivan_sk94@mail.ru

Received: 21.04.2026.
Accepted: 14.09.2026.

УДК 621.396.6
Шифр специальности ВАК: 2.2.14

Ф.А. Зотов, А.А. Меркулов, А.А. Цитович

Проектирование вторичного источника питания для АФАР Ка-диапазона

Актуальность. Разработка компактных и энергоэффективных источников питания с низким уровнем электромагнитных помех является ключевой задачей при создании современных приемопередающих модулей активных фазированных антенных решеток (АФАР) Ка-диапазона. ***Цель исследования.*** Разработка и обоснование схемотехнических решений для построения вторичного источника питания АФАР Ка-диапазона. ***Методы.*** В работе применялись методы схемотехнического проектирования, многокритериального выбора компонентной базы на основе весовых коэффициентов, а также Spice-моделирование режимов работы электрических цепей с учетом паразитных параметров компонентов. ***Научная новизна.*** Предложена и обоснована двухуровневая распределенная архитектура питания с физическим разделением каналов промежуточной шины и применением высокочастотных синхронных преобразователей в инвертирующем включении, что минимизирует влияние помех через общий импеданс на чувствительные СВЧ-цепи. ***Результаты.*** Разработана микроконтроллерная система алгоритмического управления циклограммой запуска, предотвращающая электрический пробой СВЧ-транзисторов. Результаты моделирования подтвердили обеспечение КПД силовых каскадов 86–91% при выходных токах до 15 А, при этом амплитуда пульсаций на низковольтных каналах не превышает 15 мВ. ***Практическая значимость.*** Предложенные схемотехнические решения могут быть внедрены при проектировании бортовой аппаратуры и малогабаритных радиолокационных комплексов для обеспечения высокой надежности и электромагнитной совместимости плотно упакованных антенных модулей.
Ключевые слова: вторичный источник питания, АФАР, DC/DC-преобразователь, электромагнитная совместимость, схемотехническое моделирование, циклограмма, микроконтроллер, энергоэффективность.
DOI: 10.21293/1818-0442-2026-29-2-39-47

Введение

Развитие современных радиолокационных систем неразрывно связано с переходом к АФАР миллиметрового (Ka) и сантиметрового (X, Ku) диапазонов волн [1–3]. Переход в область высоких частот неизбежно влечет за собой уменьшение габаритных размеров излучающих элементов и критическое сокращение шага антенной решетки. Это приводит к резкому повышению плотности компоновки приемопередающих модулей (ППМ) и росту удельного тепловыделения на единицу объема [4–6].

Современные ППМ представляют собой сложные гибридные устройства, включающие высокоразрядные АЦП/ЦАП, программируемые логические интегральные схемы (ПЛИС) и прецизионные СВЧ-усилители [7]. Функционирование такой компонентной базы требует формирования ряда разнополярных напряжений питания с жесткими допусками по амплитуде пульсаций. При этом традиционные подходы к построению вторичных источников питания (ВИП), например, использование крупногабаритных серийных модулей (VPX/VITA 62) [8–11], становятся малоэффективными из-за избыточных массогабаритных характеристик, ограниченной гибкости в настройке выходных номиналов и высокого уровня кондуктивных помех, требующих установки громоздких внешних LC-фильтров [12, 13].

Особую сложность при проектировании ВИП для АФАР представляет выбор частоты коммутации силовых ключей. Для минимизации габаритов фильтрующих компонентов необходимо повышение частоты преобразования до единиц мегагерц [14]. При этом необходимо обеспечить конструктивное ослабление генерируемых электромагнитных помех для исключения их влияния на промежуточные частоты приемного тракта радиолокатора, что накладывает ограничения на топологию печатной платы и требует применения экранирования [15, 16].

Кроме того, для предотвращения электрического пробоя затворов СВЧ-транзисторов при включении ППМ необходимо строгое соблюдение временной циклограммы подачи питающих напряжений [17]. Нарушение последовательности включения стоковых и затворных цепей приводит к необратимому выходу из строя дорогостоящих СВЧ-компонентов.

Таким образом, **цель работы** – представить разработку малогабаритной распределенной бестрансформаторной архитектуры ВИП, сочетающей в себе высокий КПД, эффективное подавление помех и наличие интеллектуальной системы управления запуском.

Архитектура распределенной системы питания

Для достижения поставленной цели разработана двухуровневая распределенная архитектура ВИП (рис. 1). Выбранная топология позволяет отказаться от больших изолированных DC/DC-преобразователей на каждом выходном канале, перенеся основную тепловую и токовую нагрузку на первый уровень преобразования [18, 19].

Первый уровень преобразования обеспечивает формирование из первичной бортовой сети 27 В базовых промежуточных шин. Ключевой особенностью предложенной архитектуры является физическое разделение промежуточной шины положительной по-

лярности на два независимых канала. Данное решение реализовано на базе двух отечественных изолированных импульсных модулей серии МНМ15 [20].

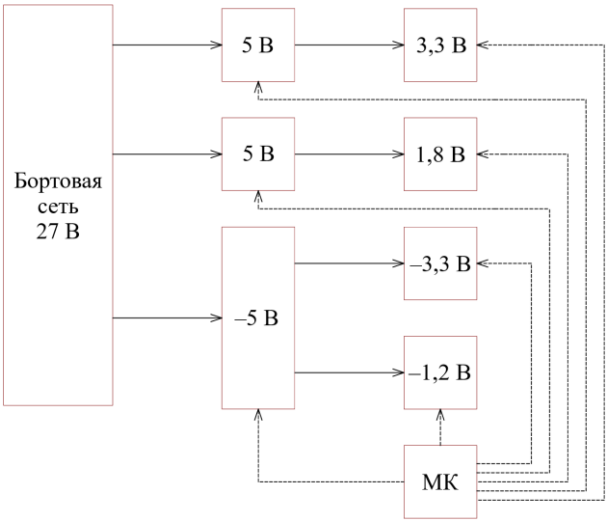


Рис. 1. Структурная схема ВИП

Разделение каналов DC/DC-модулей позволяет гальванически изолировать цепи питания цифровой логики ПЛИС от цепей питания чувствительных аналоговых каскадов и системы управления, что значительно снижает влияние помех через общий импеданс по цепям питания.

Вторая ветвь первого уровня предназначена для формирования отрицательного напряжения –5 В. Для минимизации габаритов применена синхронная микросхема LTC7149, работающая в топологии инвертирующего понижающе-повышающего преобразователя (Inverting Buck-Boost) [21]. Данный подход позволяет получить отрицательный потенциал с высоким КПД без применения многообмоточных трансформаторов [14, 18].

Второй уровень системы состоит из набора независимых высокочастотных синхронных понижающих преобразователей, формирующих конечные напряжения непосредственно вблизи потребителей: +3,3 В (ADP2386), +1,8 В (L6983). Питание преобразователей отрицательной полярности (–1,2 и –3,3 В) осуществляется от промежуточной шины –5 В с использованием интегральных схем (ИС) (TPS62827 и TPS62130A) в инвертирующем включении [22, 23]. Принцип реализации данной топологии заключается в изменении точки привязки потенциалов. Вывод GND микросхемы подключается к входной шине, а вход питания – к системной земле (0 В). При этом разность потенциалов, приложенная к выводам питания контроллера, соответствует сумме входного и абсолютного значения выходного напряжения ($V_{in} + |V_{out}|$).

Методика многокритериального выбора компонентной базы

В рамках работы применена методика комплексной оценки качества интегральных преобразователей на основе весовых коэффициентов. Данный подход позволяет сравнить компоненты различных производителей, приведя их характеристики к единому интегральному показателю. Для каждой микросхемы или модуля рассчитывался итоговый показатель качества Q по формуле

$$Q = \sum_{i=1}^{n} (w_i \cdot p_i), \quad (1)$$

где w_i – весовой коэффициент i-го параметра; p_i – оценка i-го параметра по 10-балльной шкале; n – общее количество оцениваемых параметров.

Наибольшие веса присвоены коэффициенту полезного действия (0,3) и нагрузочной способности (0,25), так как они определяют тепловой режим модуля, что является главным критерием обеспечения высокой надежности работы аппаратуры АФАР [4, 19]. Также учитывались уровень генерируемых шумов (0,15), массогабаритные характеристики (0,1), рабочий диапазон температур (0,1), ценовая категория (0,05) и надежность (0,05). На основе расчетов по (1) для каждого каскада определен оптимальный компонент. Итоговый перечень выбранной элементной базы представлен в табл. 1.

Таблица 1

Перечень отобранных микросхем для проектируемого ВИП

Преобразование	Выбранная микросхема	Показатель качества (Q)
27 В → +5 В	АЕДОН МНМ15-1В0,85,0Т	8,35
5 В → +3,3 В	Analog Devices ADP2386	8,30
5 В → +1,8 В	STMicroelectronics L6983NQTR	9,30
27 В → –5 В	Linear Technology LTC7149	8,55
–5 В → –1,2 В	Texas Instruments TPS62827	8,30
–5 В → –3,3 В	Texas Instruments TPS62130A-Q1	8,40

Схемотехническая реализация каскадов преобразования

Одной из основных задач при проектировании системы питания для АФАР является формирование ряда отрицательных напряжений (–5; –3,3; –1,2 В) с высоким КПД в условиях ограниченного объема. В разработанном устройстве применено схемотехническое решение на основе синхронных понижающих ИС, работающих по топологии инвертирующего понижающе-повышающего преобразователя [14, 18].

Формирование маломощных каналов реализовано путем изменения точки привязки потенциалов. В то же время для формирования базового мощного канала –5 В применена специализированная микросхема LTC7149, имеющая встроенные цепи сдвига уровней, что позволяет сохранить привязку ее порта GND к системной земле, существенно упрощая трассировку и повышая надежность.

Расчет индуктивности, которая влияет на режим работы преобразователя и уровень пульсаций, производился, исходя из условия работы в режиме неразрывных токов и допустимого уровня пульсаций тока ΔI_L (выбрано 30% от тока нагрузки), по формуле

$$L = \frac{V_{in} \cdot |V_{out}|}{\Delta I_L \cdot f_{sw} \cdot (V_{in} + |V_{out}|)}, \quad (2)$$

где V_{in}, V_{out} – входное и выходное напряжения; f_{sw} – частота переключения; ΔI_L – пульсации тока дросселя.

Для реализации наиболее мощного отрицательного канала (–5 В) использована ИС LTC7149. Особенностью данной микросхемы является токовый метод задания выходного напряжения через вывод ISET, из которого вытекает прецизионный ток I_{SET} = 50 мкА. Номинал задающего резистора R_{25} рассчитывался следующим образом:

$$R_{25} = \frac{|V_{out-}|}{I_{SET}}, \quad (3)$$

где $V_{out–}$ – выходное отрицательное напряжение; I_{SET} – задающий ток 50 мкА.

Рабочая частота преобразователя f_{sw} задается резистором R_{22}, подключенным между выводами RT и SVOUT–. Для обеспечения сбалансированных показателей КПД и габаритов выбрана частота 500 кГц. Требуемый номинал резистора рассчитывается по формуле и составляет 200 кОм:

$$R_{22} = \frac{10^{11}}{f_{sw}} = \frac{10^{11}}{500 \cdot 10^3}, \quad (4)$$

где 10^{11} – коэффициент преобразования, заложенный в архитектуру внутреннего генератора ИС.

Особое внимание уделено проектированию выходных фильтров, так как пульсации напряжения напрямую влияют на фазовые шумы приемопередающего тракта [2, 16]. Применен комбинированный подход к построению емкостного фильтра. Компенсация динамических просадок обеспечивается полимерными конденсаторами, обладающими высокой удельной емкостью (100–150 мкФ). Фильтрация высокочастотных помех реализована на базе параллельно включенных многослойных керамических конденсаторов различной емкости (22; 10; 0,1 мкФ). Выбор данного решения обусловлен необходимостью пропускания значительных высокочастотных токов пульсаций и минимизацией совокупного ESR, что затруднительно при использовании компактных помехоподавляющих чип-фильтров. Использование конденсаторов с диэлектриком X7R/X5R позволяет минимизировать паразитную индуктивность (ESL) и снизить эквивалентное последовательное сопротивление (ESR) фильтра до единиц миллиом. Принципиальная схема реализации четвертого каскада представлена на рис. 2.

Реализация циклограммы

Одновременный запуск всех каскадов системы питания АФАР вызывает критические броски пускового тока, способные привести к просадке первичной сети 27 В и активации систем защиты. Кроме того, специфика СВЧ-усилителей (в частности, на базе нитрида галлия, GaN) требует строгой очередности подачи напряжений смещения [1, 15]. Для решения данной задачи разработана микроконтроллерная система алгоритмического управления напряжениями питания на базе 32-разрядного микроконтроллера (МК) с ядром ARM Cortex-M0+ (STM32G031G8) в виброустойчивом безвыводном корпусе UFQFPN-28.

Питание самого МК организовано от промежуточной шины, формируемой модулем входного каскада (включается автоматически при подаче 27 В). Снижение напряжения до 3,3 В, необходимых для работы логики МК, осуществляется с помощью линейного стабилизатора напряжения (DA8).

Реализация циклограммы запуска сопряжена с необходимостью согласования логических уровней разнополярных каскадов. Выходы управления (Enable) преобразователей положительной полярности подключены напрямую к портам ввода-вывода МК (GPIO). Однако для каскадов, формирующих напряжения –1,2 и –3,3 В, прямое подключение недопустимо.

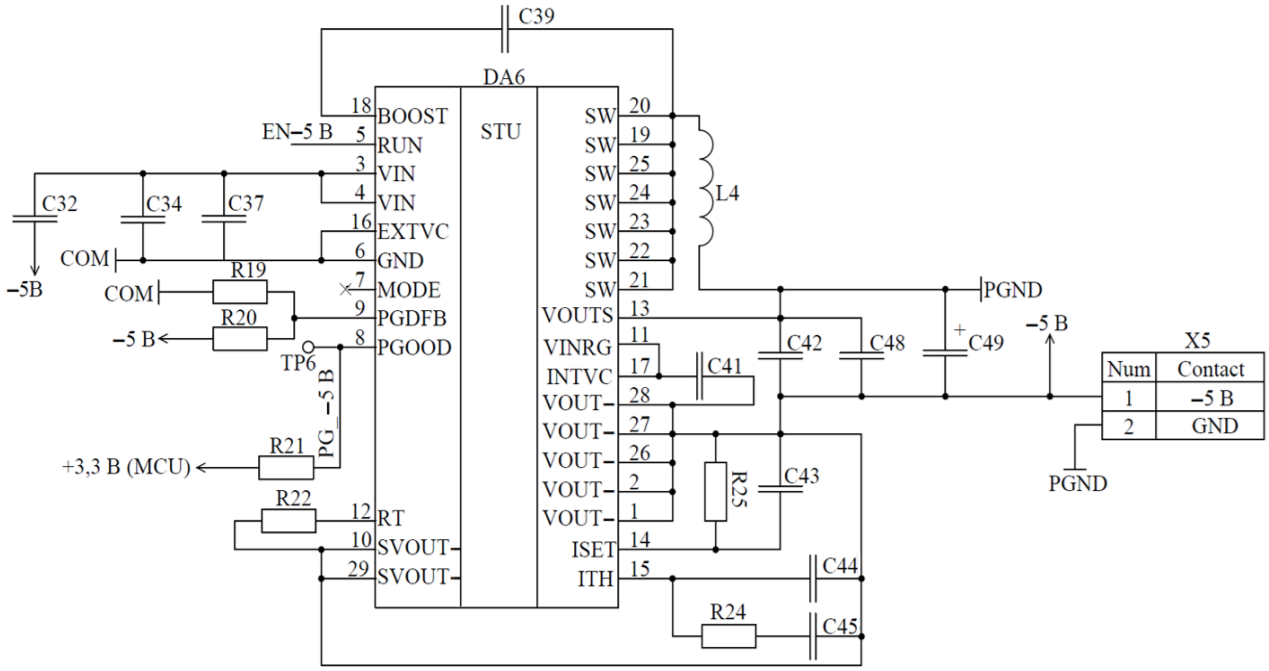


Рис. 2. Схема подключения четвертого каскада ВИП

Поскольку локальной землей данных микросхем является шина –5 В, подача управляющего логического сигнала +3,3 В от МК привела бы к возникновению разности потенциалов 8,3 В на выводе Enable.

Во избежание электрического пробоя в цепи управления инвертирующих каскадов внедрены резистивные делители напряжения. Они обеспечивают сдвиг управляющего сигнала в безопасный диапазон (около –0,85 В относительно системного нуля), что гарантированно включает преобразователи, не превышая допустимых пределов напряжения относительно их локальной земли. В качестве обратной связи для подтверждения выхода каскадов на номинальный режим используются сигналы статуса питания (Power Good). Выводы PG подтянуты к шине питания МК (+3,3 В). Для повышения отказоустойчивости и упрощения схемы сопряжения контроль запуска инвертирующих низковольтных каскадов реализован программным методом задержек. Это исключает необходимость трансляции сигналов с отрицательным смещением в порты МК, предотвращая протекание возвратных токов через защитные диоды портов ввода-вывода МК. На рис. 3 представлена схема подключения МК и его обвязки.

Результаты моделирования алгоритма управления, представленные на рис. 4, подтверждают реализацию детерминированной последовательности активации каналов.

Временная диаграмма демонстрирует трехступенчатый запуск. На первом этапе (0–8 мс) происходит стабилизация шин положительной полярности, обеспечивающих питание системы управления и промежуточных звеньев. На втором этапе (14 мс) активируется базовый инвертирующий каскад –5 В. Завершающая фаза циклограммы (28–30 мс) обеспечивает подачу напряжений смещения и питания цифровых портов ПЛИС. Выдержанные временные интервалы между включением каскадов позволяют избежать наложения переходных процессов, тем самым минимизируя суммарный импульсный ток потребления от первичной сети 27 В и гарантируя безопасные режимы работы затворных цепей СВЧ-транзисторов.

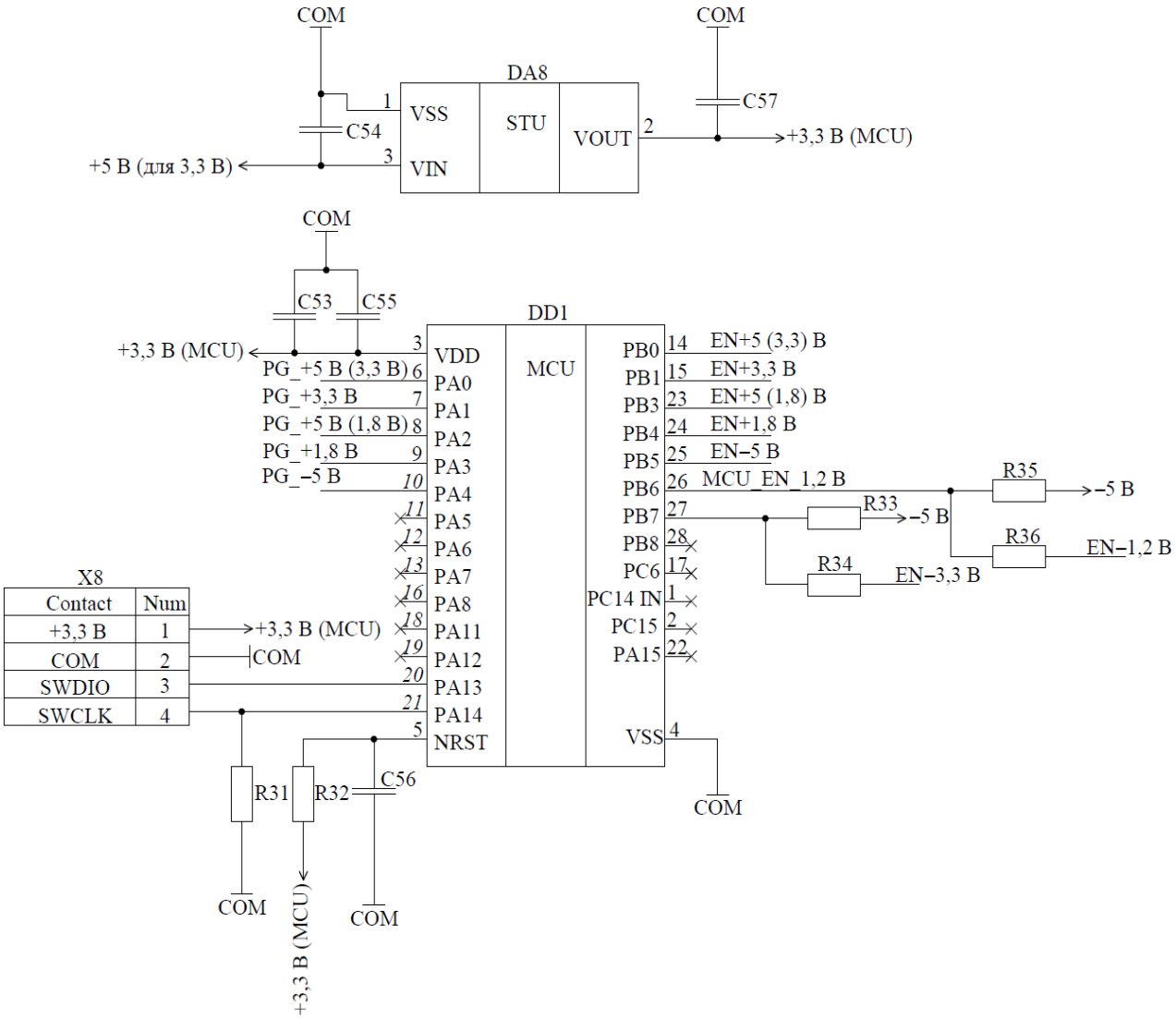


Рис. 3. Схема подключения МК STM32G031G8

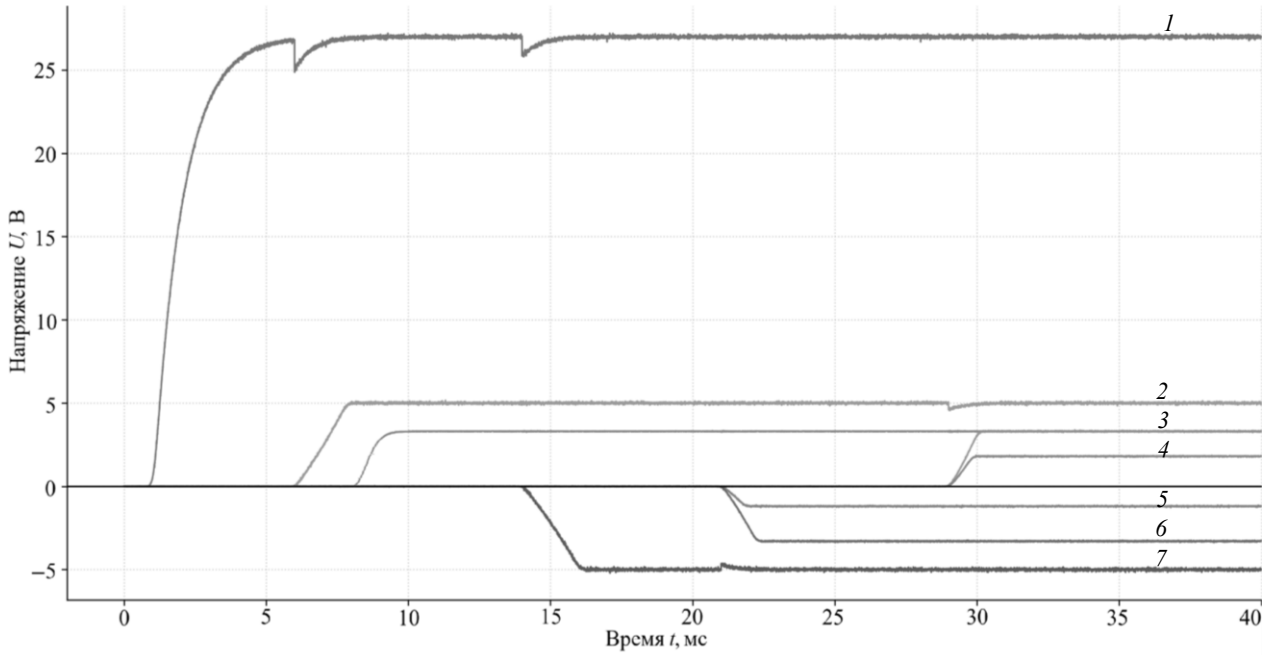


Рис. 4. График процесса последовательного включения каналов ВИП:
1 – шина +27 В; *2* – шина +5 В; *3* – шина +3,3 В; *4* шина – 1,8 В; *5* – шина –1,2 В; *6* – шина –3,3 В; *7* – шина –5 В

Моделирование динамических характеристик и переходных процессов

Подтверждение корректности подключения микросхем проводилось методом схемотехнического моделирования в программной среде LTspice. Модели пассивных компонентов задавались с учетом их паразитных параметров (ESR и ESL конденсаторов, активное сопротивление обмоток дросселей), что позволило максимально приблизить результаты к реальным условиям эксплуатации.

Анализ переходного процесса при включении (рис. 5) демонстрирует корректную работу функции плавного пуска.

Выходное напряжение достигает номинального значения за время около 3,5 мс без перерегулирования и опасных выбросов. Наблюдаемая на графике (см. рис. 5) широкая полоса значений тока дросселя в установившемся режиме обусловлена высокочастотными пульсациями с размахом около 1,3 А, детальная форма которых приведена на рис. 7.

Анализ установившегося режима при полной расчетной нагрузке подтвердил высокую эффективность выбранной комбинированной схемы фильтрации. На рис. 6 приведена осциллограмма пульсаций выходного напряжения. Амплитуда пульсаций составляет 56 мВ, что соответствует 1,1% от номинала (что для высоковольтного канала допустимо). Характерная форма пульсаций обусловлена ESR выходных конденсаторов, а частота следования соответствует частоте коммутации (500 кГц).

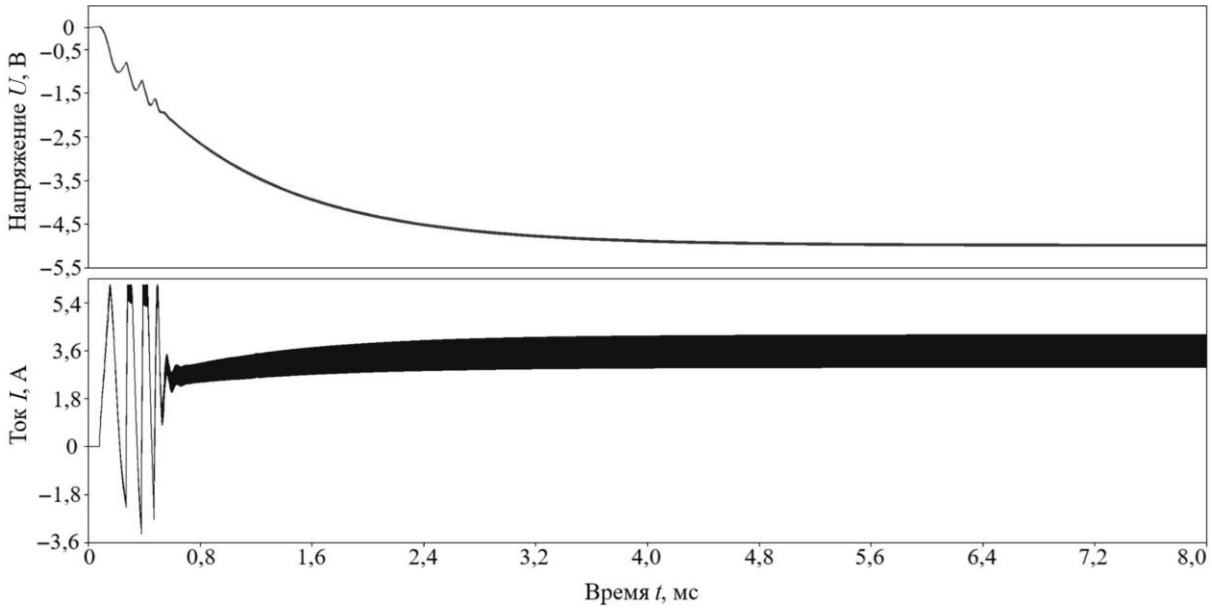


Рис. 5. Переходный процесс установления выходного напряжения и тока дросселя

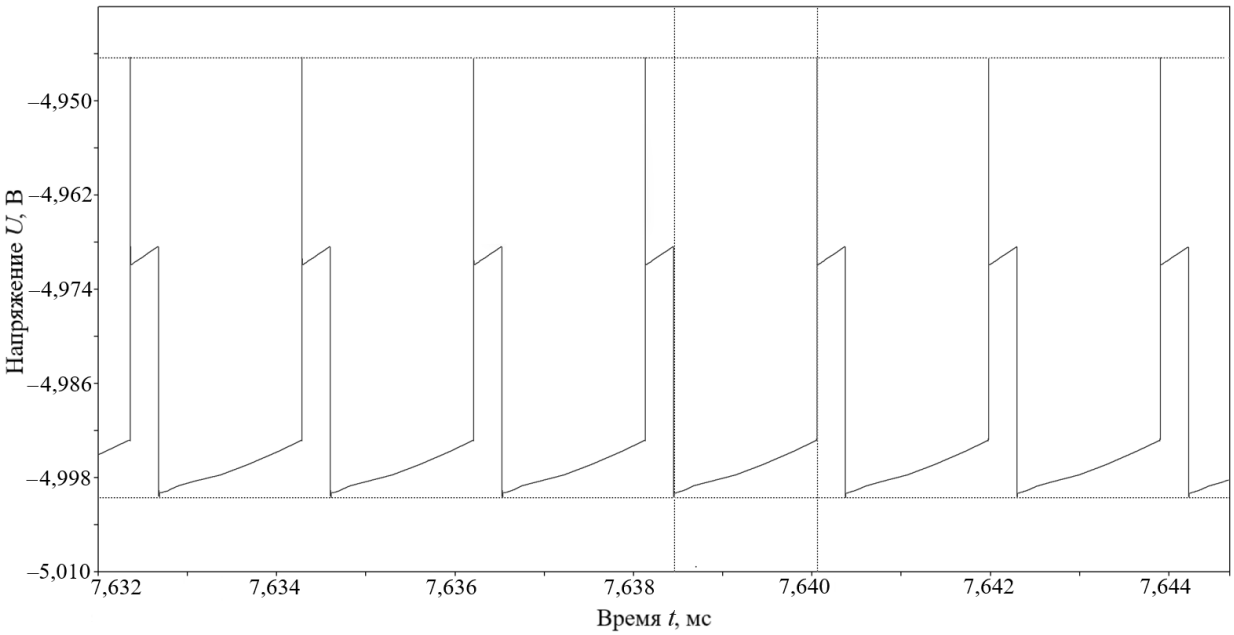


Рис. 6. Оценка пульсаций выходного напряжения

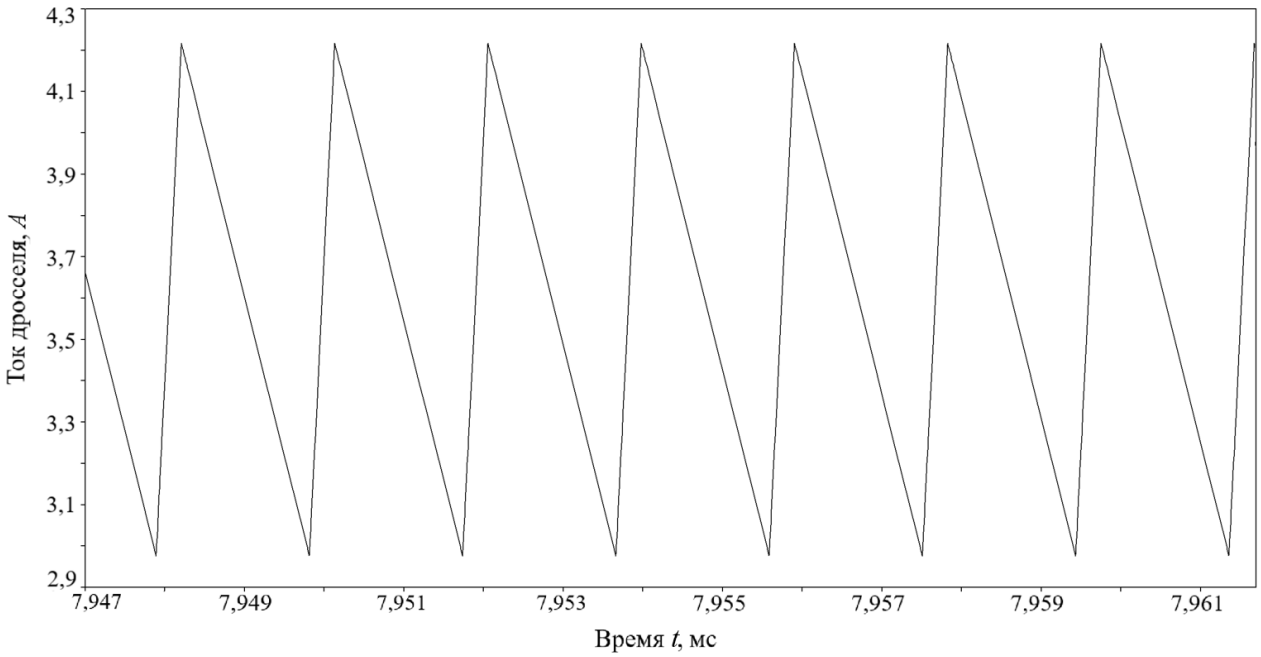


Рис. 7. Форма тока дросселя в режиме неразрывных токов

На рис. 7 представлена осциллограмма тока, протекающего через силовой дроссель каскада (–5 В). Форма тока имеет выраженный треугольный характер с постоянной составляющей, что подтверждает работу преобразователя в режиме неразрывных токов. Отсутствие искажений на вершинах треугольника свидетельствует о том, что пиковые значения тока не превышают ток насыщения выбранного дросселя, что гарантирует сохранение индуктивности и предотвращает неконтролируемый рост тока ключей.

Итоговые характеристики всех спроектированных каналов системы питания сведены в табл. 2. Оценка коэффициента полезного действия производилась в среде моделирования путем отношения средней выходной мощности к входной в установившемся режиме с учетом потерь на переключение ключей и активном сопротивлении обмоток дросселей.

Таблица 2

Сводные результаты моделирования параметров ВИП

$U_{вых}$, В	I_{max}, А	$\Delta U_{вых}$, мВ	δU, %	η, %
+5,0	15,0	45,0	0,9	89,5
+3,3	6,0	12,1	0,4	91,0
+1,8	3,0	9,0	0,5	90,5
–5,0	4,0	56,0	1,1	85,5
–3,3	3,0	10,6	0,3	89,0
–1,2	4,0	7,9	0,7	88,0

Заключение

В ходе работы спроектирована и промоделирована система вторичного электропитания для АФАР Ка-диапазона. Обоснован переход к распределенной двухуровневой архитектуре с промежуточной шиной +5 В, что обеспечило гибкость настройки каналов и исключило концентрацию всей рассеиваемой мощности в едином габаритном модуле питания. Разработана методика выбора компонентов на основе интегрального критерия качества, позволившая обоснованно применить отечественный модуль питания во входном каскаде и высокочастотные ИС в низковольтных каскадах стабилизации.

Предложены и реализованы схемотехнические решения для формирования ряда отрицательных напряжений (–5; –3,3; –1,2 В) на базе синхронных понижающих контроллеров в инвертирующем включении. Это позволило унифицировать элементную базу и отказаться от использования габаритных решений.

Разработанная микроконтроллерная система управления гарантирует строгое соблюдение требуемой циклограммы запуска каскадов. В отличие от традиционных решений на базе жесткой аппаратной логики, такой подход обеспечивает гибкость программной настройки временных интервалов без изменения топологии печатной платы.

Результаты схемотехнического моделирования подтвердили высокую эффективность предложенных решений: амплитуда пульсаций выходных напряжений для низковольтных каналов не превышает 15 мВ, обеспечен устойчивый запас по фазе в контурах регулирования. Достигнутые показатели энергоэффективности (КПД 86–91%) позволяют минимизировать перегрев плотно упакованных антенных модулей.

Литература

1. Белоус А.И. СВЧ-электроника в системах радиолокации и связи. Техническая энциклопедия. – Кн. 1 / А.И. Белоус, М.К. Мерданов, С.В. Шведов. – М.: Техносфера, 2021. – 724 с.

2. Батаев В.Я. Разработка излучающего полотна АФАР X-диапазона на основе печатного вибраторного излучателя / В.Я. Батаев, В.В. Демшевский, А.А. Цитович // Электронная техника. – Сер. 1: СВЧ-техника. – 2019. – № 3 (542). – С. 76–87.

3. Комплексированная антенная система X- и Ku-диапазонов / В.В. Кузнецов, А.А. Цитович, А.Д. Кучмий, А.А. Измайлов // Электроника и микроэлектроника СВЧ. – 2024. – Т. 1. – С. 309–313.

4. Глушко Е.В. Высокоплотный импульсный источник вторичного электропитания активных фазированных антенных решёток / Е.В. Глушко, Г.С. Юрченко, Д.А. Петрищев // Актуальные проблемы физики, электроники и энергетики : сб. ст. I Междунар. науч.-практ. конф., Новополоцк, 27–28 окт. 2022 г. / Полоцкий гос. ун-т им. Евфросинии Полоцкой. – Новополоцк, 2023. – С. 97–104.

5. Ненартович Н.Э. Из практики разработки активных фазированных антенных решеток / Н.Э. Ненартович, М.Б. Митяшев // Вестник МГТУ МИРЭА. – 2014. – № 3 (4). – С. 173–188.

6. Панеева А.П. Приёмопередающий модуль для активной фазированной антенной решетки малого космического аппарата / А.П. Панеева, Е.В. Комиссарова // Журнал радиоэлектроники. – 2024. – № 2. – С. 1–12.

7. Карасев М.С. Оперативный контроль электрических параметров приемопередающих модулей X-диапазона частот // Электронная техника. – Сер. 1: СВЧ-техника. – 2021. – № 3 (550). – С. 6–14.

8. Vicor Corporation. VIT028x6U1000y000 6U VITA 62 Military DC-DC Converter: Datasheet [Электронный ресурс]. – URL: https://www.vicorpower.com, свободный (дата обращения: 06.02.2026).

9. SynQor Inc. VPX-3U-DC28P-004 Vita 62 3U VPX Power Supply: Datasheet [Электронный ресурс]. – URL: https://www.synqor.com, свободный (дата обращения: 10.02.2026).

10. Behlman Electronics Inc. VPXtra 500M 3U VPX DC to DC Power Supply: Datasheet [Электронный ресурс]. – URL: https://www.behlman.com, свободный (дата обращения: 07.02.2026).

11. TDK-Lambda Corporation. PXF60 Series 60W DC/DC-Converters: Datasheet [Электронный ресурс]. – URL: https://www.tdk.com, свободный (дата обращения: 08.02.2026).

12. Пат. 2012/0313818 США, МПК H 01 Q 21/00. Active Electronically Scanned Array (AESA) Card / A.M. Puzella (США), D.A. Bozza (США), P.S. Dupuis (США). № 13/295 437; заявл. 14.11.2011; опубл. 13.12.2012. – 17 с. Патент на англ. яз.

13. Пат. 3319172 ЕПВ, МПК G 01 S 7/02. Airborne / Spaceborne Distributed Aperture Multi Sensing Pulsed RF System / A. Aprile (Италия), G. Isella (Италия), A. Montanari (Италия). № 17 200 470.7; заявл. 07.11.2017; опубл. 09.05.2018. – 22 с. Патент на англ. яз.

14. High-Efficiency Soft-Switching Technique for a Cascaded Buck-Boost-Converter / L. Liu, J. Dai et al. // Electronics. – 2025. – Vol. 14 (22). – P. 4499.

15. EMC-Friendly Gate Driver Design in GaN-Based DC-DC Converters for Automotive Electronics: A Review / X. Wu et al. // Electronics. – 2026. – Vol. 15 (2). – P. 283.

16. Wu Y. A 13 G to 24.8 GHz Broadband Power Amplifier with 23% PAE for Sensor Applications / Y. Wu, J. Cui // Engineering Proceedings. – 2025. – Vol. 118 (1). – P. 67.

17. Design of AD Converters in 0.35 μm SiGe BiCMOS Technology for Ultra-Wideband M-Sequence Radar Sensors / M. Sokol et al. // Sensors. – 2024. – Vol. 24 (9). – P. 2838.

18. Мелешин В.И. Транзисторная преобразовательная техника. – М.: Техносфера, 2005. – 632 с.

19. Гончаров А.Ю. Особенности построения матричных систем распределенного электропитания для АФАР // Компоненты и технологии. – 2016. – № 12. – С. 100–104.

20. Модули питания низкопрофильные серии МНМ: Технические условия [Электронный ресурс]: сайт ООО «АЕДОН». – URL: https://aedon.ru, свободный (дата обращения: 10.02.2026).

21. Analog Devices. LTC7149: 60V: Datasheet. – 2022 [Электронный ресурс]. – URL: https://www.analog.com, свободный (дата обращения: 9.02.2026).

22. Texas Instruments. TPS6282x: Step-Down Converter with 1% Device Accuracy: Datasheet. – 2021 [Электронный ресурс]. – URL: https://www.ti.com, свободный (дата обращения: 25.01.2026).

23. Texas Instruments. TPS62130A-Q1: Datasheet. – 2021. [Электронный ресурс]. – URL: https://www.ti.com, свободный (дата обращения: 25.01.2026).

Зотов Федор Андреевич
Магистрант каф. передовых технологий (КПТ)
МИРЭА – Российского технологического ун-та
Вернадского пр-т, 78, г. Москва, Россия, 119454
ORCID: 0009-0009-6476-963X
Тел.: +7-915-236-37-97
Эл. почта: zotov.f.a@mirea.ru

Меркулов Андрей Александрович
Канд. техн. наук, доцент, зав. каф. КПТ
МИРЭА – Российского технологического ун-та
Вернадского пр-т, 78, г. Москва, Россия, 119454
Тел.: +7 (499) 600-80-80, доб. 20762
Эл. почта: merkulov@mirea.ru

Цитович Алексей Александрович
Директор центра проектирования и тестирования
СВЧ-устройств МИРЭА –
Российского технологического ун-та
Вернадского пр-т, 78, г. Москва, Россия, 119454
Тел.: +7 (499) 600-80-80, доб. 20774
Эл. почта: tsitovich@mirea.ru

Поступила в редакцию: 01.05.2026.
Принята к публикации: 13.08.2026.

Zotov F.A., Merkulov A.A., Tsitovich A.A.

Design of a Secondary Power Supply for a Ka-Band AESA

Relevance. The development of compact and energy-efficient power supplies with a low level of electromagnetic interference is a critical task in the design of modern Ka-band active phased array radar (AESA) transceiver modules. ***Purpose of the study.*** This study aims to develop and substantiate circuit-design solutions for the construction of a secondary power supply intended for Ka-band AESA. ***Methods.*** The research employed methods of circuit design, multi-criteria selection of electronic components based on weight coefficients, and computer SPICE simulation of electrical circuit operation modes taking into account the parasitic parameters of components. ***Scientific novelty.*** A two-level distributed power supply architecture is proposed and justified, featuring physical separation of intermediate bus channels and the integration of high-frequency synchronous converters in an inverting configuration. This topology effectively minimizes the impact of common-impedance coupling interference on sensitive microwave circuits. ***Results.*** A microcontroller-based algorithmic control system for power-up sequencing has been developed to prevent the electrical breakdown of microwave transistors. Simulation results confirm a power-stage efficiency of 86–91% at output currents up to 15 A, while the ripple amplitude on low-voltage channels does not exceed 15 mV. ***Practical significance.*** The proposed circuit solutions can be implemented in the design of onboard equipment and small-sized radar systems to ensure high reliability and electromagnetic compatibility within densely integrated antenna array modules.

Keywords: secondary power supply, AESA, DC/DC-converter, electromagnetic compatibility, circuit simulation, power sequencing, microcontroller, energy efficiency.
DOI: 10.21293/1818-0442-2026-29-2-39-47

References

1. Belous A.I., Merdanov M.K., Shvedov S.V. *SVCh-elektronika v sistemakh radiolokatsii i svyazi. Tekhnicheskaya entsiklopediya. Kn. 1* [Microwave electronics in radar and communication systems. Technical encyclopedia. Book 1]. M., Tekhnosfera Publ., 2021, 724 pp. (in Russ.).

2. Bataev V.Ya., Demshevskiy V.V., Tsitovich A.A. [Design of printed dipole array for X-band AESA]. *Elektronnaya tekhnika*. *Seriya 1: SVCh-tekhnika*, 2019, no. 3 (542), pp. 76–87 (in Russ.).

3. Kuznetsov V.V., Tsitovich A.A., Kuchmiy A.D., Izmaylov A.A. *Kompleksirovannaya antennaya sistema X- i Ku-diapazonov* [Integrated X- and Ku-band antenna system]. *Elektronika i mikroelektronika SVCH*, 2024, vol. 1, pp. 309–313 (in Russ.).

4. Glushko E.V., Yurchenko G.S., Petrishche D.A. *Vysokoplotnyi impul'snyi istochnik vtorichnogo elektropitaniya aktivnykh fazirovannykh antennykh reshetok* [High-density pulse secondary power supply for active phased antenna arrays]. *Aktual'nye problemy fiziki, elektroniki i energetiki* [Actual Problems of Physics, Electronics and Energy: Proc. of the 1st Int. Sci. and Pract. Conf.]. Novopolotsk, PSU Publ., 2023, pp. 97–104 (in Russ.).

5. Nenartovich N.E., Mityashev M.B. [From the practice of developing active phased antenna arrays]. *Russian Technological Journal*, 2014, no. 3 (4), pp. 173–188 (in Russ.).

6. Paneeva A.P., Komissarova E.V. [A transceiver module for an active phased array antenna of a small spacecraft]. *Journal of Radio Electronics*, 2024, no. 2, pp. 1–12 (in Russ.).

7. Karasev M.S. [Online monitoring of electrical parameters of X-frequency band transceiving modules]. *Elektronnaya tekhnika*. *Seriya 1: SVCh-tekhnika*, 2021, no. 3 (550), pp. 6–14 (in Russ.).

8. Vicor Corporation. VIT028x6U1000y000 6U VITA 62 Military DC-DC Converter: Datasheet. Available at: https://www.vicorpower.com (accessed: 6 February 2026).

9. SynQor Inc. VPX-3U-DC28P-004 Vita 62 3U VPX Power Supply: Datasheet. Available at: https://www.synqor.com (accessed: 10 February 2026).

10. Behlman Electronics Inc. VPXtra 500M 3U VPX DC to DC Power Supply: Datasheet. Available at: https://www.behlman.com (accessed: 7 February 2026).

11. TDK-Lambda Corporation. PXF60 Series 60W DC/DC-Converters: Datasheet. Available at: https://www.tdk.com (accessed: 8 February 2026).

12. Puzella A.M., Bozza D.A., Dupuis P.S. Active Electronically Scanned Array (AESA) Card. Patent US, no. 2012/0313818 A1, 2012.

13. Aprile A., Isella G., Montanari A. Airborne / Spaceborne Distributed Aperture Multi Sensing Pulsed RF System. Patent EP, no. 3319172 A1, 2018.

14. Liu L., Dai J. et al. High-Efficiency Soft-Switching Technique for a Cascaded Buck-Boost-Converter. *Electronics*, 2025, vol. 14, no. 22, pp. 4499.

15. X. Wu et al. EMC-Friendly Gate Driver Design in GaN-Based DC-DC Converters for Automotive Electronics: A Review. *Electronics*, 2026, vol. 15, no. 2, pp. 283.

16. Wu Y., Cui J. A 13 G to 24.8 GHz Broadband Power Amplifier with 23% PAE for Sensor Applications. *Engineering Proceedings*, 2025, vol. 118, no. 1, pp. 67.

17. Sokol M., et al. Design of AD Converters in 0.35 µm SiGe BiCMOS Technology for Ultra-Wideband M-Sequence Radar Sensors. *Sensors*, 2024, vol. 24, no. 9, pp. 2838.

18. Meleshin V.I. *Tranzistornaya preobrazovatel'naya tekhnika* [Transistor converter technology]. M., Tekhnosfera Publ., 2005. 632 pp. (in Russ.).

19. Goncharov A.Yu. *Osobennosti postroeniya matrichnykh sistem raspredelennogo elektropitaniya dlya AFAR* [Features of building matrix distributed power supply systems for active phased antenna arrays]. *Komponenty i tekhnologii*, 2016, no. 12, pp. 100–104 (in Russ.).

20. OOO AEDON. *Moduli pitaniya nizkoprofil'nye serii MNM: Tekhnicheskie usloviya* [Low-profile power modules of MNM series: Technical specifications]. Available at: https://aedon.ru (accessed: 10 February 2026) (in Russ.).

21. Analog Devices. LTC7149: 60V, 4A Synchronous Step-Down Regulator for Inverting Output Applications: Datasheet. 2022. Available at: https://www.analog.com (accessed: 9 February 2026).

22. Texas Instruments. TPS6282x: Step-Down Converter with 1% Device Accuracy: Datasheet. 2021. Available at: https://www.ti.com (accessed: 25 January 2026).

23. Texas Instruments. TPS62130A-Q1: Datasheet. 2021. Available at: https://www.ti.com (accessed: 25 January 2026).

Fedor A. Zotov
Master's Student, Dept. of Advanced Technologies
MIREA – Russian Technological University
78, Vernadsky Ave., Moscow, 119454, Russia
ORCID: 0009-0009-6476-963X
Phone: +7-915-236-37-97
Email: zotov.f.a@mirea.ru

Andrey A. Merkulov
Candidate of Engineering Sciences, Associate Professor, Head of the Dept. of Advanced Technologies
MIREA – Russian Technological University
78, Vernadsky Ave., Moscow, 119454, Russia
Phone: +7 (499) 600-80-80, ext. 20762
Email: merkulov@mirea.ru

Aleksey A. Tsitovich
Director of the Microwave Device Design and Testing Center
MIREA – Russian Technological University
78, Vernadsky Ave., Moscow, 119454, Russia
Phone: +7 (499) 600-80-80 ext. 20774
Email: tsitovich@mirea.ru

Received: 01.05.2026.
Accepted: 13.08.2026.

УДК 621.396.67
Шифр специальности ВАК: 2.2.14

С.Б. Чинь, Т.Р. Газизов

Влияние аппроксимации оптимальной токовой сеткой на массу и моностатическую эффективную площадь рассеяния параболической рефлекторной проводно-сеточной антенны

Актуальность. Актуальной задачей в области телекоммуникационных и радиолокационных систем является снижение массы и моностатической эффективной площади рассеяния (МЭПР) антенн при сохранении излучательных характеристик на практически приемлемом уровне. Для снижения массы эффективно использование антенн из проводной сетки (ПС) с её разреживанием посредством аппроксимации оптимальной токовой сеткой (АОТС) на основе метода моментов (МоМ), когда удаляются провода с модулем тока ниже заданного допуска удаления элементов сетки (ДУЭС). ***Цель исследования.*** Оценка влияния АОТС на МЭПР разреженной антенны. ***Методы.*** Исследование выполнено на примере параболической рефлекторной антенны из ПС, возбуждаемой диполем и согласованной в полосе 3,25–3,85 ГГц. Моделирование выполнялось в среде MATLAB. Характеристики излучения оценивались в режиме возбуждения антенны диполем, а МЭПР для линейной и круговой поляризаций падающей волны вычислялась при замене возбуждающего сегмента сопротивлением 50 Ом, учитываемым в матрице импедансов МоМ. ***Научная новизна.*** Впервые исследована МЭПР проводно-сеточной параболической антенны, разреженной с помощью АОТС. Показано, что АОТС позволяет снизить не только массу антенны, но и её МЭПР в главном направлении, причём без дополнительных вычислительных затрат на учёт критерия уменьшения МЭПР при сохранении излучательных характеристик на практически приемлемом уровне. ***Результаты.*** При ДУЭС = 30, 35, 40% достигнуто уменьшение массы в 1,57; 1,65 и 1,73 раза при приемлемом сохранении согласования и коэффициента усиления (снижение не более 0,11 дБ) и стабильной ширины главного лепестка. Кроме того, в расширенном диапазоне 3–5 ГГц средняя МЭПР в главном направлении снизилась на 4,4–4,8 дБ (максимально до 13,9 дБ на частоте 3,8 ГГц) при линейной поляризации и на 1,7–2,53 дБ (максимально до 4,6 дБ на частоте 4 ГГц) при круговой. При этом в отдельных приосевых направлениях ($\theta \approx \pm 12°$) МЭПР возрастает, поскольку критерий по МЭПР не входит в процесс разреживания, что указывает направление дальнейшего совершенствования подхода. ***Практическая значимость.*** Полученные результаты показывают, что разреживание снижает массу антенны и сопровождается снижением МЭПР в главном направлении, где МЭПР исходной структуры максимальна, и открывает путь к интеграции критерия уменьшения МЭПР в процесс разреживания.
Ключевые слова: рефлекторная антенна, проводная сетка, аппроксимация оптимальной токовой сеткой, метод моментов, моностатическая эффективная площадь рассеяния, разреженная антенна.
DOI: 10.21293/1818-0442-2026-29-2-48-56

Введение

В современных телекоммуникационных и радиолокационных системах антенна является одним из важных элементов. Различные типы антенн отличаются характеристиками излучения и массой, что определяет области их применения. Среди них рефлекторная антенна широко используется как в военной [1, 2], так и гражданской сферах, например, в системах спутниковой связи [3, 4], благодаря её способности эффективно фокусировать и концентрировать электромагнитные волны [5]. Рефлекторы антенны обычно изготавливаются из сплошных металлических плоскостей, что приводит к относительно большой массе, требующей прочных опорных конструкций [6, 7]. Данная проблема особенно важна в космических применениях с жёсткими требованиями к массе [8, 9], поэтому снижение массы антенны в целом, а рефлекторных антенн в частности, при сохранении их излучательных характеристик является актуальной задачей.

Кроме того, в последних исследованиях антенн значительное внимание уделяется снижению моностатической эффективной площади рассеяния (МЭПР), поскольку она напрямую влияет на скрытность всей системы [10, 11]. В ряде работ предложены методы снижения МЭПР на основе метаповерхностей [12–14], частотно-селективных поверхностей (FSS) [15, 16], радиопоглощающих материалов [17] и анализа характеристических мод [18, 19], которые позволили достичь заметных результатов. Однако перечисленные методы, как правило, требуют интеграции дополнительных структур или этапов проектирования, что усложняет создание антенн. В ряде случаев дополнительные структуры ухудшают характеристики излучения антенны [20]. Таким образом, в современном моделировании и проектировании антенн актуально одновременное снижение массы и МЭПР при сохранении излучательных характеристик антенны.

Для снижения массы можно заменить сплошную поверхность проводной сеткой (ПС), а затем применить аппроксимацию оптимальной токовой сеткой (АОТС) для разреживания на основе метода моментов (МоМ). Исследования АОТС показали эффективность снижения массы антенн [21–24], рассеивателей [25] и даже электромагнитных экранов [26] из ПС удалением отдельных проводов ПС. При этом результаты моделирования были многократно подтверждены результатами экспериментов. Для разреженных антенн из ПС удаляются провода с малым током,

а значит, и вкладом в излучение. Благодаря этому получаются разреженные, более лёгкие структуры при сохранении характеристик антенны на уровне, близком к исходному. Однако разреженные антенны после АОТС рассматривались лишь с точки зрения снижения массы и сохранения излучательных характеристик на практически приемлемом уровне, тогда как их МЭПР не рассматривалась.

Цель работы – исследовать влияние АОТС на массу и МЭПР параболической рефлекторной проводно-сеточной антенны.

Применение АОТС к параболической антенне с проводной сеткой

Суть АОТС заключается в удалении проводов с малым током, поскольку мал их вклад в характеристики антенны. Другими словами, после разреживания сохраняются только провода, вносящие значительный вклад в излучение, что подтверждено в работах [21–24]. АОТС следует подходу, изложенному в [23]: антенна из ПС анализируется МоМ [27] со ступенчатыми базисными функциями.

Каждый провод представляется одним сегментом, а возбуждающий элемент (где требуется более детальное описание распределения тока) разбивается на несколько сегментов. Вектор возбуждения **v** имеет только один ненулевой элемент, соответствующий напряжению 1 В, приложенному к возбуждающему сегменту. Решается система уравнений МоМ $\mathbf{Zi} = \mathbf{v}$, где **Z** – матрица импедансов, а **i** – вектор распределения тока, элементы которого соответствуют токам на отдельных проводах. После получения **i** на частоте f модули токов нормируются по их максимальному или среднему значению, а затем сравниваются с заданным допуском удаления элементов сетки (ДУЭС). Если провод имеет нормированный модуль тока меньше ДУЭС, он считается проводом с малым током и удаляется.

Таким образом, изменяя ДУЭС, можно получить различные разреженные антенны: чем больше ДУЭС, тем более разреженной и лёгкой становится структура, однако характеристики антенны могут ухудшаться, поэтому ДУЭС следует выбирать так, чтобы обеспечить баланс между снижением массы и сохранением характеристик антенны на приемлемом уровне.

В данной работе АОТС применяется к параболической рефлекторной антенне [28, 29], согласованной в полосе 3,25–3,85 ГГц с резонансной частотой f = 3,5 ГГц. Рефлектор имеет диаметр D = 680 мм, глубину h = 245 мм и фокусное расстояние F = 118 мм (рис. 1). Поверхность рефлектора разбита на $N_г$ = 76 радиальных и $N_с$ = 25 концентрических рядов проводов для формирования поверхности ПС (3800 проводов). Сетка с диаметром проводов d = 0,02λ (λ – длина волны на 3,5 ГГц) (рис. 2, *а*). Антенна возбуждается диполем с диаметром d = 0,02λ и длиной l = 0,35λ, расположенным вдоль оси x в фокусе параболического рефлектора и разделённым на 3 сегмента длиной около 0,12λ, что удовлетворяет условиям тонкопроводной аппроксимации. Возбуждающее напряжение 1 В приложено к среднему сегменту диполя, выполняющему роль возбуждающего сегмента в векторе возбуждения **v** системы уравнений МоМ.

В среде MATLAB АОТС применяется на частоте f = 3,5 ГГц, при этом токи в проводах нормируются по среднему значению. В работе рассматриваются исходная и разреженные антенны при ДУЭС = 30, 35, 40% (см. рис. 2).

Как исходная, так и разреженные антенны хорошо согласованы по импедансу на 3,5 ГГц (рис. 3, *а*) с очень малым отличием (не более 0,19 дБ относительно исходной при ДУЭС = 40%). Коэффициент отражения $|s_{11}|$ = –14,96; –14,8; –14,82; –14,77 дБ для исходной антенны и при ДУЭС = 30, 35, 40% соответственно. Более того, зависимость $|s_{11}|$ разреженных антенн близко повторяет исходную с полосой 3,25–3,85 ГГц (около 17%). Таким образом, после АОТС разреженные антенны сохраняют хорошее согласование импеданса, близкое к исходной структуре, во всём рабочем диапазоне.

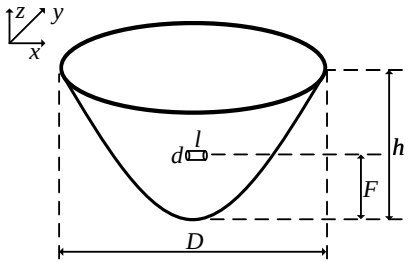
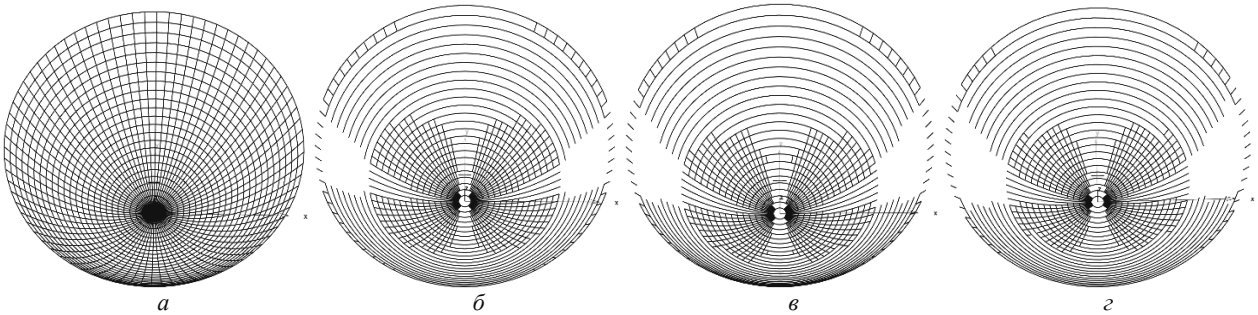


Рис. 1. Изометрический вид рефлекторной антенны



а *б* *в* *г*

Рис. 2. Рефлекторные антенны из ПС: исходная (*а*), ДУЭС = 30 (*б*), 35 (*в*), 40% (*г*)

Излучательные характеристики также сохраняются на хорошем уровне. Так, максимальные коэффициенты усиления $КУ_{макс}$ разреженных и исходной антенн очень близки во всем рабочем диапазоне (рис. 3, *б*). На резонансной частоте $КУ_{макс}$ = 22,89; 22,86; 22,81 дБи при ДУЭС = 30, 35, 40% соответственно, а у исходной $КУ_{макс}$ = 22,92 дБи. $КУ_{макс}$ немного снижается с ростом ДУЭС, что объяснимо тем, что разреживание изменяет эффективную отражающую площадь параболического рефлектора. Такой эффект также наблюдался в исследованиях [22–25], но снижение в данном случае (не более 0,11 дБ при ДУЭС = 40%) может считаться приемлемым.

Анализ диаграмм направленности (ДН) на резонансной частоте в плоскостях φ = 0° и 90° (рис. 3, *в*, *г*) показывает, что разреженные антенны очень хорошо сохраняют излучательные характеристики исходной. Максимальный КУ в главном направлении (θ = 0°) изменяется очень незначительно, как отмечено выше. Ширина ДН $\theta_{0,5р}$ также стабильна в обеих плоскостях (около 13,1° при φ = 0° и 7,7° при φ = 90°). Это подтверждает, что разреживание данной антенны для снижения массы мало влияет на главный лепесток (ГЛ).

Примечательно, что в боковых лепестках (БЛ) разреженные антенны проявляют преимущество в ряде направлений. Так, при φ = 0° вблизи направления θ ≈ 90°/270° уровень БЛ снижается более чем на 10 дБ при ДУЭС = 40%. Это также наблюдается в направлении θ ≈ 120°/240° при ДУЭС = 30, 35%. В плоскости φ = 90° наибольший уровень БЛ (около θ ≈ 46°/314°) практически не изменяется и даже немного снижается при ДУЭС = 30, 35% (с 3,77 дБи до 3,72 и 3,68 дБи). Таким образом, в ряде направлений вне ГЛ разреживание способствует снижению БЛ. Эти полезные эффекты могут быть связаны с фазой токов на проводах. Поскольку АОТС выбирает удаляемые провода только по модулю тока, не учитывая фазу, влияние разреживания на разные направления неоднородно. В направлении ГЛ, где вклады проводов практически синфазны и определяются в основном проводами с большим током, удаление проводов с малым током почти не изменяет характеристики излучения. Уровни БЛ же формируются интерференцией множества вкладов с различными фазами, поэтому удаление части проводов может снижать БЛ в одном направлении, незначительно изменяя его в другом.

Вместе с тем есть некоторое ухудшение, хотя и незначительное, и сосредоточенное главным образом в области обратного излучения плоскости φ = 0°. В частности, отношение излучения вперёд-назад в этой плоскости снижается с 35,3 дБ (исходная) до примерно 33,7 дБ (при ДУЭС = 35, 40%).

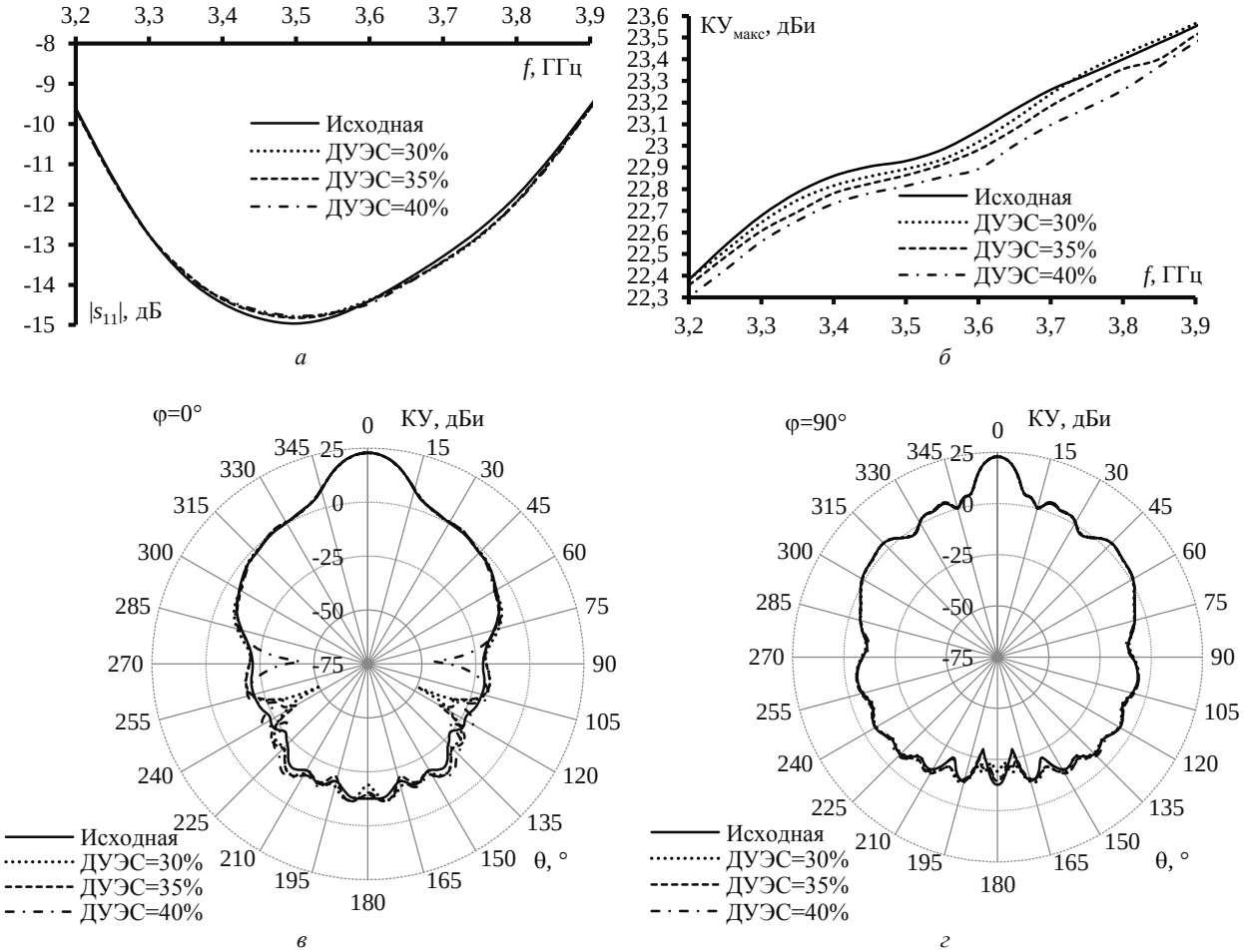


Рис. 3. Частотные зависимости $|s_{11}|$ (*а*), $КУ_{макс}$ (*б*) и ДН в плоскостях φ = 0° (*в*), 90° (*г*) для исходной и разреженных антенн

При этом максимальный уровень обратного излучения возрастает с −12,3 дБи до примерно −10,8 дБи (ухудшается примерно на 1,5 дБ). В плоскости φ = 90° отношение излучения вперёд-назад практически сохраняется (≈ 29 дБ). Здесь отношением излучения вперёд-назад называется отношение уровня излучения в направлении главного максимума (θ = 0°) к уровню излучения в обратном направлении (θ = 180°), выраженное в дБ. Причиной данного ухудшения считается то, что после удаления части проводов поверхность становится более разреженной и небольшая доля энергии проникает назад. Степень ухудшения также имеет тенденцию слегка возрастать с ростом ДУЭС, но общая форма ДН не изменяется.

Масса антенны заметно снизилась. Для оценки этого вводится параметр, называемый отношением масс (M) проводов исходной и разреженной антенн. После АОТС отношение масс равно 1,57; 1,65; 1,73 при ДУЭС = 30, 35, 40% соответственно. Таким образом, масса разреженных антенн значительно уменьшилась при сохранении важных характеристик антенны.

В итоге разреживание антенны приемлемо сохраняет основные излучательные характеристики и даже снижает БЛ в ряде направлений. Одновременно достигается довольно большое отношение масс при незначительном ухудшении $КУ_{макс}$. В частности, на резонансной частоте $КУ_{макс}$ только снижается на 0,03; 0,06; 0,11 дБ относительно исходной при ДУЭС = 30, 35, 40% соответственно, и примерно на 1,5 дБ ухудшается отношение излучения вперёд-назад в плоскости φ = 0°. Для оценки МЭПР после АОТС для каждой антенны МЭПР определялась для двух поляризаций падающей волны: линейной и круговой. Падающая плоская волна задаётся в терминах единичных векторов $\left(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}\right)$ сферической системы координат и фазора $e^{+j\mathbf{kr}}$, где $\mathbf{k}$ – волновой вектор, направленный вдоль направления распространения волны, а $\mathbf{r}$ – радиус-вектор точки наблюдения. При линейной поляризации θ-поляризованная падающая волна имеет вектор электрического поля, направленный вдоль $\hat{\boldsymbol{\theta}}$, и записывается как

$$\mathbf{E}_{\text{inc}} = \hat{\boldsymbol{\theta}}\, e^{+j\mathbf{kr}}, \qquad (1)$$

а при круговой

$$\mathbf{E}_{\text{inc}} = \frac{1}{\sqrt{2}}(\hat{\boldsymbol{\theta}} - j\hat{\boldsymbol{\varphi}})\, e^{+j\mathbf{kr}}, \qquad (2)$$

что соответствует волне с правой круговой поляризацией (ПКП) – волне, у которой вектор электрического поля вращается по часовой стрелке при наблюдении в направлении распространения волны.

Следует отметить, что при исследовании МЭПР антенна не возбуждается в месте расположения источника: вместо этого сегмент заменяется сопротивлением 50 Ом в соответствии с волновым сопротивлением тракта. В МоМ это реализуется добавлением 50 Ом к диагональному элементу матрицы импедансов Z, соответствующему сегменту источника. В итоге получаемая МЭПР включает как структурную составляющую рассеяния от ПС, так и переизлучение антенны через оконечную нагрузку.

При падающей волне θ-поляризации в диапазоне 3–5 ГГц в направлении (0,0) наблюдается снижение МЭПР (рис. 4, *а*). Средняя МЭПР снижается с 12,23 дБм^2 для исходной структуры до 7,78; 7,68 и 7,47 дБм^2 при ДУЭС = 30, 35, 40% (примерно на 4,4–4,8 дБ) (таблица). На некоторых частотах снижение очень велико (до 13,9 дБ на 3,8 ГГц при ДУЭС = 40%). Снижение может быть обусловлено тем, что АОТС удалила часть проводов, существенно влияющих на рассеяние. Одновременно разреживание изменяет эффективную отражающую площадь параболического рефлектора, что также способствует снижению МЭПР. Однако снижение неравномерно по частоте: в области 3,0–3,2 ГГц оно весьма мало.

Причину этого можно объяснить следующим образом. Отражающие свойства ПС зависят от отношения размера ячейки к длине волны λ. Когда ячейка мала по сравнению с λ, сетка отражает лучше, а когда частота возрастает и размер ячейки сравним с λ, отражение более чувствительно к плотности проводов. Ввиду этого в области низких частот ячейки даже после разреживания в большинстве своём остаются достаточно малыми, так что удаление части проводов мало изменяет отражённое поле и МЭПР снижается незначительно. Напротив, при возрастании частоты тот же уровень разреживания вызывает более заметные изменения и МЭПР ниже.

Исследование по углу на частоте 3,5 ГГц в плоскости φ = 0°, содержащей диполь (см. рис. 4, *б*), показывает наиболее впечатляющий результат именно в направлении (0,0): МЭПР снижается с 12,93 до 4,70; 4,59; 4,05 дБм^2. Однако в направлениях, отклонённых от нормали, особенно вблизи θ ≈ ±12…±15°, МЭПР разреженных антенн выше, чем исходной (при θ = ±12° возрастает с 7,4 до 12,1 дБм^2). Тем самым энергия рассеяния от зеркального максимума перераспределяется в другие направления.

Следует отметить, что АОТС выполняет разреживание только на основе модуля тока, не учитывая характеристику рассеяния в критерии удаления, поэтому возрастание МЭПР в некоторых приосевых направлениях является неконтролируемым следствием процесса разреживания. Это может быть связано с нарушением фазовой компенсации токов проводов, удерживавшей рассеяние в этих направлениях низким. На частотах 3,8 и 4,8 ГГц (см. рис. 4, *в*, *г*) получается аналогично.

В главном направлении МЭПР снижается с 11,91 дБм^2 до −0,21; −0,92 и −2,01 дБм^2 на 3,8 ГГц и с 13,47 дБм^2 до 7,05; 7,13 и 7,03 дБм^2 на 4,8 ГГц при ДУЭС = 30, 35, 40% соответственно. При этом максимум МЭПР по углу не возрастает, а снижается: с 12,93 до 12,11–12,18 дБм^2 на 3,5 ГГц, с 11,91 до 10,45–10,60 дБм^2 на 3,8 ГГц и с 13,47 до 11,98–12,18 дБм^2 на 4,8 ГГц.

Таким образом, обнаруженный эффект наблюдается в рассмотренном диапазоне, включая частоты за пределами рабочей полосы антенны.

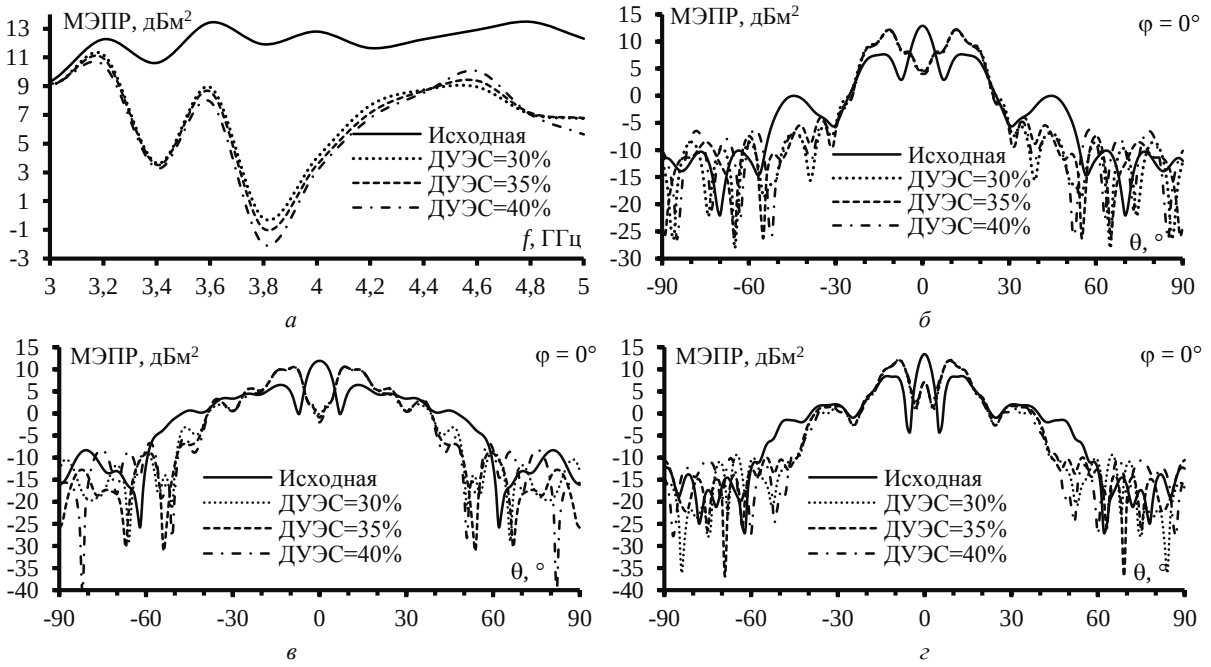


Рис. 4. МЭПР при падающей волне θ-поляризации для исходной и разреженных антенн в диапазоне 3–5 ГГц (*а*) и в плоскости φ = 0° на 3,5 (*б*), 3,8 (*в*) и 4,8 ГГц (*г*)

$|S_{11}|$, $КУ_{макс}$, средняя МЭПР в главном направлении и отношение масс для исходной и разреженных антенн

Структура	$\|S_{11}\|$ на 3,5 ГГц, дБ	$КУ_{макс}$ на 3,5 ГГц, дБи	Средняя МЭПР, 3–5 ГГц, θ-поляризация, дБм²	Средняя МЭПР, 3–5 ГГц, ПКП, дБм²	*M*
Исходная ДУЭС, %	–14,96	22,92	12,23	10,23	–
30	–14,8	22,89	7,78	8,53	1,57
35	–14,82	22,86	7,68	8,05	1,65
40	–14,77	22,81	7,47	7,7	1,73

При падающей волне ПКП разреженные структуры также дают снижение МЭПР в главном направлении в диапазоне 3–5 ГГц, за исключением области 3,0–3,2 ГГц, где снижение практически отсутствует (рис. 5, *а*). По сравнению с исходной (10,23 дБм²) средняя МЭПР при ДУЭС = 30, 35, 40% снижается до 8,53; 8,05 и 7,7 дБм², что меньше на 1,7; 2,18 и 2,53 дБ. Максимальное снижение до 4,6 дБ на частоте 4 ГГц при ДУЭС = 35%. Исследование по углу в плоскости φ = 0° (см. рис. 5, *б*) показывает, что наиболее выраженный эффект по-прежнему в главном направлении (0,0): МЭПР снижается с 10,21 дБм² до 6,57; 6,2 и 5,92 дБм², на 3,64–4,29 дБ. Аналогично случаю линейной поляризации часть энергии рассеяния смещается в приосевые направления: вблизи θ ≈ ±12° МЭПР разреженных антенн возрастает с 5,48 дБм² до примерно 9,6–9,9 дБм², тогда как в области больших углов (θ = 30…40°) разреженные структуры имеют заметно меньшую МЭПР, чем исходная. Таким образом, АОТС снижает МЭПР в наиболее важном направлении и для ПКП, что подтверждает результаты, полученные для линейной поляризации. Однако снижение в главном направлении для ПКП меньше: примерно 4 дБ против 8–9 дБ при линейной поляризации. Это меньшее снижение можно объяснить тем, что разреживание выполняется по току при возбуждении диполем: удаляются провода, ориентированные почти ортогонально его поляризации. Волна θ-поляризации совпадает с выделенной поляризацией, а ПКП содержит две ортогональные составляющие, из которых лишь одна ей соответствует; кроме того, критерий уменьшения МЭПР не учитывается в разреживании.

На частотах 3,8 и 4,8 ГГц (см. рис. 5, *в*, *г*) в главном направлении МЭПР снижается с 9,67 дБм² до 6,59; 6,01 и 5,30 дБм² на 3,8 ГГц и с 11,28 дБм² до 8,28; 7,99 и 7,45 дБм² на 4,8 ГГц при ДУЭС = 30, 35, 40% соответственно. Максимум МЭПР по углу также снижается: с 10,21 до 9,57–9,86 дБм² на 3,5 ГГц, с 9,67 до 8,40–8,44 дБм² на 3,8 ГГц и с 11,28 до 9,47–9,89 дБм² на 4,8 ГГц. Следовательно, эффект наблюдается при обеих поляризациях падающей волны в рассмотренном диапазоне.

Следует подчеркнуть, что при разреживании ток каждого провода определяется в режиме возбуждения антенны от источника и провода с малым током удаляются исходя только из этого распределения тока. При этом рассеяние в критерии удаления проводов не учитывается. Тем не менее полученные разреженные антенны одновременно со снижением массы (отношение масс до 1,73) обеспечивают снижение МЭПР в главном направлении (в среднем на 4,4–4,8 дБ при θ-поляризации и на 1,7–2,53 дБ при ПКП). Таким образом, снижение МЭПР в главном направлении представляет собой сопутствующий эффект разреживания по току, что показано впервые.

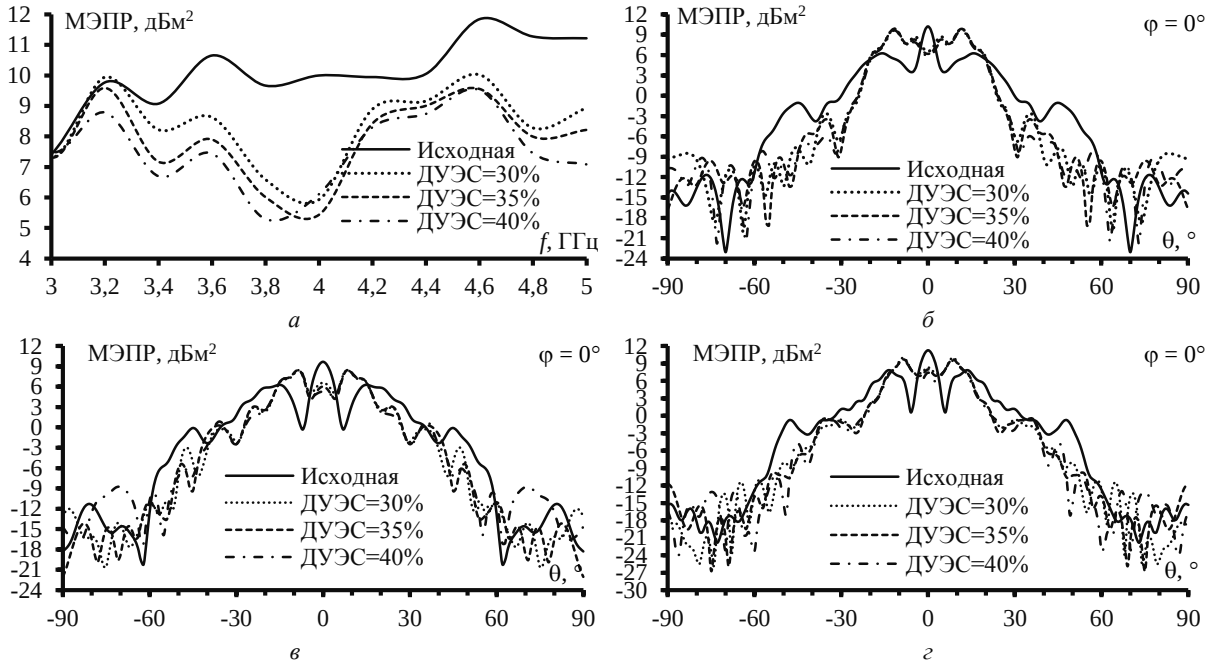


Рис. 5. МЭПР при падающей волне ПКП для исходной и разреженных антенн в диапазоне 3–5 ГГц (*а*) и в плоскости φ = 0° на 3,5 (*б*), 3,8 (*в*) и 4,8 ГГц (*г*)

Вместе с тем этот положительный эффект сопровождается возрастанием МЭПР в отдельных приосевых направлениях: поскольку удаление проводов не контролируется по критерию рассеяния, то и перераспределение рассеянной энергии остаётся неконтролируемым. Отсюда вытекают направления дальнейших исследований: учёт критерия уменьшения МЭПР при разреживании и фазы тока наряду с его амплитудой, что позволит управлять перераспределением рассеянной энергии и снижать массу и МЭПР согласованно без ухудшения излучательных характеристик и согласования импеданса. Кроме того, как отмечено в [24], удаление проводов может приводить к появлению свободных проводов, не связанных с сеткой, что усложняет изготовление, поэтому подход АОТС целесообразно дополнить механизмом восстановления связности оставшихся проводов.

Заключение

АОТС была применена к параболической рефлекторной антенне из ПС в диапазоне 3,25–3,85 ГГц с целью снижения массы, после чего впервые исследовано влияние разреживания на МЭПР. При ДУЭС = 30, 35, 40% разреженные структуры достигают отношения масс M = 1,57; 1,65; 1,73, сохраняя при этом хорошие излучательные характеристики: $|s_{11}|$ и $КУ_{макс}$ мало изменяются (снижаются не более 0,11 дБ), ширина $\theta_{0,5р}$ стабильна, а БЛ в ряде направлений даже уменьшаются. Одновременно разреживание снижает МЭПР: при падающей волне линейной поляризации МЭПР в главном направлении снижается в среднем на 4,4–4,8 дБ в расширенном диапазоне частот 3–5 ГГц, который шире рабочего диапазона антенны (3,25–3,85 ГГц), и до 8–9 дБ на 3,5 ГГц, а её максимальное снижение достигает 13,9 дБ на частоте 3,8 ГГц (при ДУЭС = 40%). При ПКП снижение меньше, но заметно, и максимальное снижение достигает 4,6 дБ на частоте 4 ГГц (при ДУЭС = 35%). В то же время МЭПР возрастает в некоторых приосевых направлениях. Важно подчеркнуть, что снижение МЭПР получено без дополнительных вычислительных затрат на оптимизацию по критерию уменьшения рассеяния. Таким образом, впервые исследована МЭПР проводно-сеточной антенны, разреженной с помощью АОТС. Результаты показали, что разреживание АОТС снижает массу антенны и сопровождается снижением МЭПР в главном направлении, где МЭПР исходной структуры максимальна, и открывает путь к интеграции критерия уменьшения МЭПР в процесс разреживания.

Работа выполнена при финансовой поддержке Минобрнауки России по проекту FEWM-2024-0005.

Литература

1. Development of highly constrained 1 m Ka-band mesh deployable offset reflector antenna for next generation cubesat radars / Y. Rahmat-Samii, V. Manohar, J.M. Kovitz, R.E. Hodges, G. Freebury, E. Peral // IEEE Transactions on Antennas and Propagation. – 2019. – Vol. 67, No. 10. – P. 6254–6266. DOI: 10.1109/TAP.2019.2920223.

2. Marty L. From passive to active radar reflectors and beyond / L. Marty, P. Gerbaud, F. Christophe // 2021 IEEE Conference on Antenna Measurements & Applications (CAMA). – France: IEEE, 2021. – P. 605–607. DOI: 10.1109/CAMA49227.2021.9703473.

3. Shaped prime-focus reflector antenna for satellite communication / R.C. Gupta, S.K. Sagi, K.P. Raja, N.K. Sharma, R. Jyoti // IEEE Antennas and Wireless Propagation Letters. – 2017. – Vol. 16. – P. 1945–1948. DOI: 10.1109/LAWP.2017.2689800.

4. Rotshild D. Ultra-wideband reconfigurable X-band and Ku-band metasurface beam-steerable reflector for satellite communications / D. Rotshild, A. Abramovich // Electronics. –

2021. – Vol. 10. – Art. 2165. DOI: 10.3390/electronics 10172165.

5. Rahmat-Samii Y. Reflector antenna developments: a perspective on the past, present and future / Y. Rahmat-Samii, R. Haupt // IEEE Antennas and Propagation Magazine. – 2015. – Vol. 57, No. 2. – P. 85–95. DOI: 10.1109/MAP.2015.2414534.

6. 3D-printed dual-reflector antenna with self-supported dielectric subreflector / A. Rebollo, A.F. Vaquero, M. Arrebola, M.R. Pino // IEEE Access. – 2020. – Vol. 8. – P. 209091–209100. DOI: 10.1109/ACCESS.2020.3038739.

7. Chinese deep space stations: a brief review [Antenna Applications Corner] / G. He, M. Liu, X. Gao, X. Du, H. Zhou, H. Zhu // IEEE Antennas and Propagation Magazine. – 2022. – Vol. 64, No. 1. – P. 102–111. DOI: 10.1109/MAP.2021.3129690.

8. Antenna designs for CubeSats: a review / S. Abulgasem, F. Tubbal, R. Raad, P.I. Theoharis, S. Lu, S. Iranmanesh // IEEE Access. – 2021. – Vol. 9. – P. 45289–45324. DOI: 10.1109/ACCESS.2021.3066632.

9. A review of multibeam phased array antennas as LEO satellite constellation ground station / G. He, X. Gao, L. Sun, R. Zhang // IEEE Access. – 2021. – Vol. 9. – P. 147142–147154. DOI: 10.1109/ACCESS.2021.3124318.

10. 3D-printed microwave patch antenna via fused deposition method and ultrasonic wire mesh embedding technique / M. Liang, C. Shemelya, E. MacDonald, R. Wicker, H. Xin // IEEE Antennas and Wireless Propagation Letters. – 2015. – Vol. 14. – P. 1346–1349. DOI: 10.1109/LAWP.2015.2405054.

11. Design of low-RCS antenna using antenna array / Y. Shi, X.F. Zhang, Z.K. Meng, L. Li // IEEE Transactions on Antennas and Propagation. – 2019. – Vol. 67, No. 10. – P. 6484–6493. DOI: 10.1109/TAP.2019.2925180.

12. Low-cost, low-profile wide-band radar cross section reduction using dual-concentric phase-gradient modulated surface / Y. Azizi, M. Soleimani, S.H. Sedighy, L. Matekovits // Electronics. – 2021. – Vol. 10. – Art. 1552. DOI: 10.3390/electronics10131552.

13. Ultra-wideband low-RCS and gain-enhanced phase gradient metasurface-based antenna using characteristic mode analysis / T. Zhang, J. Ren, Y. Yang, X. Pang, Q. Zheng // AEU – International Journal of Electronics and Communications. – 2024. – Vol. 183. – Art. 155370. DOI: 10.1016/j.aeue.2024.155370.

14. Broadband low-profile and low-RCS circular polarization metasurface antenna based on hybrid mechanism with characteristic mode analysis / Z. Tian, L. Yang, Z. Cheng, B. Cai, Y. Cheng, H. Luo, X. Li // AEU – International Journal of Electronics and Communications. – 2025. – Vol. 202. – Art. 156043. DOI: 10.1016/j.aeue.2025.156043.

15. Tilted beam Fabry–Perot antenna with enhanced gain and broadband low backscattering / H. Umair, T.B.A. Latef, Y. Yamada, W.N.L.B.W. Mahadi, M. Othman, K. Kamardin, M.I. Hussein, A.I. Najam // Electronics. – 2021. – Vol. 10. – Art. 267. DOI: 10.3390/electronics10030267.

16. Wideband RCS reduction in microstrip antennas using integrated artificial magnetic conductors and frequency-selective surface / S.A. Harooni, M. Naser-Moghadasi, B. Ghalamkari, M. Khatir // AEU – International Journal of Electronics and Communications. – 2026. – Vol. 205. – Art. 156135. DOI: 10.1016/j.aeue.2025.156135.

17. Saxena G. Design of metasurface absorber for low RCS and high isolation MIMO antenna for radio location & navigation / G. Saxena, Y.K. Awasthi, P. Jain // AEU – International Journal of Electronics and Communications. – 2021. – Vol. 133. – Art. 153680. DOI: 10.1016/j.aeue.2021.153680.

18. Design of broadband low-RCS array antennas based on characteristic mode cancellation / J. Han, D. Jia, B. Du, G. Han, Y. Jia, Z. Zhao // Electronics. – 2023. – Vol. 12. – Art. 1536. DOI: 10.3390/electronics12071536.

19. Liu Y. Design of broad-angle low-RCS microstrip antenna for dual-polarization using characteristic mode analysis / Y. Liu, B. Du, D. Jia // Electronics. – 2025. – Vol. 14. – Art. 2121. DOI: 10.3390/electronics14112121.

20. A progression in the techniques of reducing RCS for the targets / M.U. Ullah, T.B.A. Latef, M. Othman, M.I. Hussein, H.M. Alkhoori, Y. Yamada, K. Kamardin, R. Khalid // Alexandria Engineering Journal. – 2024. – Vol. 100. – P. 153–169. DOI: 10.1016/j.aej.2024.05.001.

21. Alhaj Hasan A.F. Modelling and designing wire-grid sparse antennas using MoM-based approaches for enhanced performance and reduced cost / A.F. Alhaj Hasan, M.T. Nguyen, T.R. Gazizov // Microwave Review. – 2023. – Vol. 29, No. 2. – P. 83–94. DOI: 10.18485/mtts_mr.2023.29.2.10.

22. Alhaj Hasan A.F. Novel MoM-based approaches for generating wire-grid sparse antenna structures / A.F. Alhaj Hasan, T.M. Nguyen, T.R. Gazizov // 2023 IEEE 24th International Conference of Young Professionals in Electron Devices and Materials (EDM). – Novosibirsk: IEEE, 2023. – P. 570–576. DOI: 10.1109/EDM58354.2023.10225219.

23. Nguyen M.T. Generating sparse wire-grid antennas using maximum current-based optimal current grid approximation / M.T. Nguyen, A.F. Alhaj Hasan, T.R. Gazizov // IEEE Open Journal of Antennas and Propagation. – 2026. – Vol. 7, No. 2. – P. 652–665. DOI: 10.1109/OJAP.2025.3543559.

24. Nguyen M.T. Design of monopole antennas with various grounds for 2.4 GHz applications / M.T. Nguyen, T.R. Gazizov // AEU – International Journal of Electronics and Communications. – 2025. – Vol. 202. – Art. 156050. DOI: 10.1016/j.aeue.2025.156050.

25. Данг Т.Ф. Усовершенствование аппроксимации оптимальной токовой сеткой для разреженных рассеивающих структур / Т.Ф. Данг, Т.Р. Газизов // Радиотехника. – 2025. – Т. 89, № 9. – С. 181–192. DOI: 10.18127/j00338486-202509-20.

26. Лукас К.С. Разреженные электромагнитные экраны из проводной сетки / К.С. Лукас, Т.Ф. Данг, Т.Р. Газизов // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2025. – Т. 28, № 4. – С. 13–19. DOI: 10.21293/1818-0442-2025-28-4-13-19.

27. Harrington R.F. Matrix methods for field problems // Proceedings of the IEEE. – 1967. – Vol. 55, No. 2. – P. 136–149. DOI: 10.1109/PROC.1967.5433.

28. Parabolic antenna JRB-25 MIMO. Jirous antennas [Электронный ресурс]. – URL: https://en.jirous.com/antenna-3,5ghz/jrb-25_mimo/ (дата обращения: 15.05.2026).

29. Нгуен М.Т. Сравнительный анализ характеристик разреженных структур рефлекторной антенны X-диапазона на разных частотах / М.Т. Нгуен, А.Ф. Алхадж Хасан, Т.Р. Газизов // Антенны и распространение радиоволн: сб. докл. Всерос. науч.-техн. конф., Санкт-Петербург, 23–25 октября 2024 г. – СПб.: СПбГЭТУ «ЛЭТИ» им. В.И. Ульянова (Ленина), 2024. – С. 20–27.

Чинь Суан Ба
Аспирант каф. телевидения и управления (ТУ)
Томского государственного университета
систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0009-0000-5711-5878
Тел.: +7-920-115-04-68
Эл. почта: xuanba6420@gmail.com

Газизов Тальгат Рашитович
Д-р техн. наук, проф., зав. каф. ТУ ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0002-1192-4853
Тел.: +7-913-826-07-24
Эл. почта: talgat.r.gazizov@tusur.ru

Поступила в редакцию: 23.06.2026.
Принята к публикации: 01.09.2026.

Trinh X.B., Gazizov T.R.

Effect of optimal current grid approximation on the mass and monostatic radar cross-section of a wire grid parabolic reflector antenna

Relevance. Reducing the mass of antennas and their monostatic radar cross-section (MRCS) while preserving the radiation characteristics at a practically acceptable level is a relevant problem in telecommunication and radar systems. To reduce the mass, it is effective to use wire grid (WG) antennas with subsequent sparsification of the grid by optimal current grid approximation (OCGA) based on the method of moments (MoM), in which wires with a normalized current magnitude below a specified grid element elimination tolerance (GEET) are removed. ***Purpose of the study.*** Evaluating the effect of OCGA on the MRCS of a sparse antenna. ***Methods.*** The study is carried out on a dipole-fed WG parabolic reflector antenna matched over 3.25–3.85 GHz. The simulation is performed in MATLAB. The radiation characteristics are evaluated with the antenna excited by the dipole, while the MRCS under linearly and circularly polarized incident waves is computed by replacing the feed segment with a 50 Ohm resistance included in the MoM impedance matrix. ***Novelty.*** The MRCS of a WG parabolic antenna sparsified by OCGA is investigated for the first time. It is shown that OCGA reduces not only the antenna mass but also its MRCS in the main direction, without additional computational cost for accounting for an MRCS reduction criterion and while preserving the radiation characteristics at a practically acceptable level. ***Results.*** At GEET = 30, 35, 40%, the mass was reduced by factors of 1.57, 1.65 and 1.73 with acceptable preservation of impedance matching and gain (degradation no more than 0.11 dB) and a stable main lobe width. In addition, over 3–5 GHz band the average MRCS in the main direction decreased by 4.4–4.8 dB (maximum up to 13.9 dB at 3.8 GHz) under linear polarization and by 1.7–2.53 dB (maximum up to 4.6 dB at 4 GHz) under circular polarization. At the same time, the MRCS increased in some near-axis directions ($\theta \approx \pm 12°$), since an MRCS reduction criterion is not included in the sparsification process, which indicates a direction for further improvement of the approach. ***Practical significance.*** The results show that sparsification reduces the antenna mass and is accompanied by a reduction of the MRCS in the main direction, where the MRCS of the initial structure is maximum, and they open a path to integrating an MRCS reduction criterion into the sparsification process.

Keywords: reflector antenna, wire grid, optimal current grid approximation, method of moments, monostatic radar cross-section, sparse antenna.
Funding: This research was funded by the Ministry of Science and Higher Education of the Russian Federation project FEWM-2024-0005.
DOI: 10.21293/1818-0442-2026-29-2-48-56

References

1. Rahmat-Samii Y., Manohar V., Kovitz J.M., Hodges R.E., Freebury G., Peral E. Development of highly constrained 1 m Ka-band mesh deployable offset reflector antenna for next generation cubesat radars. IEEE Transactions on Antennas and Propagation, 2019, vol. 67, no. 10, pp. 6254–6266. DOI: 10.1109/TAP.2019.2920223.

2. Marty L., Gerbaud P., Christophe F. From passive to active radar reflectors and beyond. 2021 IEEE Conference on Antenna Measurements & Applications (CAMA), France, 2021, pp. 605–607. DOI: 10.1109/CAMA49227.2021.9703473.

3. Gupta R.C., Sagi S.K., Raja K.P., Sharma N.K., Jyoti R. Shaped prime-focus reflector antenna for satellite communication. IEEE Antennas and Wireless Propagation Letters, 2017, vol. 16, pp. 1945–1948. DOI: 10.1109/LAWP.2017.2689800.

4. Rotshild D., Abramovich A. Ultra-wideband reconfigurable X-band and Ku-band metasurface beam-steerable reflector for satellite communications. Electronics, 2021, vol. 10, art. 2165. DOI: 10.3390/electronics10172165.

5. Rahmat-Samii Y., Haupt R. Reflector antenna developments: a perspective on the past, present and future. IEEE Antennas and Propagation Magazine, 2015, vol. 57, no. 2, pp. 85–95. DOI: 10.1109/MAP.2015.2414534.

6. Rebollo A., Vaquero A.F., Arrebola M., Pino M.R. 3D-printed dual-reflector antenna with self-supported dielectric subreflector. IEEE Access, 2020, vol. 8, pp. 209091–209100. DOI: 10.1109/ACCESS.2020.3038739.

7. He G., Liu M., Gao X., Du X., Zhou H., Zhu H. Chinese deep space stations: a brief review [Antenna Applications Corner]. IEEE Antennas and Propagation Magazine, 2022, vol. 64, no. 1, pp. 102–111. DOI: 10.1109/MAP.2021.3129690.

8. Abulgasem S., Tubbal F., Raad R., Theoharis P.I., Lu S., Iranmanesh S. Antenna designs for CubeSats: a review. IEEE Access, 2021, vol. 9, pp. 45289–45324. DOI: 10.1109/ACCESS.2021.3066632.

9. He G., Gao X., Sun L., Zhang R. A review of multibeam phased array antennas as LEO satellite constellation ground station. IEEE Access, 2021, vol. 9, pp. 147142–147154. DOI: 10.1109/ACCESS.2021.3124318.

10. Liang M., Shemelya C., MacDonald E., Wicker R., Xin H. 3D-printed microwave patch antenna via fused deposition method and ultrasonic wire mesh embedding technique. IEEE Antennas and Wireless Propagation Letters, 2015, vol. 14, pp. 1346–1349. DOI: 10.1109/LAWP.2015.2405054.

11. Shi Y., Zhang X.F., Meng Z.K., Li L. Design of low-RCS antenna using antenna array. IEEE Transactions on Antennas and Propagation, 2019, vol. 67, no. 10, pp. 6484–6493. DOI: 10.1109/TAP.2019.2925180.

12. Azizi Y., Soleimani M., Sedighy S.H., Matekovits L. Low-cost, low-profile wide-band radar cross section reduction using dual-concentric phase-gradient modulated surface. Electronics, 2021, vol. 10, art. 1552. DOI: 10.3390/electronics10131552.

13. Zhang T., Ren J., Yang Y., Pang X., Zheng Q. Ultra-wideband low-RCS and gain-enhanced phase gradient metasurface-based antenna using characteristic mode analysis. AEU – International Journal of Electronics and Communications, 2024, vol. 183, art. 155370. DOI: 10.1016/j.aeue.2024.155370.

14. Tian Z., Yang L., Cheng Z., Cai B., Cheng Y., Luo H., Li X. Broadband low-profile and low-RCS circular polarization metasurface antenna based on hybrid mechanism with characteristic mode analysis. AEU – International Journal of Electronics and Communications, 2025, vol. 202, art. 156043. DOI: 10.1016/j.aeue.2025.156043.

15. Umair H., Latef T.B.A., Yamada Y., Mahadi W.N.L.B.W., Othman M., Kamardin K., Hussein M.I., Najam A.I. Tilted beam Fabry-Perot antenna with enhanced gain and broadband low backscattering. Electronics, 2021, vol. 10, art. 267. DOI: 10.3390/electronics10030267.

16. Harooni S.A., Naser-Moghadasi M., Ghalamkari B., Khatir M. Wideband RCS reduction in microstrip antennas using integrated artificial magnetic conductors and frequency-selective surface. AEU – International Journal of Electronics and Communications, 2026, vol. 205, art. 156135. DOI: 10.1016/j.aeue.2025.156135.

17. Saxena G., Awasthi Y.K., Jain P. Design of metasurface absorber for low RCS and high isolation MIMO antenna for radio location & navigation. AEU – International Journal of Electronics and Communications, 2021, vol. 133, art. 153680. DOI: 10.1016/j.aeue.2021.153680.

18. Han J., Jia D., Du B., Han G., Jia Y., Zhao Z. Design of broadband low-RCS array antennas based on characteristic mode cancellation. Electronics, 2023, vol. 12, art. 1536. DOI: 10.3390/electronics12071536.

19. Liu Y., Du B., Jia D. Design of broad-angle low-RCS microstrip antenna for dual-polarization using characteristic mode analysis. Electronics, 2025, vol. 14, art. 2121. DOI: 10.3390/electronics14112121.

20. Ullah M.U., Latef T.B.A., Othman M., Hussein M.I., Alkhoori H.M., Yamada Y., Kamardin K., Khalid R. A progression in the techniques of reducing RCS for the targets. Alexandria Engineering Journal, 2024, vol. 100, pp. 153–169. DOI: 10.1016/j.aej.2024.05.001.

21. Alhaj Hasan A.F., Nguyen M.T., Gazizov T.R. Modelling and designing wire-grid sparse antennas using MoM-based approaches for enhanced performance and reduced cost. Mikrotalasna Revija (Microwave Review), 2023, vol. 29, no. 2, pp. 83–94. DOI: 10.18485/mtts_mr.2023.29.2.10.

22. Alhaj Hasan A.F., Nguyen M.T., Gazizov T.R. Novel MoM-based approaches for generating wire-grid sparse antenna structures. 2023 IEEE 24th International Conference of Young Professionals in Electron Devices and Materials (EDM), Novosibirsk, Russian Federation, IEEE, 2023, pp. 570–576. DOI: 10.1109/EDM58354.2023.10225219.

23. Nguyen M.T., Alhaj Hasan A.F., Gazizov T.R. Generating sparse wire-grid antennas using maximum current-based optimal current grid approximation. IEEE Open Journal of Antennas and Propagation, 2026, vol. 7, no. 2, pp. 652–665. DOI: 10.1109/OJAP.2025.3543559.

24. Nguyen M.T., Gazizov T.R. Design of monopole antennas with various grounds for 2.4 GHz applications. AEU – International Journal of Electronics and Communications, 2025, vol. 202, art. 156050. DOI: 10.1016/j.aeue.2025.156050.

25. Dang T.P., Gazizov T.R. [Improving the optimal current grid approximation for sparse scattering structures]. *Radiotekhnika*, 2025, vol. 89, no. 9, pp. 181–192. DOI: 10.18127/j00338486-202509-20 (in Russ.).

26. Lukas K.S., Dang T.F., Gazizov T.R. [Sparse electromagnetic shields obtained from wire grid]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2025, vol. 28, no. 4, pp. 13–19. DOI: 10.21293/1818-0442-2025-28-4-13-19 (in Russ.).

27. Harrington R.F. Matrix methods for field problems. Proceedings of the IEEE, 1967, vol. 55, no. 2, pp. 136–149. DOI: 10.1109/PROC.1967.5433.

28. Parabolic antenna JRB-25 MIMO. Jirous antennas. Available at: https://en.jirous.com/antenna-3,5ghz/jrb-25_mimo/ (accessed: 15.05.2026).

29. Nguyen M.T., Alhaj Hasan A.F., Gazizov T.R. *Sravnitel'nyy analiz kharakteristik razrezhennykh struktur reflektornoy antenny X-diapazona na raznykh chastotakh* [Comparative analysis of the characteristics of sparse X-band reflector antenna structures at different frequencies]. Antenny i rasprostranenie radiovoln [Antennas and radio wave propagation]. Proc. of the All-Russian Scientific and Technical Conf., St. Petersburg, 2024, pp. 20–27. EDN QLBCKE.

Xuan Ba Trinh
Postgraduate Student, Department of Television
and Control (TC), Tomsk State University
of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0009-0000-5711-5878
Phone: +7-920-115-04-68
Email: xuanba6420@gmail.com

Talgat R. Gazizov
Doctor of Engineering, Professor,
Head of the Department of TC, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0002-1192-4853
Phone: +7-913-826-07-24
Email: talgat.r.gazizov@tusur.ru

Received: 23.06.2026.
Accepted: 01.09.2026.

УДК 621.396.24
Шифр специальности ВАК: 2.2.13, 2.3.1

А.Д. Скорик

Использование мониторинга помех и неоднородностей ионосферы для повышения надежности коротковолновой связи

Актуальность. Надёжность радиосвязи в диапазоне коротких волн является ключевым параметром систем дальней связи, а статические модели расчёта, регламентированные рекомендациями ITU-R, не могут гарантировать точность прогноза из-за изменчивости реальных условий ионосферного канала. ***Цель исследования.*** Провести анализ существующих методик оценки надёжности КВ-радиолиний, основанных на рекомендациях ITU-R, показать разброс оценок при различной полноте учёта факторов и обосновать необходимость перехода к адаптивным системам, использующим данные мониторинга ионосферы. ***Методы.*** Сравнительное моделирование трёх подходов на примере радиолинии с телеграфией F1B: базового (по Рекомендации МСЭ-R P.842-5 с фиксированным допустимым отношением сигнал/шум из Рекомендации МСЭ-R F.339), уточнённого (с учётом частотной зависимости вариаций помех) и объединённого (включающего также частотно-зависимое допустимое отношение сигнал/шум, обусловленное диффузностью ионосферы). ***Научная новизна***. Разработан объединённый подход к расчету надежности КВ-связи, отличающийся от известных изолированных подходов (прежде всего ITU-R) установлением сквозной аналитической связи между динамическим энергетическим порогом $Z_{доп}(f, \beta_и)$, рабочей частотой (f), текущим состоянием диффузности ионосферы ($\beta_и$) и частотно-зависимым СКО флуктуаций помех ($\sigma_z(f)$). Впервые показано, что совместный учет этих факторов выявляет немонотонный характер зависимости надежности от частоты, что позволяет предиктивно определять оптимальную рабочую частоту для конкретных гелиогеофизических условий. ***Результаты.*** Базовый подход даёт максимальную надёжность около 64% на максимально применимой частоте; учёт частотной зависимости помех повышает оценки на низких частотах; объединённый подход приводит к значениям надёжности 80–87% с выраженным оптимумом в области 2–3 МГц. Разброс оценок достигает десятков процентных пунктов. ***Практическая значимость.*** Результаты расчётов обосновывают достижимость надёжности КВ-связи на уровне 86,8% для рассмотренных гелиогеофизических условий и подтверждают необходимость перехода к проактивным адаптивным системам КВ-связи с постоянным пассивным мониторингом состояния канала.

Ключевые слова: КВ-радиосвязь, надёжность связи, диффузность ионосферы, допустимое отношение сигнал/шум, атмосферные помехи, адаптивная радиосвязь, пассивный мониторинг ионосферы, рекомендации ITU-R.
DOI: 10.21293/1818-0442-2026-29-2-57-63

Введение

Надёжность радиосвязи в диапазоне коротковолновых (КВ) частот (1,5–30 МГц) является ключевым параметром для систем дальней связи, особенно в условиях отсутствия альтернативных каналов. Ионосферный канал характеризуется значительной изменчивостью, обусловленной зависимостью параметров ионосферы от солнечной активности, времени суток, сезона и географического положения [1]. Для описания регулярных (медианных) свойств ионосферы разработаны глобальные эмпирические модели, например, международная справочная модель ионосферы IRI [2]. Прогнозирование надёжности требует учёта множества факторов, включая статистические характеристики сигнала и помех, поглощение радиоволн в ионосфере [3], замирания и возможности разнесённого приёма.

Международным союзом электросвязи (МСЭ) разработан ряд рекомендаций, регламентирующих методы расчёта надёжности КВ-радиолиний: МСЭ-R P.533 [4] – прогнозирование медианных уровней сигнала, МСЭ-R P.372 [5] – статистика шумов (помех), МСЭ-R P.842 [6] – расчёт основной надёжности линии (BCR – basic circuit reliability). В рекомендации МСЭ-R F.339 [7] приведены требуемые отношения сигнал/шум для различных классов излучения и условий.

Однако, как показывают современные исследования [8, 9], прямое использование табличных значений без учёта частотной зависимости как допустимого отношения сигнал/помеха, так и статистических характеристик помех может приводить к существенным погрешностям расчёта надёжности КВ-связи. В работе [8] установлена связь допустимого отношения сигнал/помеха с уровнем диффузности ионосферы и рабочей частотой, а в [9] показано, что среднеквадратическое отклонение (СКО) отношения сигнал/помеха также зависит от частоты из-за вариаций атмосферных помех.

Цель настоящей статьи – на примере конкретной радиолинии продемонстрировать разброс оценок надёжности КВ-радиолинии, получаемых при различной полноте учёта факторов, обосновать новый объединенный критерий оперативной надежности, а также доказать необходимость перехода к адаптивным системам, использующим пассивный мониторинг ионосферы для оперативной коррекции рабочих параметров.

Описание общих подходов

Согласно [6], основная надёжность КВ-радиолинии BCR определяется как вероятность превышения $(S/N)_{med} \geq (S/N)_r$ или непревышения $(S/N)_{med} < (S/N)_r$ медианным отношением сигнал/помеха $(S/N)_{med}$ на входе приёмника требуемого (допустимого) отношения сигнал/шум $(S/N)_r$ согласно выражению (1):

$$\mathrm{BCR}=\begin{cases}130-\dfrac{80}{1+\dfrac{\left(\dfrac{S}{N}\right)_{\mathrm{med}}-\left(\dfrac{S}{N}\right)_{\mathrm{r}}}{D_{\mathrm{u}}(SN)}}\%, \\ \text{при}\left(\dfrac{S}{N}\right)_{\mathrm{med}}\geq\left(\dfrac{S}{N}\right)_{\mathrm{r}}; \\ \max\left(\dfrac{80}{1+\dfrac{\left(\dfrac{S}{N}\right)_{\mathrm{med}}-\left(\dfrac{S}{N}\right)_{\mathrm{r}}}{D_{\mathrm{u}}(SN)}}-30{,}0\right)\%, \\ \text{при}\left(\dfrac{S}{N}\right)_{\mathrm{med}}<\left(\dfrac{S}{N}\right)_{\mathrm{r}}, \end{cases} \quad (1)$$

где $D_{\mathrm{u}}(SN)$ – верхнее децильное отклонение среднего отношения сигнал/помеха, характеризующее разброс значений сигнал/помеха. Величины $(S/N)_{\mathrm{med}}$ и $D_{\mathrm{u}}(SN)$ рассчитываются с использованием данных о сигнале по [4], о помехах и шумах – по [5], а также децильных отклонений сигнала и шума.

Рассмотрим КВ-радиолинию с параметрами, близкими по значениям к использованным в примерах расчётов из [8, 9]:

1) координаты точки отражения 56,84° с.ш., 53,78° в.д.;

2) дата 08.08.2019 (лето);

3) время 00:00 местного (21:00 UTC);

4) дальность 2000 км;

5) мощность передатчика 1 кВт (30 дБВт), коэффициенты усиления передающей G_{t} = 3 дБ и приёмной G_{r}=2 дБ антенн (индексы t и r – от англ. transmitting и receiving), полоса приёмника B = 180 Гц, сельская местность (мощность индустриальных помех 5 дБ).

Анализ проведём в диапазоне частот от f = 1 МГц до 10 МГц; максимальная применимая частота (МПЧ) $f_{\mathrm{МПЧ}}$ = 10 МГц. В настоящей работе значения $(S/N)_{\mathrm{med}}(f)$ получены расчётом по методике [4]. Результаты расчёта представлены в табл. 1.

Для определения требуемого $(S/N)_{\mathrm{r}}$ воспользуемся данными рекомендации [7, табл. 1] для класса излучения F1B MFSK 6-tone ITA2 10 знаков/с – одного из широко распространённых сервисов Глобальной морской системы связи при бедствии и для обеспечения безопасности судоходства (ГМССБ). Документ устанавливает, что при обеспечении вероятности ошибки P_{c} = 0,01 и использовании двойного разнесённого приёма (dual diversity) необходимое отношение сигнал/помеха $(S/N)_{\mathrm{r}}$ = 31 дБ, а без разнесённого приёма это значение должно быть увеличено на 10 дБ, т.е. составит 41 дБ.

Для обеспечения корректности последующего сравнительного анализа параметры сигнала, использованные при расчетах, были максимально согласованы с условиями из [8, 9] и практических ситуаций для ГМССБ. В [8, 9] использовались сигналы частотной телеграфии с шириной полосы сигнала не более 300 Гц, что физически сопоставимо с предписанной рекомендацией МСЭ-R F.339-7 и стандартами ГМССБ полосой приемника 180 Гц для класса излучения F1B MFSK 6-tone ITA2 10 знаков/с, причем требования к верности приема приняты на уровне $P_{\mathrm{ош}} = 10^{-2}$.

Переход к требуемой вероятности ошибки $P_{\mathrm{ош}} = 10^{-3}$ ужесточит требования не менее чем на 4 дБ. При этом МСЭ-R F.339-7 однозначно задает требования лишь к одиночному и двойному разнесенному приему, а состояние ионосферы учитывается качественно (или обычное, или возмущенное), предлагаемый в методике критерий $D_{\mathrm{св}}$ через учет параметра диффузности $\beta_{\mathrm{и}}$ позволяет проводить такую оценку в более широких пределах, опираясь на конкретные данные пассивного мониторинга ионосферы, что подтверждает дополнительные преимущества от перехода к динамическому критерию $D_{\mathrm{св}}$.

Децильные отклонения сигнала $D_{\mathrm{u}}S_{\mathrm{d}}$ определены по табл. 2 Рекомендации [6] для геомагнитной широты < 60°: для $f/f_{\mathrm{МПЧ}} \leq 0{,}8$ принимается $D_{\mathrm{u}}S_{\mathrm{d}} = 6$ дБ (суммарное верхнее децильное отклонение отношения сигнал/помеха, возникающее из-за изменений условий ото дня ко дню в течение месяца), при $f/f_{\mathrm{МПЧ}} = 0{,}9\ldots1$ – $D_{\mathrm{u}}S_{\mathrm{d}} = 8$ дБ. Краткосрочные отклонения (в пределах одного конкретного часа) $D_{\mathrm{u}}S_{\mathrm{h}} = 8$ дБ, как того требует Рекомендация [6]. В первом приближении вариации шума не учитывались ($N_{\mathrm{u}} = N_{\mathrm{med}}$), поэтому

$$D_{\mathrm{u}}(SN) = \sqrt{\left(D_{\mathrm{u}}S_{\mathrm{d}}\right)^2 + \left(D_{\mathrm{u}}S_{\mathrm{h}}\right)^2}. \quad (2)$$

Таблица 1

Медианные значения $(S/N)_{\mathrm{med}}$ для различных частот (расчёт по [4])

f, МГц	$f/f_{\mathrm{МПЧ}}$	$(S/N)_{\mathrm{med}}$, дБ
1,0	0,1	22,5
2,0	0,2	25,5
3,0	0,3	27,4
4,0	0,4	28,4
5,0	0,5	28,5
6,0	0,6	29,4
7,0	0,7	31,6
8,0	0,8	32,4
9,0	0,9	33,4
10,0	1,0	33,5

Таблица 2

Базовый упрощённый расчёт надёжности по методике [6] (методика 1)

f, МГц	$f/f_{\mathrm{МПЧ}}$	$(S/N)_{\mathrm{med}}$, дБ	$D_{\mathrm{u}}(SN)$, дБ	$\Delta = (S/N)_{\mathrm{med}} - 31$, дБ	$\Delta = 31 - (S/N)_{\mathrm{med}}$, дБ	BCR, %
1,0	0,1	22,5	10,0	–	8,5	13,2
2,0	0,2	25,5	10,0	–	5,5	21,6
3,0	0,3	27,4	10,0	–	3,6	28,8
4,0	0,4	28,4	10,0	–	2,6	33,5
5,0	0,5	28,5	10,0	–	2,5	34,0
6,0	0,6	29,4	10,0	–	1,6	39,0
7,0	0,7	31,6	10,0	0,6	–	54,5
8,0	0,8	32,4	10,0	1,4	–	59,8
9,0	0,9	33,4	11,3	2,4	–	64,0
10,0	1,0	33,5	11,3	2,5	–	64,5

На основе частотных зависимостей $(S/N)_{med}(f)$ и $D_u(SN)(f)$ производился расчёт основной надёжности КВ-связи BCR(f) согласно (1).

Результаты расчёта BCR с применением базового подхода (методика 1)

Результаты расчёта по базовой методике [6] приведены в табл. 2.

Максимальная надёжность, рассчитанная по методике 1, достигается на частоте $f_{МПЧ}$ = 10 МГц (64,5%), однако на практике работа на частотах, равных МПЧ или превышающих её, часто запрещена, и диапазон рабочих частот обычно ограничен значением $f/f_{МПЧ}$ = 0,9. Следует отметить, что при отсутствии разнесённого приёма требуемое $(S/N)_г$ было бы на 10 дБ выше, что привело бы к существенно более низким значениям BCR. Таким образом, применение разнесённого приёма является важным условием достижения приемлемой надёжности в данном случае.

Расчёт надёжности с учётом частотной зависимости помех (методика 2)

В работе [9] показано, что децильные отклонения отношения сигнал/помеха существенно зависят от частоты из-за частотной зависимости вариаций атмосферных помех. Согласно [6; табл. 1, этап 6], полное выражение для $D_u(SN)$ должно включать вклад от флуктуаций шума

$$D_u(SN) = \sqrt{(D_u Sd)^2 + (D_u Sh)^2 + \left[10\lg\left(\frac{N_u}{N_{med}}\right)\right]^2}, \quad (3)$$

где N_u и N_{med} – мощности шума, соответствующие верхней децили и медиане соответственно. Эти величины зависят от частоты, времени суток, сезона и географического положения и определяются по Рекомендации [5]. В работе [9] на основе Рекомендации [5] получены частотно-зависимые значения СКО отношения сигнал/помеха $\sigma_z(f)$ для условий летней ночи [9; рис. 3, *б*]. Верхнее децильное отклонение $D_u(SN)$ согласно соотношению, обоснованному в [9], рассчитывается как (4)

$$D_u(SN)(f) = 1{,}28\cdot\sigma_z(f). \quad (4)$$

Полученные значения $\sigma_z(f)$ приведены в третьем столбце табл. 3, а результаты расчёта BCR – в последнем столбце. Расчёт возможен, поскольку моделируемые условия весьма близки; BCR рассчитывается по формуле (1) с теми же $(S/N)_{med}$.

Максимум расчётной надёжности достигается на частоте $f_{МПЧ}$ = 10 МГц (61,8%). Сравнение с табл. 2 показывает, что учёт помеховых вариаций повышает оценку на низких частотах (например, с 13,2 до 20,1% на частоте 1 МГц) и незначительно изменяет её на высоких. Это объясняется более реалистичным описанием разброса отношения сигнал/шум, который на низких частотах увеличивается из-за сильных вариаций атмосферных помех.

Расчёт BCR на основе объединённого подхода (методика 3)

В работе [8] допустимое отношение сигнал/помеха $Z_{доп}$ (фактически это $(S/N)_г$ по [6]) ставится в зависимость от рабочей частоты и уровня диффузности ионосферы $\beta_и$ (интенсивности мелкомасштабных неоднородностей). Для умеренной диффузности ($\beta_и = 10^{-2}$) значения $Z_{доп}(f)$ взяты из [8, рис. 4, *а*] для ночных условий; для частот, не указанных явно на графике, использована линейная интерполяция. Подставив их в формулу надёжности вместе с $\sigma_z(f) = D_u(SN)/1{,}28$ из [9], получим

$$D_{св} = F\left(\frac{\left(S/N\right)_{med} - Z_{доп}(f, \beta_и)}{\sigma_z(f)}\right), \quad (5)$$

где $F(\cdot)$ – функция Лапласа (интеграл вероятности). Результаты расчёта для частот f = 1–10 МГц приведены в табл. 4.

Необходимо подчеркнуть, что в отличие от базовых рекомендаций ITU-R, использующих консервативный статический порог для ситуации типа «наихудший случай», и частных методик [8, 9], предложенный критерий оперативной надежности ($D_{св}$) (5) выступает функцией динамического состояния канала и зависит от состояния ионосферы в конкретный момент времени, т.е. он позволяет связать параметры радиолинии с физикой среды распространения, что, в свою очередь, позволяет непосредственно в процессе пассивного мониторинга использовать оценку надежности как предиктивный инструмент управления частотным и энергетическим ресурсом.

Результат демонстрирует качественно иное поведение расчётной надёжности КВ-радиолинии: появляется ярко выраженный максимум в области 2–3 МГц (86,8% на 3 МГц, 86,5% на 2 МГц), и в целом диапазон расчётных значений надёжности (80–87%) значительно превышает полученные в табл. 2 и 3 значения (13–64%). Это объясняется двумя обстоятельствами:

– адаптивный (нестатичный) учёт требуемого $(S/N)_г$, зависящего от реального уровня диффузности в канале, позволяет использовать значительно меньшее допустимое отношение сигнал/шум, которое оказывается гораздо ниже консервативной оценки 31 дБ, заложенной в [7];

– учёт влияния уровня помех и частотной зависимости СКО сигнал/помеха приводит к дополнительному выигрышу в области низких частот.

Столь существенное влияние зависимости допустимого отношения сигнал/шум $Z_{доп}(f, \beta_и)$ от выбора рабочей частоты f и уровня диффузности ионосферы $\beta_и$ на основную надёжность КВ-связи обусловливает необходимость использования в комплексах КВ-связи аппаратно-программных комплексов (АПК) пассивного мониторинга интенсивности мелкомасштабных неоднородностей электронной концентрации в ионосфере [10, 11] во всём диапазоне рабочих частот.

Обсуждение результатов

Сравнительный анализ трёх методик (табл. 5, рис. 1) показывает, что разброс оценок надёжности для одной и той же радиолинии может достигать десятков процентных пунктов. Методики 1 и 2, использующие фиксированный порог $(S/N)_г$ = 31 дБ (напомним, достигнутый за счёт двойного разнесённого приёма), дают близкие результаты, причём их кривые почти совпадают.

Таблица 3

Расчёт надёжности с учётом частотной зависимости помех (по данным [9])

f, МГц	$(S/N)_{med}$, дБ	$\sigma_z(f)$ [9], дБ	$D_u(SN)=1{,}28\sigma_z$, дБ	$31-(S/N)_{med}$, дБ	$(S/N)_{med}-31$, дБ	BCR, %
1,0	22,5	11,4	14,6	8,5	–	20,1
2,0	25,5	11,3	14,5	5,5	–	27,8
3,0	27,4	11,1	14,2	3,6	–	33,6
4,0	28,4	10,8	13,8	2,6	–	37,1
5,0	28,5	10,6	13,6	2,5	–	37,3
6,0	29,4	10,3	13,2	1,6	–	41,2
7,0	31,6	10,2	13,1	–	0,6	53,2
8,0	32,4	10,3	13,2	–	1,4	57,5
9,0	33,4	10,8	13,8	–	2,4	61,6
10,0	33,5	10,8	13,8	–	2,5	61,8

Таблица 4

Расчёт надёжности с учётом объединения частотно-зависимых $Z_{доп}$ и σ_z (методика 3), по данным [8, 9]

f, МГц	$(S/N)_{med}$, дБ	$Z_{доп}(f)$, дБ	$D_u(SN)$ [9], дБ	$\sigma_z(f)$, дБ	ξ	$D_{св}$, %
1,0	22,5	12	14,6	11,4	0,921	82,1
2,0	25,5	13	14,5	11,3	1,106	86,6
3,0	27,4	15	14,2	11,1	1,117	86,8
4,0	28,4	17	13,8	10,8	1,056	85,4
5,0	28,5	18	13,6	10,6	0,991	83,9
6,0	29,4	19	13,2	10,3	1,010	84,4
7,0	31,6	21	13,1	10,2	1,039	85,1
8,0	32,4	22	13,2	10,3	1,010	84,4
9,0	33,4	23	13,8	10,8	0,963	83,2
10,0	33,5	24	13,8	10,8	0,88	81,1

Таблица 5

Сравнение рассчитанной основной надёжности для различных методик, %

Модель	1 МГц	2 МГц	3 МГц	4 МГц	5 МГц	6 МГц	7 МГц	8 МГц	9 МГц	10 МГц
Базовая ITU-R (см. табл. 2)	13,2	21,6	28,8	33,5	34,0	39,0	54,5	59,8	64,0	64,5
С учётом помех [9] (см. табл. 3)	20,1	27,8	33,6	37,1	37,3	41,2	53,2	57,5	61,6	61,8
Объединённая [8, 9] (см. табл. 4)	82,1	86,6	86,8	85,4	83,9	84,4	85,1	84,4	83,2	81,1

Небольшие различия между ними обусловлены только учётом частотной зависимости вариаций помех в методике 2, что несколько корректирует разброс, но не меняет принципиально характер зависимости: монотонный рост с частотой и максимум на МПЧ. Обращает на себя внимание небольшой перегиб в районе 7 МГц, связанный с переходом от формулы для $(S/N)_{med}<(S/N)_r$ к формуле для случая превышения порога $(S/N)_{med}\geq(S/N)_r$.

Методика 3 принципиально отличается: она учитывает частотно-зависимое допустимое отношение сигнал/шум $Z_{доп}(f)$, которое на низких и средних частотах оказывается значительно ниже фиксированного порога 31 дБ.

Благодаря этому на частотах 2–7 МГц достигается высокая надёжность (80–87%) даже без использования разнесённого приёма. На частотах, близких к МПЧ, значение $Z_{доп}(f)$ возрастает, приближаясь к 31 дБ, и надёжность методики 3 сближается с надёжностью методик 1 и 2.

Важно подчеркнуть, что методика 3 не является простым математическим усложнением. Она отражает принципиальный сдвиг в подходе к проектированию КВ-радиолиний: от жёсткого нормирования, основанного на фиксированных таблицах, к динамическому учёту реальной физической обстановки.

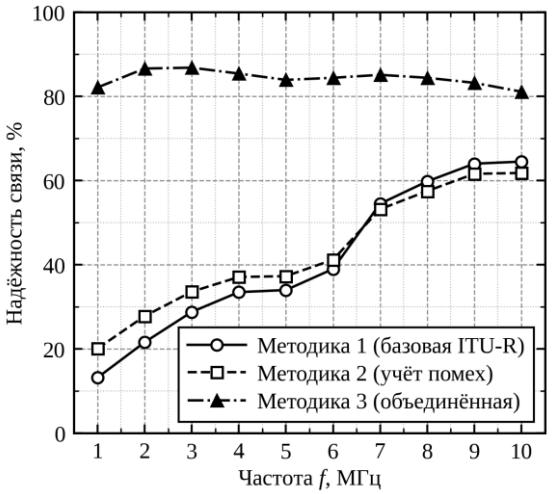


Рис. 1. Частотные зависимости надёжности КВ-связи, рассчитанные по различным методикам

В методиках 1 и 2 допустимое отношение сигнал/шум задаётся априори (как константа, достигнутая аппаратной реализацией), тогда как методика 3 исходит из того, что этот порог сам является функцией состояния ионосферы и помеховой обстановки. Это отражение физической реальности: ионосфера нестационарна, и её свойства меняются во времени и

по частоте. Именно поэтому реализация такого подхода невозможна без непрерывного мониторинга: только зная текущие значения диффузности и вариаций помех, можно обоснованно выбирать рабочую частоту, мощность и другие параметры канала. Мониторинг ионосферы и помеховой обстановки превращает потенциальные преимущества методики 3 в реальный инструмент повышения надёжности и эффективности использования частотного ресурса.

Современные комплексы радиоконтроля, такие как VRB-12-141 [12], обеспечивают автоматическое измерение отношения сигнал/шум, анализ спектра и идентификацию типов сигналов, что создаёт техническую основу для внедрения адаптивных алгоритмов выбора частоты. Похожие возможности реализованы и в отечественных разработках (например, в комплексе адаптивной радиосвязи ПИРС АО «РИМР» или в комплексах радиоконтроля типа ЧВА-001-04РП-Э НИИ «Вектор»), что подтверждает практическую реализуемость предлагаемого подхода.

Таким образом, практическая надёжность перестаёт быть фиксированной характеристикой, закладываемой на этапе проектирования, и становится функцией эффективности адаптивных алгоритмов. Представленные в статье модели служат важным инструментом предварительного анализа и выбора потенциально пригодных диапазонов частот, однако окончательное решение о параметрах связи должно приниматься с учётом данных реального времени.

Заключение

Выполнен анализ существующих методов оценки надёжности КВ-радиолиний, основанных на рекомендациях МСЭ-R. На примере линии с телеграфией F1B MFSK 6-tone ITA2 продемонстрировано следующее.

1. Базовая методика МСЭ-R P.842-5 с фиксированным допустимым отношением сигнал/шум из МСЭ-R F.339 даёт максимальную надёжность около 64% на МПЧ с учётом двойного разнесения по приёму (см. табл. 2).

2. Учёт частотной зависимости децильных отклонений (за счёт вариаций помех) по [9] повышает оценки на низких частотах, но максимум остаётся на МПЧ – 61,8% (см. табл. 3).

3. Введение частотно-зависимого допустимого отношения сигнал/шум, обусловленного диффузностью ионосферы [8], в комбинации с уточнённым СКО приводит к изменениям зависимости надежности связи от частоты. Для конкретных гелиогеофизических условий достигнуты значения надёжности 80–87% с ярко выраженным оптимумом в области 2–3 МГц, т.е. при $f/f_{МПЧ} = 0{,}2–0{,}3$ (см. табл. 4).

Разброс оценок, достигающий десятков процентных пунктов, указывает на ограниченную применимость статических моделей для гарантированного обеспечения качества связи. Переход к адаптивным системам, использующим пассивный мониторинг ионосферы и помеховой обстановки, позволяет подстраиваться под текущие условия и реализовывать потенциальные возможности, заложенные в физически обоснованных моделях. Мониторинг даёт возможность динамически выбирать рабочую частоту, мощность и схему связи, обеспечивая требуемую надёжность даже при изменении условий распространения. Современные технические средства [11, 12] и отечественные разработки подтверждают возможность практической реализации такого подхода.

Расширение и распространение результатов на другие гелиогеофизические условия (иные уровни диффузности, время суток, сезоны) составляет предмет дальнейших исследований.

Необходимость постоянного мониторинга состояния ионосферы в канале связи и оптимизации параметров непосредственно под текущие условия отмечается и в последних публикациях зарубежных исследователей [13–16].

Литература

1. Coster A.J. Ionospheric irregularities, space weather impact on GNSS, and spacecraft anomalies // Proceedings of the United Nations/Germany Workshop on the International Space Weather Initiative. ISWI 2024. – Singapore: Springer, 2026. – P. 131–141. DOI: 10.1007/978-981-95-1121-1_15.

2. The International Reference Ionosphere model: a review and description of an ionospheric benchmark / D. Bilitza, M. Pezzopane, V. Truhlik, D. Altadill, B.W. Reinisch, A. Pignalberi // Reviews of Geophysics. – 2022. – Vol. 60, No. 4. – Art. no. e2022RG000792. DOI: 10.1029/2022RG000792.

3. Calculating the absorption of HF radio waves in the ionosphere / K.A. Zawdie, D.P. Drob, D.E. Siskind, C. Coker // Radio Science. – 2017. – Vol. 52, No. 6. – P. 767–783. DOI: 10.1002/2017RS006256.

4. ITU-R Recommendation P.533-14. Method for the prediction of the performance of HF circuits. – Geneva: ITU, 2019. – 42 p.

5. ITU-R Recommendation P.372-15. Radio noise. – Geneva: ITU, 2021. – 58 p.

6. ITU-R Recommendation P.842-5. Computation of reliability and compatibility of HF radio systems. – Geneva: ITU, 2013. – 16 p.

7. ITU-R Recommendation F.339-7. Bandwidths, signal-to-noise ratios and fading allowances in complete systems. – Geneva: ITU, 2006. – 12 p.

8. Зависимость надежности связи в декаметровой радиолинии от выбора рабочей частоты с учетом сигнально-помеховой обстановки и диффузности ионосферы / В.П. Пашинцев, А.Д. Скорик, С.А. Коваль, Д.П. Киселев, М.А. Сенокосов // Системы управления, связи и безопасности. – 2019. – № 4. – С. 300–322. DOI: 10.24411/2410-9916-2019-10412.

9. Методика определения зависимости среднеквадратического отклонения отношения сигнал-помеха в декаметровой радиолинии от выбора частоты / В.П. Пашинцев, Е.М. Гринев, А.Д. Скорик, Д.А. Белоконь // Системы управления, связи и безопасности. – 2023. – № 2. – С. 63–80. – DOI: 10.24412/2410-9916-2023-2-63-80.

10. Метод определения величины интенсивности неоднородностей по данным ионосферного зондирования / В.П. Пашинцев, А.В. Омельчук, С.А. Коваль, Ю.И. Галушко // Двойные технологии. – 2009. – № 1(46). – С. 38–41.

11. Аппаратно-программный комплекс для оперативного мониторинга ионосферы в интересах обеспечения коротковолновой связи / В.М. Смирнов, Е.В. Смирнова,

В.Н. Скобелкин, С.И. Тынянкин // Радиолокация и радиосвязь: матер. IV Всерос. конф. – М.: ИРЭ РАН, 2010. – С. 43–47.

12. VRB-12-141/A(V/UHF) V/UHF Radio Monitoring [Электронный ресурс]. – URL: https://mindex-center.ir/en/products/vrb-12-141-avuhf-vuhf-radio-monitoring (дата обращения: 08.03.2026).

13. Capacity and reliability of ionosphere communication channel based on multi-carrier modulation technique and LUF–MUF variation / V. Dear, A.S. Mardiani, N. Dedi, P. Abadi, B.H. Prananto, Iskandar // IEICE Transactions on Communications. – 2024. – Vol. E107-B, No. 4. – P. 357–367. DOI: 10.1587/transcom.2023EBP3099.

14. A frequency selecting method for high-frequency communication based on ionospheric oblique backscatter sounding / C. Cai, G. Yang, T. Liu, C. Jiang // Remote Sensing. – 2024. – Vol. 16, No. 21. – Art. no. 4095. DOI: 10.3390/rs16214095.

15. Development of methods for spectral monitoring of HF interference to determine the availability of ionospheric radio channels / V.A. Ivanov, D.V. Ivanov, N.V. Ryabova, R.R. Belgibaev // 2024 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF). – St. Petersburg, Russia: IEEE, 2024. – P. 1–5. DOI: 10.1109/WECONF61770.2024.10581396.

16. Chen Z. Adaptive modulation and coding based on data-driven learning algorithm in polar HF OFDM-CPM system / Z. Chen, X. Zhang // 2024 5th International Conference on Computer Engineering and Application (ICCEA). – Hangzhou, China: IEEE, 2024. – P. 485–488. DOI: 10.1109/ICCEA62105.2024.10604266.

Скорик Александр Дмитриевич
Генеральный директор АО «Концепт»
Лесная ул. (мкр. Юбилейный), 14а, г. Королёв,
Московская обл., Россия, 141092
ORCID: 0009-0005-6204-9445
Тел.: +7-929-638-63-60
Эл. почта: alexander_skorik@mail.ru

Поступила в редакцию: 26.07.2026.
Принята к публикации: 21.09.2026.

Skorik A.D.

Using monitoring of interference and ionospheric irregularities to improve the reliability of high-frequency communication

Relevance. The reliability of high-frequency (HF) radio communication is a key parameter of long-range communication systems. The ionospheric channel is highly variable, and static models regulated by ITU-R recommendations cannot guarantee prediction accuracy under real conditions. ***Purpose of the study***. To analyze the existing ITU-R-based methods for estimating the reliability of HF radio links, to demonstrate the spread of estimates under differing degrees of factor consideration, and to justify the transition to adaptive systems employing ionospheric monitoring data. ***Methods.*** The study employs comparative modeling of three approaches using an F1B radiotelegraphy link as a case study: the basic one (according to Recommendation ITU-R P.842-5 with a fixed required signal-to-noise ratio from Recommendation ITU-R F.339), a refined one (taking into account the frequency dependence of interference variations) and a combined one (also incorporating the frequency-dependent required signal-to-noise ratio caused by ionospheric diffuseness). ***Scientific Novelty.*** A combined approach to calculating HF communication reliability has been developed. Unlike existing isolated methods (primarily ITU-R), this approach establishes an end-to-end analytical relationship between the dynamic energy threshold $Z_{perm}(f, \beta_i)$, the operating frequency (f), the current state of ionospheric diffuseness (β_i), and the frequency-dependent standard deviation of interference fluctuations ($\sigma_z(f)$). For the first time, it is demonstrated that the combined consideration of these factors reveals a non-monotonic relationship between reliability and frequency, enabling the predictive determination of the optimum operating frequency for specific heliogeophysical conditions. ***Results.*** The basic approach yields a maximum reliability of approximately 64% at the maximum usable frequency; accounting for the frequency dependence of interference improves estimates at lower frequencies; the combined approach results in reliability values of 80–87% with a pronounced optimum at 2–3 MHz. The spread of estimates reaches tens of percentage points. ***Practical significance.*** The calculation results substantiate that an HF communication reliability of 86.8% is achievable under the considered heliogeophysical conditions and confirm the necessity of transitioning to proactive adaptive HF communication systems featuring continuous passive channel state monitoring.
Keywords: HF radio communication, communication reliability, ionospheric diffuseness, required signal-to-noise ratio, atmospheric radio noise, adaptive radio communication, passive ionospheric monitoring, ITU-R recommendations.
DOI: 10.21293/1818-0442-2026-29-2-57-63

References

1. Coster A.J. Ionospheric irregularities, space weather impact on GNSS, and spacecraft anomalies. Proceedings of the United Nations/Germany Workshop on the International Space Weather Initiative. ISWI 2024. Singapore, Springer, 2026, pp. 131–141. DOI: 10.1007/978-981-95-1121-1_15.

2. Bilitza D., Pezzopane M., Truhlik V., Altadill D., Reinisch B.W., Pignalberi A. The International Reference Ionosphere model: a review and description of an ionospheric benchmark. *Reviews of Geophysics*, 2022, vol. 60, no. 4, e2022RG000792. DOI: 10.1029/2022RG000792.

3. Zawdie K.A., Drob D.P., Siskind D.E., Coker C. Calculating the absorption of HF radio waves in the ionosphere. *Radio Science*, 2017, vol. 52, no. 6, pp. 767–783. DOI: 10.1002/2017RS006256.

4. ITU-R Recommendation P.533-14. Method for the prediction of the performance of HF circuits. Geneva, ITU, 2019, 42 p.

5. ITU-R Recommendation P.372-15. Radio noise. Geneva, ITU, 2021, 58 p.

6. ITU-R Recommendation P.842-5. Computation of reliability and compatibility of HF radio systems. Geneva, ITU, 2013, 16 p.

7. ITU-R Recommendation F.339-7. Bandwidths, signal-to-noise ratios and fading allowances in complete systems. Geneva, ITU, 2006, 12 p.

8. Pashintsev V.P., Skorik A.D., Koval S.A., Kiselev D.P., Senokosov M.A. [Decameter radio line communication reliability dependence of the operating frequency choice taking into account the signal-noise situation and ionosphere diffuseness]. *Systems of Control, Communication and Security*, 2019, no. 4, pp. 300–322 (in Russ.). DOI: 10.24411/2410-9916-2019-10412.

9. Pashintsev V.P., Grinev E.M., Skorik A.D., Belokon D.A. [Method for determining the dependence of the standard

deviation from the signal-interference in a decameter radio link on the choice of frequency]. *Systems of Control, Communication and Security*, 2023, no. 2, pp. 63–80 (in Russ.). DOI: 10.24412/2410-9916-2023-2-63-80.

10. Pashintsev V.P., Omelchuk A.V., Koval S.A., Galushko Yu.I. [Method of irregularity intensity value determination according to ionosphere sounding]. *Dvoynye tekhnologii*, 2009, no. 1(46), pp. 38–41 (in Russ.).

11. Smirnov V.M., Smirnova E.V., Skobelkin V.N., Tynyankin S.I. *Apparatno-programmnyy kompleks dlya operativnogo mo-nitoringa ionosfery v interesakh obespecheniya korotko-volnovoy svyazi* [Hardware-software complex for operational monitoring of the ionosphere in the interests of providing HF communication]. *Radiolokatsiya i radiosvyaz* [Radar and Radio Communication], Proc. of the IV All-Russian conf., Moscow, IRE RAS, 2010, pp. 43–47 (in Russ.).

12. VRB-12-141/A(V/UHF) V/UHF Radio Monitoring. Available at: https://mindex-center.ir/en/products/vrb-12-141-avuhf-vuhf-radio-monitoring (accessed: 08 March 2026).

13. Dear V., Mardiani A.S., Dedi N., Abadi P., Prananto B.H., Iskandar. Capacity and reliability of ionosphere communication channel based on multi-carrier modulation technique and LUF–MUF variation. *IEICE Transactions on Communications*, 2024, vol. E107-B, no. 4, pp. 357–367. DOI: 10.1587/transcom.2023EBP3099.

14. Cai C., Yang G., Liu T., Jiang C. A frequency selecting method for high-frequency communication based on ionospheric oblique backscatter sounding. *Remote Sensing*, 2024, vol. 16, no. 21, art. no. 4095. DOI: 10.3390/rs16214095.

15. Ivanov V.A., Ivanov D.V., Ryabova N.V., Belgibaev R.R. Development of methods for spectral monitoring of HF interference to determine the availability of ionospheric radio channels. *2024 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, St. Petersburg, Russia, IEEE, 2024, pp. 1–5. DOI: 10.1109/WECONF61770.2024.10581396.

16. Chen Z., Zhang X. Adaptive modulation and coding based on data-driven learning algorithm in polar HF OFDM-CPM system. *2024 5th International Conference on Computer Engineering and Application (ICCEA)*, Hangzhou, China, IEEE, 2024, pp. 485–488. DOI: 10.1109/ICCEA62105.2024.10604266.

Alexander D. Skorik
General Director, JSC «Concept»
14a, Lesnaya St. (Yubileyny microdistrict), Korolev,
Moscow region, Russia, 141092
ORCID: 0009-0005-6204-9445
Phone: +7-929-638-63-60
Email: alexander_skorik@mail.ru

Received: 26.07.2026.
Accepted: 21.09.2026.

УДК 621.396.41
Шифр специальности ВАК: 2.2.13

А.Н. Кравченко, Д.А. Покаместов

Методы кодирования случайных кодов с малой плотностью проверок на четность

Актуальность. Разработка и исследование методов кодирования случайных кодов с малой плотностью проверок на четность (МППЧ) является важной задачей при построении высокоскоростных систем связи, поскольку различные алгоритмы формирования кодовых слов существенно отличаются по вычислительным затратам и требованиям к реализации. ***Цель исследования.*** Сравнительный анализ методов кодирования МППЧ-кодов, заданных случайными разреженными проверочными матрицами, с оценкой их вычислительной сложности и практической производительности. ***Методы.*** Использованы методы линейной алгебры, аналитическая оценка вычислительной сложности, имитационное моделирование и метод Монте-Карло. В качестве исследуемых методов кодирования рассмотрены метод Нила, алгоритмы на основе LU- и LUP-разложений, приведение проверочной матрицы к ступенчатой форме, а также методы нижней и приближенной нижней триангуляции. ***Научная новизна.*** Проведено сравнительное исследование методов кодирования случайных МППЧ-кодов, включающее совместную оценку их аналитической вычислительной сложности и времени программной реализации при одинаковых исходных условиях. ***Результаты.*** Установлены различия в характере роста вычислительных затрат рассмотренных методов при увеличении размера проверочной матрицы. Показано, что методы Нила, LU и приведения к ступенчатой форме характеризуются меньшим числом арифметических операций, тогда как результаты программного моделирования существенно зависят от способа реализации вычислений и возможностей оптимизации матричных операций в MatLab. ***Практическая значимость.*** Полученные результаты могут использоваться при выборе алгоритма кодирования случайных МППЧ-кодов и разработке программных и аппаратных кодеров для высокоскоростных систем связи.

Ключевые слова: LDPC, МППЧ, случайные матрицы, помехоустойчивое кодирование, вычислительная сложность.
DOI: 10.21293/1818-0442-2026-29-2-64-73

Введение

С развитием систем беспроводной связи наблюдается рост требований к пропускной способности, надежности и энергоэффективности передачи, обусловленный увеличением числа подключенных устройств и разнообразием сценариев применения. В рамках перехода к системам связи шестого поколения (6G) предполагается дальнейшее расширение возможностей сетей, включая поддержку массовых машинных коммуникаций, ультранадежной связи с малыми задержками и интеграцию с интеллектуальными сервисами [1]. В этих условиях ключевую роль играют методы помехоустойчивого кодирования, обеспечивающие эффективную передачу информации в сложных каналах связи. Коды с малой плотностью проверок на четность (МППЧ; англ. LDPC, low density parity check) занимают важное место среди современных кодовых конструкций благодаря своей способности обеспечивать производительность, близкую к пределу Шеннона, при относительно низкой вычислительной сложности декодирования [2]. В системах 5G LDPC-коды используются для кодирования пользовательских данных, а в контексте 6G рассматриваются их дальнейшие модификации и адаптивные версии, ориентированные на работу в сценариях, характерных для перспективных сетей беспроводной связи [3].

В настоящее время существует большое количество книг, публикаций, патентов, посвященных теории, построению, алгоритмам кодирования/декодирования, практической аппаратной реализации МППЧ-кодов. Основные понятия и определения, характеризующие МППЧ-коды, можно найти в книге П.В. Трифонова [4]. Также в данной книге находится достаточно обширная библиография, позволяющая сформировать представление о МППЧ-кодах. Как известно [5], конструкции кодов можно разделить на две общие категории: случайные и структурированные. В свою очередь, структурированные коды построены на основе алгебраических и комбинаторных методов [5, 6].

Помимо МППЧ-кодов, в теории помехоустойчивого кодирования широко известны совершенные и циклические коды. Совершенные коды для определенных наборов параметров достигают границы Хэмминга и обеспечивают максимально эффективное использование кодового пространства с точки зрения гарантированного исправления заданного числа ошибок. Циклические коды благодаря своей алгебраической структуре допускают эффективную реализацию процедур кодирования и декодирования, в том числе с использованием регистров сдвига с обратными связями. Вместе с тем МППЧ-коды обладают большей гибкостью при выборе длины и скорости кода, допускают построение кодовых слов большой длины и обеспечивают высокую помехоустойчивость, близкую к пределу Шеннона, что обусловило их широкое применение в современных системах связи.

Для случайных МППЧ-кодов, задаваемых разреженными проверочными матрицами, важной задачей является выбор эффективного метода формирования кодового слова.

Случайное построение МППЧ-кодов заключается в конструировании разреженных случайных матриц проверки четности с использованием компьютерного поиска на основе определенного набора правил проектирования. К графам Таннера таких кодов предъявляется ряд требований: распределение степеней символьных и проверочных узлов, а также контроль циклов малой кратности. Классическая работа [7] показала, что коды, построенные на основе разреженных случайных матриц, обладают возможностью практической реализации декодирования и близостью к пределу Шеннона. Использование таких проверочных матриц позволяет формировать ансамбль кодов, для которого при увеличении длины кода вероятность ошибки стремится к нулю при скорости передачи, не превышающей пропускную способность канала.

Разреженность проверочных матриц **H** обеспечивает возможность эффективного итеративного декодирования с использованием алгоритма распространения доверия (BP) или его приближений [2]. Ряд работ последних лет посвящен разработке эффективных алгоритмов декодирования МППЧ-кодов [8, 9].

К настоящему времени опубликовано множество работ, посвященных конструированию МППЧ-кодов на основе случайных проверочных матриц. В качестве примера можно привести работу Мак-Кея [7]. Мак-Кей и Нил создали обширную библиотеку разреженных случайных проверочных матриц [10], пригодных для практического использования. Кроме того, коды из данной библиотеки могут быть использованы для оценки эффективности и сравнения различных алгоритмов декодирования.

В данной работе излагаются наиболее эффективные и пригодные для практической реализации алгоритмы и методы кодирования случайных МППЧ-кодов. Несмотря на значительное число исследований, посвященных вопросам построения и декодирования МППЧ-кодов, задачи кодирования для случайных проверочных матриц остаются недостаточно проработанными. В общем случае кодирование сводится к решению системы линейных уравнений большой размерности, что может приводить к высокой вычислительной сложности и численным проблемам.

Современные системы связи предъявляют требования высокой скорости обработки при ограниченных вычислительных ресурсах. В связи с этим разработка эффективных алгоритмов кодирования является важной задачей. В работе проводится анализ методов кодирования случайных МППЧ-кодов с точки зрения их применимости на практике. Рассматриваемые методы сопровождаются подробными примерами, что способствует лучшему пониманию их реализации и особенностей применения.

Представлена оценка вычислительной сложности, основанная на подсчете числа арифметических операций, что позволяет выполнить сравнение методов с точки зрения их масштабируемости. Кроме того, приведены результаты имитационного моделирования среднего времени кодирования в среде MatLab, позволяющие оценить практическую производительность алгоритмов при работе со случайными проверочными матрицами различного размера.

Основные методы кодирования двоичных случайных МППЧ-кодов

Коды с малой плотностью проверок на четность представляют собой класс двоичных линейных блоковых кодов, задаваемых разреженной проверочной матрицей $\mathbf{H} \in \{0,1\}^{m,n}$, где m – число проверочных (избыточных) битов, n – длина кодового слова. Скорость МППЧ-кодов определяется соотношением $R = k/n = 1 - m/n$, где k – длина сообщения [2].

Для практической реализации кодирования МППЧ-кодов предложен ряд методов, отличающихся по вычислительной сложности, требованиям к структуре проверочной матрицы и особенностям аппаратной реализации. В основе этих подходов лежат различные способы преобразования проверочной матрицы и решения соответствующих систем линейных уравнений. Далее рассмотрим основные из них.

Кодирование методом Нила [11]. Предположим, для кодирования используется проверочная матрица

$$\mathbf{H} = \left[\begin{array}{ccccc|ccccc} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right], \quad (1)$$

которую можно разделить на две части $\mathbf{H} = [\mathbf{A}|\mathbf{B}]$, где $\mathbf{A}$ и $\mathbf{B}$ имеют размерности $m\times(n-m)$ и $m\times m$ соответственно. Кодовое слово $\mathbf{c}$ также делится на две части $\mathbf{c} = [\mathbf{u}|\mathbf{p}]$, где $\mathbf{u}$ – вектор из $(n-m)$ информационных битов, а $\mathbf{p}$ – вектор из m проверочных битов. Проверочное соотношение согласно [11] имеет вид

$$[\mathbf{A}\,|\,\mathbf{B}]\cdot\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix} = \mathbf{0}, \quad (2)$$

которое можно записать следующим образом:

$$\mathbf{A}\cdot\mathbf{u} + \mathbf{B}\cdot\mathbf{p} = \mathbf{0} \quad (3)$$

или

$$\mathbf{p} = \mathbf{B}^{-1}\cdot\mathbf{A}\cdot\mathbf{u}. \quad (4)$$

Предварительно можно вычислить элементы матрицы $\mathbf{C}$, которые для проверочной матрицы (1) будут равны

$$\mathbf{C} = \mathbf{B}^{-1}\cdot\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}. \quad (5)$$

Тогда проверочные биты можно вычислить в соответствии с выражением

$$\mathbf{p} = \mathbf{C}\cdot\mathbf{u}. \quad (6)$$

Пример 1. Допустим, имеется следующая последовательность информационных битов: $\mathbf{u} = [1\ 1\ 0\ 1\ 0]$. Тогда проверочные биты $\mathbf{p}$, вычисленные согласно уравнению (6), имеют вид $\mathbf{p} = [0\ \ 1\ \ 1\ \ 1\ \ 1]$. Вери-

фикация формирования кодового слова выполняется путем проверки $\mathbf{c}\cdot\mathbf{H}^{\mathrm{T}}=\mathbf{0}$, где

$$\mathbf{c}=[1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 1\ \ 1].$$

Метод кодирования, основанный на LU-разложении квадратной матрицы. Как известно [12], LU-разложение широко используется при решении систем линейных алгебраических уравнений. Методы и примеры разложения матрицы **A** на подматрицы **L** и **U** подробно описаны в [12]. Методы решения систем уравнений на основе данного метода описаны в [13]. Кратко опишем суть метода.

LU-разложением квадратной матрицы **A** называется представление матрицы в виде произведения $\mathbf{A}=\mathbf{L}\cdot\mathbf{U}$, где $\mathbf{L}=(l_{ij})$ – нижняя треугольная матрица, а $\mathbf{U}=(u_{ij})$ – верхняя треугольная матрица с единицами на главной диагонали, например:

$$\mathbf{A}=\mathbf{LU}=\begin{bmatrix}1 & 0 & 0\\ l_{21} & 1 & 0\\ l_{31} & l_{32} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & u_{13}\\ 0 & u_{22} & u_{23}\\ 0 & 0 & u_{33}\end{bmatrix}. \qquad (7)$$

Допустим, необходимо решить систему линейных уравнений (7) с помощью LU-разложения.

$$\mathbf{A}\cdot\mathbf{x}=\mathbf{b} \qquad (8)$$

или

$$\mathbf{L}\cdot\mathbf{U}\cdot\mathbf{x}=\mathbf{b}. \qquad (9)$$

Перепишем (9) в виде $\mathbf{L}\cdot(\mathbf{U}\cdot\mathbf{x})=\mathbf{b}$ и обозначим

$$\mathbf{U}\cdot\mathbf{x}=\mathbf{y}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}. \qquad (10)$$

Первым действием в решении является нахождение **y** из уравнения

$$\mathbf{L}\cdot\mathbf{y}=\mathbf{b}, \qquad (11)$$

затем нахождение **x** из уравнения

$$\mathbf{U}\cdot\mathbf{x}=\mathbf{y}. \qquad (12)$$

Уравнение (11) решается методом прямой подстановки [13]

$$y_i=b_i-\sum_{j=1}^{i-1}l_{ij}\cdot y_j. \qquad (13)$$

Уравнение (12) решается методом обратной подстановки

$$x_i=y_i-\sum_{j=i+1}^{n}u_{ij}\cdot x_j. \qquad (14)$$

Как отмечается в [13], LU-разложение осуществляется только тогда, когда главные миноры матрицы отличны от нуля. На практике это условие выполняется не всегда. Поэтому в прикладных задачах вместо LU-разложений чаще используются так называемые LUP-разложения [14]. Матрица **A** с помощью матрицы перестановок **P** модифицируется так, что матрицы **L** и **U** имеют описанную ранее структуру. Пример решения систем уравнений (9), (10) с учетом матрицы **P** будет описан далее.

Применим LU-разложение для кодирования МППЧ-кода.

Пример 2. Для кодирования кода будем использовать матрицу (5) из примера 1. Главные миноры данной матрицы отличны от нуля. Исходя из этого, для разложения будем использовать стандартную функцию [*L*,*U*] = *lu*(*C*) из библиотеки MatLab.

$$\mathbf{L}=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 1 & 1 & 0 & 0 & 0\\ 1 & 1 & 1 & 0 & 0\\ 1 & 0 & 1 & 1 & 0\\ 0 & 1 & 0 & 0 & 1\end{bmatrix}, \qquad (15)$$

$$\mathbf{U}=\begin{bmatrix}1 & 0 & 0 & 1 & 1\\ 0 & 1 & 1 & 0 & 1\\ 0 & 0 & 1 & 0 & 1\\ 0 & 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}. \qquad (16)$$

Систему уравнений (11), (12) перепишем в виде

$$\mathbf{y}=\mathbf{U}\cdot\mathbf{u}, \qquad (17)$$

где **u** – вектор информационных битов.

$$\mathbf{p}=\mathbf{L}\cdot\mathbf{u}, \qquad (18)$$

где **p** – вектор проверочных битов.

Алгоритм вычисления кодового слова:

1. Формируется вектор сообщения **u**

u=[1 1 0 1 0]

2. Вектор **y** вычисляется путем обратной подстановки

```
K = m;
y(1:K) = 0;
for i = K: -1:1
   sum = 0;
   for j = K:-1:i sum = mod(sum + mod(U(i,j) * u(j),2),2);
end
   y(i) = mod(sum,2);
end
```

y = [0 1 0 1 0]

3. Вектор проверочных битов **p** вычисляется путем прямой подстановки

```
p(1:K) = 0;
for i = 1:K
   sum = 0;
   for j = 1:i sum = mod(sum + mod(L(i,j) * y(j),2),2);
end
   p(i) = sum;
end
```

p = [0 1 1 1 1]

4. c = [1 1 0 1 0 0 1 1 1 1]

5. mod(c*H',2) = [0 0 0 0 0] %{Проверка достоверности кодового слова %}

Метод кодирования, основанный на LUP-разложении квадратной матрицы. Как отмечалось выше, не для всех квадратных матриц существует LU-разложение. В таких случаях не все главные миноры матрицы отличны от нуля. В таком случае строки/столбцы квадратной матрицы переставляются так, чтобы результирующие матрицы имели строго треугольную форму.

LUP-разложением квадратной матрицы **A** называется представление матрицы в виде произведения **PA** = **LU**. В процессе выполнения LUP-разложения генерируется матрица перестановок **P**, которая ис-

пользуется для коррекции решения системы уравнений. Продемонстрируем данный метод для кодирования МППЧ-кода на примере.

Пример 3. Допустим, при построении кода используется проверочная матрица

$$\mathbf{H}=\begin{bmatrix}1&1&1&0&0&1&0&1&1&0\\1&0&0&1&1&1&1&0&0&1\\0&1&1&0&1&0&1&1&1&0\\0&0&1&1&0&1&1&0&1&1\\1&1&0&1&1&0&0&1&0&1\end{bmatrix}. \quad (19)$$

Матрица **C**, вычисленная в соответствии с преобразованием (5), имеет вид

$$\mathbf{C}=\begin{bmatrix}1&0&0&0&0\\0&0&0&0&1\\1&1&0&0&1\\1&0&1&0&1\\0&0&0&1&0\end{bmatrix}. \quad (20)$$

Вычислим LU-разложение для матрицы (20)

$$\mathbf{L}=\begin{bmatrix}1&0&0&0&0\\0&0&0&0&1\\1&1&0&0&0\\1&0&1&0&0\\0&0&0&1&0\end{bmatrix}, \quad (21)$$

$$\mathbf{U}=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}. \quad (22)$$

Как видно из (21), матрица **L** имеет нетреугольную форму. Анализ матрицы (20) показывает, что не все главные миноры отличны от нуля. Миноры 2, 3, 4 равны нулю, следовательно, процесс кодирования невозможен. В таком случае вычислим LUP-разложение. Результат разложения имеет следующий вид:

$$\mathbf{L}=\begin{bmatrix}1&0&0&0&0\\1&1&0&0&0\\1&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}, \quad (23)$$

$$\mathbf{U}=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}, \quad (24)$$

$$\mathbf{P}=\begin{bmatrix}1&0&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&1&0&0&0\end{bmatrix}. \quad (25)$$

Алгоритм вычисления кодового слова:

1. %u = double(rand(K,1)>0.5);
u = [1 0 0 0 1]
2. Вектор **y** вычисляем путем обратной подстановки
y = [1 1 1 0 1]
3. Вектор проверочных битов **p** (промежуточный результат) вычисляем путем прямой подстановки
p = [1 0 0 0 1]
4. Коррекция проверочных битов
p = double(p)
p = mod(p***P**,2)
p = [1 1 0 0 0]
5. c = [1 0 0 0 1 1 1 0 0 0] %{Вычисленное кодовое слово%}
6. mod(c*H',2) = [0 0 0 0 0] %{Проверка достоверности кодового слова %}

Алгоритмы кодирования, построенные на методах LU и LUP, являются последовательными. Алгоритм кодирования, основанный на LU-разложении, выполняется за 2 шага. Алгоритм кодирования, основанный на LUP-разложении, выполняется за 3 шага.

Данное свойство дает возможность использовать технику конвейерной обработки данных [15] и позволяет значительно увеличить производительность кодера по сравнению, например, с кодером Нила.

Кроме того, аппаратная реализация данных методов очень простая, не требующая параллельных матричных вычислений. Примеры блоков, реализующих вычисления прямой и обратной подстановки, можно найти в [16].

Метод кодирования, основанный на приведении проверочной матрицы к ступенчатой форме. Для приведения проверочной матрицы **H** к систематической ступенчатой $\mathbf{H}_{\mathbf{sys}}$ используется метод исключения Гаусса–Жордана [12]. Систематическая проверочная матрица имеет вид

$$\mathbf{H}_{\mathbf{sys}}=[\mathbf{I}_m\mathbf{P}], \quad (26)$$

где $\mathbf{I}_m$ – единичная матрица, имеющая размерность $m \times m$. Матрица **P** имеет размерность $m \times (n–m)$.

Для нахождения $\mathbf{H}_{\mathbf{sys}}$ из **H** будем использовать стандартную функцию R = rref(A) из библиотеки MatLab. Функция rref(H) приводит матрицу **H** к ступенчатой форме $\mathbf{H}_{\mathbf{sys}}$ с использованием исключения Гаусса–Жордана.

Пример 4. Допустим, для кодирования используется проверочная матрица:

$$\mathbf{H}=\begin{bmatrix}1&1&0&1&1&0&0&1&0&0\\0&1&1&0&1&1&1&0&0&0\\0&0&0&1&0&0&0&1&1&1\\1&1&0&0&0&1&1&0&1&0\\0&0&1&0&0&1&0&1&0&1\end{bmatrix}, \quad (27)$$

$$\mathbf{H}_{\mathbf{sys}}=\begin{bmatrix}1&0&0&0&0&0&1&1&1&0\\0&1&0&0&0&1&0&1&0&0\\0&0&1&0&0&1&0&1&0&1\\0&0&0&1&0&0&0&1&1&1\\0&0&0&0&1&1&1&0&0&1\end{bmatrix}. \quad (28)$$

Путем перестановки подматриц $\mathbf{I}_m\mathbf{P}$ в (28) получаем стандартную форму проверочной матрицы в ступенчатой форме

$$\mathbf{H_{std}} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (29)$$

Используя правила построения порождающей матрицы [17], из стандартной проверочной получаем стандартную порождающую матрицу в ступенчатой форме

$$\mathbf{G} = [\mathbf{I}_m\mathbf{P}^{\mathrm{T}}] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \quad (30)$$

Достоверность вычисления порождающей матрицы проверяем с помощью соотношения $[\mathbf{G}\cdot\mathbf{H}_{\mathbf{std}}^{\mathrm{T}}] = \mathbf{0}$.

Алгоритм вычисления кодового слова:

1. %u = round(rand(1,size(G,1)));
u = [1 0 1 0 1]
2. Вычисление кодового слова:
c = mod(u*G,2);
%{c = [1 0 1 0 1 1 0 1 0 0]%}
3. mod(c*H_std',2) = [0 0 0 0 0] %{Проверка достоверности кодового слова %}

Методы кодирования на основе эквивалентной проверочной матрицы в нижней треугольной форме

Метод кодирования на основе нижней триангуляции проверочной матрицы. В работе [18] предложена методика кодирования, основанная на приведении проверочной матрицы к эквивалентной проверочной матрице в нижней треугольной форме.

Средствами метода исключения Гаусса [12] проверочная матрица **H** преобразуется в эквивалентную нижнетреугольную матрицу $\mathbf{H}_\mathbf{t}$. Проверочные биты p_l, $l = 1, \ldots, m$ вычисляются в соответствии с формулой

$$p_l = \sum_{j=1}^{m-n} \mathbf{H}_{\mathbf{t}l,j} \cdot s_j + \sum_{j=1}^{l-1} \mathbf{H}_{\mathbf{t}l,j+n-m} \cdot p_j, \quad (31)$$

где s_j – информационные биты.

Пример 5. Допустим, для кодирования используется проверочная матрица

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (32)$$

Применяя исключение Гаусса для проверочной матрицы, получаем нижнюю треугольную матрицу:

$$\mathbf{H_t} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}. \quad (33)$$

Как видно из (33), матрицу $\mathbf{H_t}$ можно разделить на две части $\mathbf{H_t} = [\mathbf{A} \mid \mathbf{L}]$, где **A** и **L** имеют размерности $m\times(n-m)$ и $m\times m$ соответственно. Матрица **A** является прямоугольной (в примере 5 квадратной, поскольку скорость кода равна 1/2). Матрица **L** является треугольной. Кодовое слово **c** делится также на две части $\mathbf{c} = [\mathbf{s}\ \mathbf{p}]$. С учетом данного обстоятельства из анализа выражения (29) видно, что в вычислении кодового слова можно выделить два независимых процесса. Во время первого процесса коэффициенты текущей строки l матрицы **A** умножаются на информационные биты **s** и суммируются. Во время второго процесса вычисленные предыдущие проверочные биты умножаются на коэффициенты матрицы **L** и суммируются. Данный процесс в [18] назван методом обратной подстановки. По окончании данного процесса вычисляется текущий проверочный бит p_l.

Проверочные биты вычисляются последовательно, при этом оба процесса на каждом шаге могут выполняться параллельно. При параллельной реализации вычислений матрицы **A** и **L** должны храниться в памяти раздельно. Данный метод даёт возможность значительно повысить скорость кодирования. Отметим, что при этом требования к объему памяти выше, чем у кодера Нила или кодера, основанного на LU и LUP-разложениях.

Алгоритм вычисления кодового слова, реализующий выражение (29):

```
1. %s=double(rand(k,1)>0.5)
s =[1 0 1 1 1]'
for j = 1:m
   p(j) = 0;
end
2. Вычисление кодового слова:
for l = 1:m
   sum1 = 0;
   sum1 = mod(HT(l,1:k)*s,2);  % First process
   sum2 = 0;
   for j = 1:l sum2 = mod(sum2 +
mod(HT(l,j+k)*p(j),2),2 % Second process
   end
   p(l) = mod(sum1 + sum2,2);
end
% p = [1 1 1 1 1]
3. c = [s' p]
% c = [1 0 1 1 1 1 1 1 1 1]
mod(H*c',2) = [0 0 0 0 0] %{Проверка достоверности
кодового слова %}
```

Метод кодирования на основе приближенной нижней триангуляции проверочной матрицы. В работе [18] предложена методика кодирования, осно-

ванная на приближенной эквивалентной проверочной матрице в нижней треугольной форме.

Следует отметить, что данный подход применяется при проектировании современных кодеров квазициклических МППЧ-кодов стандарта 5G NR [22, 23]. Структура проверочных матриц кодов 5G NR позволяет использовать нижнетреугольное представление для эффективного последовательного вычисления проверочных битов. Реализации кодеров, основанные на данном принципе, рассмотрены в ряде современных работ [22, 23].

Суть метода заключается в следующем:

1) генерируется или выбирается разреженная проверочная матрица **H**, имеющая размерность $m \times n$. Матрица должна быть невырожденной;

2) матрица **H** делится на блоки (подматрицы) с соответствующими размерностями

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{T} \\ \mathbf{C} & \mathbf{D} & \mathbf{E} \end{bmatrix}. \quad (34)$$

Подматрицы имеют следующие размерности:

A – $(m–g)\times(n–m)$; **C** – $g\times(n–m)$;

B – $(m–g)\times(g)$; **D** – $g\times g$;

T – $(m–g)\times(m-g)$; **E** – $g\times(m–g)$.

Значение g должно быть по возможности минимальным [18].

Проверочное соотношение записывается в форме

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{T} \\ \mathbf{C} & \mathbf{D} & \mathbf{E} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{u} \\ \mathbf{p}_1 \\ \mathbf{p}_2 \end{bmatrix} = \mathbf{0}, \quad (35)$$

где **u** – вектор информационных битов, $\mathbf{p}_1$ и $\mathbf{p}_2$ – векторы проверочных битов.

Проверочное соотношение можно записать в иной форме

$$\mathbf{A}\cdot\mathbf{u} + \mathbf{B}\cdot\mathbf{p}_1 + \mathbf{T}\cdot\mathbf{p}_2 = \mathbf{0}, \quad (36)$$

$$\mathbf{C}\cdot\mathbf{u} + \mathbf{D}\cdot\mathbf{p}_1 + \mathbf{E}\cdot\mathbf{p}_2 = \mathbf{0}. \quad (37)$$

Система уравнений (36), (37) решается с помощью блочного исключения Гаусса–Шура [19], позволяющего исключить $\mathbf{p}_2$ в уравнении (37). Данная задача решается путем «обнуления» подматрицы **E** в уравнении (37). Описание процедуры «обнуления» подматрицы в системе уравнений и дополнительные сведения приведены в [20]. Для вычисления дополнения Шура к подматрице **E** выполняется умножение на матрицу

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{E}\cdot\mathbf{T}^{-1} & \mathbf{I} \end{bmatrix}, \quad (38)$$

$$\begin{bmatrix} \mathbf{T} \\ \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{E}\cdot\mathbf{T}^{-1} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{T} \\ \mathbf{0} \end{bmatrix}, \quad (39)$$

где

$$\mathbf{T}\cdot(-\mathbf{E}\cdot\mathbf{T}^{-1}) + \mathbf{E}\cdot\mathbf{I} = 0. \quad (40)$$

Соответственно, дополнения Шура к подматрицам **C** и **D** также вычисляются путем умножения данных подматриц на матрицу (36). Окончательные выражения, полученные авторами, приводятся

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{T} \\ -\mathbf{E}\cdot\mathbf{T}^{-1}\cdot\mathbf{A} + \mathbf{C} & -\mathbf{E}\cdot\mathbf{T}^{-1}\cdot\mathbf{B} + \mathbf{D} & 0 \end{bmatrix}. \quad (41)$$

Совмещая систему уравнений (36), (37) и выражение (41), окончательно получим систему уравнений для вычисления проверочных битов $\mathbf{p}_1$ и $\mathbf{p}_2$:

$$\mathbf{A}\cdot\mathbf{u} + \mathbf{B}\cdot\mathbf{p}_1 + \mathbf{T}\cdot\mathbf{p}_2 = 0, \quad (42)$$

$$\left(-\mathbf{E}\cdot\mathbf{T}^{-1}\cdot\mathbf{A} + \mathbf{C}\right)\cdot\mathbf{u} + \left(-\mathbf{E}\cdot\mathbf{T}^{-1}\cdot\mathbf{B} + \mathbf{D}\right)\cdot\mathbf{p}_1 = 0. \quad (43)$$

Из уравнения (41) получим выражение для вычисления проверочных битов $\mathbf{p}_1$.

$$\mathbf{p}_1 = \left(-\mathbf{E}\cdot\mathbf{T}^{-1}\cdot\mathbf{B} + \mathbf{D}\right)^{-1}\left(-\mathbf{E}\cdot\mathbf{T}^{-1}\cdot\mathbf{A} + \mathbf{C}\right)\cdot\mathbf{u}. \quad (44)$$

Подставляя вычисленный вектор $\mathbf{p}_1$ в уравнение (40), получаем вектор проверочных битов $\mathbf{p}_2$:

$$\mathbf{p}_2 = \mathbf{T}^{-1}\left(\mathbf{A}\cdot\mathbf{u} + \mathbf{B}\cdot\mathbf{p}_1\right). \quad (45)$$

Рассмотрим данную методику кодирования на примере.

Пример 6. В данном примере используем следующую проверочную матрицу.

Данная матрица невырожденная и имеет ранг $r = 10$.

$$\mathbf{H} = \begin{bmatrix} 0&1&1&0&1&0&0&0&0&1&0&0&0&1&1&0&0&1&0&0 \\ 1&0&0&0&0&0&1&0&0&0&0&1&0&0&1&0&0&1&1&0 \\ 0&0&0&0&0&0&0&0&0&0&1&1&0&1&1&1&0&0&0&1 \\ 0&1&0&0&1&1&0&1&0&1&0&0&1&0&0&1&0&0&0&0 \\ 0&0&0&0&0&1&0&1&1&0&0&1&1&0&0&0&1&0&0&0 \\ 1&0&0&1&0&0&0&0&1&1&0&0&1&0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&1&0&0&1 \\ 0&0&0&1&1&1&1&0&1&0&1&0&0&0&0&0&0&1&0&0 \\ 0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0 \\ 1&1&1&0&0&0&0&1&0&0&1&0&0&0&0&0&1&0&0&0 \end{bmatrix}. \quad (46)$$

Значение g выбираем равным 2. К первым восьми строкам матрицы применен метод исключения Гаусса для нижней треугольной матрицы. Последние две строки матрицы остаются без изменения. Полученная эквивалентная проверочная матрица **ATM** с нижней треугольной формой имеет вид

$$\mathbf{ATM} = \left[\begin{array}{cccccccccc|cc|cccccccc}
0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
- & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{array}\right]. \quad (47)$$

Алгоритм вычисления кодового слова:

1. invT = mod((inv(T)),2) =
[1 0 0 0 0 0 0 0;
0 1 0 0 0 0 0 0;
1 1 1 0 0 0 0 0;
1 0 0 1 0 0 0 0;
1 0 0 0 1 0 0 0;
0 0 0 0 0 1 0 0;
1 0 0 1 1 0 1 0;
0 0 0 1 1 0 0 1]
2. ET1 = mod(-(E*invT),2) = [1 0 0 1 1 0 1 0; 1 0 0 0 1 0 0 0]
3. cD = mod((ET1*B + D),2) = [0 1; 1 1] %{ Матрица невырожденная, что соответствует условию метода %}
4. cC = mod((ET1*A + C),2) =
[1 0 0 1 0 1 1 0 1; 1 0 0 1 0 1 0 1 1 1]
5. tmp = mod(invcD * cC,2) =
[0 0 0 0 1 1 1 0 1 0; 1 0 0 1 1 0 1 1 0 1]
6. %u = round(rand(1,m))
u = [0 1 0 1 1 1 0 0 0 1]
7. p1 = mod(tmp*u',2) = [0 1]
8. p2 = mod(invT*(A*u' + B*p1),2) = [1 1 1 1 1 1 1 0]
9. c = [u p1' p2'] = [0 1 0 1 1 1 0 0 0 1 0 1 1 1 1 1 1 1 1 0]
10. mod(c*H',2) = [0 0 0 0 0 0 0 0 0 0] %{Проверка достоверности кодового слова %}

Оценка вычислительной сложности методов кодирования

В этом разделе проводится оценка вычислительной сложности (ВС) рассмотренных методов кодирования. Оценка ВС по подготовке матриц, используемых в кодере, не рассматривается.

Под ВС будет пониматься число булевых арифметических операций, участвующих в процессе кодирования (сложений, умножений), которые рассматриваются как равнозначные. Основное вычисление, которое выполняет арифметическое устройство систематического кодера – умножение порождающей матрицы

$$\mathbf{G} = \begin{bmatrix} g_{11} & \cdots & g_{1k} \\ \vdots & \ddots & \vdots \\ g_{m1} & \cdots & g_{mk} \end{bmatrix} \quad (48)$$

на вектор информации

$$\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \ldots & u_k \end{bmatrix}^{\mathrm{T}}. \quad (49)$$

В результате произведения формируется вектор проверочных (избыточных) битов:

$$\mathbf{p} = \mathbf{G}\mathbf{u}, \quad (50)$$

$$p_i = \sum_{j}^{m} g_{ij} u_i. \quad (51)$$

Оценку ВС будем рассматривать на примере LDPC-кода с кодовой скоростью $R = 1/2$. В этом случае число информационных битов k равно числу проверочных битов m, а $\mathbf{G}$ является квадратной матрицей. Для вычисления одного проверочного бита (51) необходимо m операций умножения и $m-1$ операций сложения, или в общем $2m-1$ операций. Для вычисления всех проверочных битов требуется $m(2m-1) = 2m^2 - m$ операций. Обозначим для примера ВС умножения матрицы на вектор как $O(2m^2 - m)$. Основываясь на данном подходе, оценим ВС всех рассмотренных методов кодирования.

1. Метод Нила для вычисления избыточных битов согласно (6)

$$O(2m^2 - m). \quad (52)$$

2. Метод кодирования, основанный на LU-разложении квадратной матрицы:

$$O(m_1, m_2) = O\left(2\left(\sum_{j=1}^{m} j + \sum_{j=1}^{m} (j-1)\right)\right), \quad (53)$$

где первая сумма определяет число умножений, вторая – число сложений.

3. Метод кодирования, основанный на LUP-разложении квадратной матрицы:

$$O(m_1, m_2) = O\left(2\left(\sum_{j=1}^{m} j + \sum_{j=1}^{m} (j-1)\right) + 2m^2 - m\right), \quad (54)$$

где третье слагаемое определяет дополнительные вычисления для коррекции избыточных битов путем перемножения на перестановочную матрицу $\mathbf{P}$.

4. Метод кодирования, основанный на приведении проверочной матрицы к ступенчатой форме:

$$O(2m + (2m^2 - m)). \quad (55)$$

5. Метод кодирования на основе нижней триангуляции проверочной матрицы

$$O\left(2m^2 - m + \left(\sum_{j=1}^{m} j + \sum_{j=1}^{m} (j-1)\right)\right), \quad (56)$$

где первое слагаемое определяет вычисления на матрице **A**, второе – на матрице **L**.

6. Метод кодирования на основе приближенной нижней триангуляции проверочной матрицы:

1) вычисление выражения (44)

$$O_1\left((2m-1)g\right). \quad (57)$$

2) вычисление выражения (45) состоит из операций, имеющих следующую сложность:

$$\mathbf{A}\cdot\mathbf{u}: O_2\left((m-g)(2m-1)\right), \quad (58)$$

$$\mathbf{B}\cdot\mathbf{p}_1: O_3\left(g\left(2(m-g)-1\right)\right). \quad (59)$$

Сумма произведений

$$\mathbf{A}\cdot\mathbf{u}+\mathbf{B}\cdot\mathbf{p}_1: O_4\left(m-g\right). \quad (60)$$

Последняя операция матричного умножения в (45)

$$\mathbf{T}^{-1}\left(\mathbf{A}\cdot\mathbf{u}+\mathbf{B}\cdot\mathbf{p}_1\right): O_5(g(2(m-g)^2-m-g)). \quad (61)$$

Таким образом, алгоритм имеет следующую вычислительную сложность:

$$O\left(O_1+O_2+O_3+O_4+O_5\right). \quad (62)$$

Для сопоставления вычислительной сложности рассмотренных методов выполнено моделирование зависимости числа арифметических операций от размера проверочной матрицы m при фиксированной скорости (рис. 1).

Как видно из рис. 1, наименьшую сложность демонстрируют методы Нила, LU и RREF (приведение проверочной матрицы к ступенчатой форме), имеющие близкие зависимости с квадратичным ростом. Методы Triang (нижняя триангуляция проверочной матрицы), ALT (приближенная нижняя триангуляция проверочной матрицы) и LUP демонстрируют более быстрый рост вычислительной сложности.

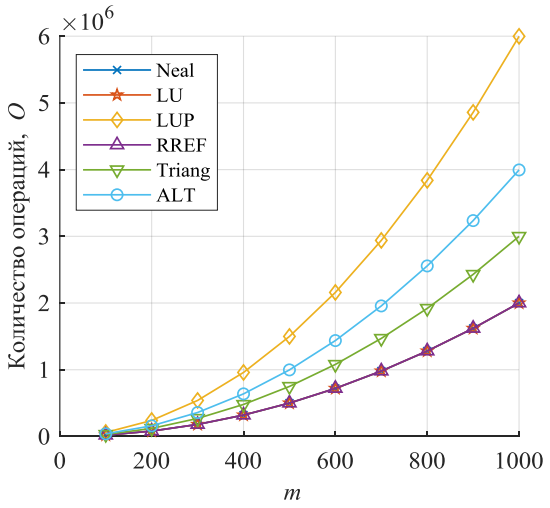


Рис. 1. Зависимость числа арифметических операций от размера проверочной матрицы m

Оценка времени вычислений

В дополнение к проведенной аналитической оценке ВС методов кодирования случайных МППЧ-кодов рассмотрим их временную сложность при работе в единой программной среде.

Для оценки среднего времени выполнения операций кодирования T_{enc} описанных методов была разработана имитационная модель в среде MatLab. Во всех сценариях моделирования рассматривались проверочные матрицы МППЧ-кодов с фиксированной скоростью кодирования R = 1/2. Каждая строка матрицы имела фиксированный вес, равный трем, а изменение размеров матрицы осуществлялось за счет управления параметром m (при $n = 2m$). Результаты усреднялись по 1000 случайным реализациям матриц, удовлетворяющих условиям реализуемости рассматриваемых методов. Для каждой матрицы имитировалось кодирование ста случайных сообщений. Результаты моделирования приведены на рис. 2.

Как видно из рис. 2, наибольшее время кодирования демонстрируют методы, основанные на LU- и LUP-разложениях, а также метод Triang. Это объясняется тем, что кодирование в реализациях алгоритмов этих методов основано на циклах, в то время как методы Нила, RREF и ALT реализованы с использованием матричных умножений. Среда MatLab оптимизирована для выполнения матричных умножений, что позволяет обеспечить низкий рост времени вычислений с увеличением размеров проверочных матриц. Вместе с тем операции матричного умножения требуют значительных аппаратных ресурсов при реализации в программируемых логических интегральных схемах (ПЛИС).

По этой причине время вычислений различными методами при аппаратной реализации в ПЛИС будет отличаться от проведенной оценки.

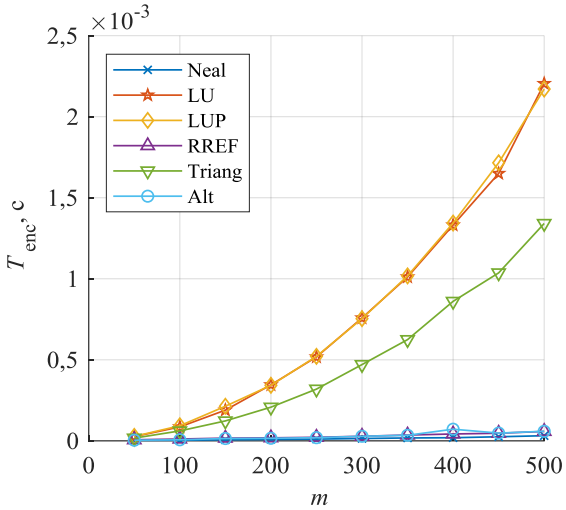


Рис. 2. Зависимость среднего времени кодирования от размера проверочной матрицы m

Заключение

В работе проведен анализ методов кодирования случайных МППЧ-кодов, основанных на различных подходах к преобразованию проверочной матрицы и решению систем линейных уравнений. Показано, что задачи эффективного кодирования для случайных разреженных матриц остаются актуальными, особенно в контексте практической реализации.

Рассмотренные методы, включая подход Нила, алгоритмы на основе LU- и LUP-разложений, приве-

дение матрицы к ступенчатой форме, а также методы нижней и приближенной триангуляции, демонстрируют различия по вычислительной сложности, требованиям к структуре матрицы и возможностям аппаратной реализации. Установлено, что методы, основанные на треугольных преобразованиях, позволяют реализовать последовательные алгоритмы кодирования с возможностью конвейерной обработки, что важно для высокоскоростных систем связи.

Показано, что выбор конкретного метода кодирования определяется компромиссом между сложностью реализации, требованиями к памяти и структурой проверочной матрицы. Представленные в работе примеры иллюстрируют практическую применимость рассмотренных алгоритмов и позволяют лучше понять особенности их реализации.

Выполнен анализ вычислительной сложности рассмотренных методов кодирования, основанный на оценке числа арифметических операций, а также проведено моделирование времени кодирования в среде MatLab. Полученные результаты демонстрируют различия между методами с точки зрения их практической реализуемости и эффективности при увеличении размерности проверочной матрицы.

Полученные результаты могут быть использованы при разработке эффективных кодеров для систем связи следующих поколений, включая 5G- и перспективные 6G-системы.

Литература

1. 6G communication systems: concept, trends, physical level technologies / D.A. Pokamestov, Y.V. Kryukov, R.R. Abenov, E.V. Rogozhnikov, A.A. Brovkin, A.S. Shinkevich, G.N. Shalin // Radio Engineering and Electronics. – 2024. – Vol. 69, № 1. – P. 3–33.

2. Gallager R. Low-density parity-check codes // IRE Transactions on Information Theory. – 1962. – Vol. 8, No. 1. – P. 21–28.

3. Channel Coding Toward 6G: Technical Overview and Outlook / M. Rowshan, M. Qiu, Y. Xie, X. Gu, J. Yuan // IEEE Open Journal of the Communications Society. – 2024. – Vol. 5. – P. 2585–2685.

4. Трифонов П.В. Основы помехоустойчивого кодирования: учеб. пособие. – СПб.: Университет ИТМО, 2022. – 231 с.

5. Lin S. Structured Low-Density Parity-Check Codes: Algebraic Constructions // IEEE 802.3an Task Force Meeting. – IEEE: Portland, OR, 13–15, 2004. – P. 1–20.

6. Donovan D. High-rate LDPC codes from partially balanced incomplete block designs / D. Donovan, A. Price, A. Rao, E. Üsküplü, E.Ş. Yazıcı // Journal of Algebraic Combinatorics. – 2022. – Vol. 55. – P. 259–275.

7. MacKay D.J.C. Good Error-Correcting Codes Based on Very Sparse Matrices // IEEE Transactions on Information Theory. – 1999. – Vol. 45, No. 2. – P. 399–431.

8. Кравченко А.Н. Снижение сложности декодирования низкоплотностного кода // Цифровая обработка сигналов. – 2010. – № 2. – С. 35–41.

9. Subhi M.I. Enhancing Data Communication Performance: A Comprehensive Review and Evaluation of LDPC Decoder Architectures / M.I. Subhi, Q. Al-Doori, O. Alani // Ingénierie des Systèmes d'Information. – 2023. – Vol. 28, No. 5. – P. 1113–1125.

10. MacKay D.J.C. Encyclopedia of Sparse Graph Codes [Электронный ресурс]. – URL: http://www.inference.org.uk/mackay/codes/data.html (дата обращения: 19.08.2026).

11. Neal R.M. Sparse Matrix Methods and Probabilistic Inference Algorithms // IMA Program on Codes, Systems and Graphical Models. – 1999. – P. 1–16.

12. Strang G. Introduction to Linear Algebra. – 6th ed. – Wellesley, MA, USA: Wellesley-Cambridge Press, 2023. – 537 p.

13. Вержбицкий В.М. Вычислительная линейная алгебра. – М.: Высшая школа, 2009. – 352 с.

14. LUP-разложение квадратных матриц [Электронный ресурс]. – URL: https://study.sfu-kras.ru/DATA/docs/ProgramTheory/recurs/lup_expn.htm (дата обращения: 19.08.2026).

15. Harris S.L. Digital design and computer architecture, RISC-V edition / S.L. Harris, D. Harris. – Elsevier: Morgan Kaufmann, 2021. – 592 p.

16. A Flexible Hardware Encoder for Low-Density Parity-Check Codes / D.-U. Lee, W. Luk, C. Wang, C. Jones, M. Smith, J. Villasenor // Proceedings of the 12th Annual IEEE Symposium on Field-Programmable Custom Computing Machines. – 2004. – P. 101–111.

17. Research and implementation of high-efficiency and low-complexity LDPC-coding algorithm / X. Liao, J. Guo, Z. Luo, Y. Xu, Y. Chu // Electronics. – 2023. – Vol. 12, No. 17. – P. 1–20.

18. Efficient Encoding of Low-Density Parity-Check Codes / T.J. Richardson, R.L. Urbanke // IEEE Transactions on Information Theory. – 2001. – Vol. 47, No. 2. – P. 638–656.

19. Уоткинс Д.С. Основы матричных вычислений / пер. с англ. В.Е. Кондрашова, С.Б. Королева. – М.: БИНОМ. Лаборатория знаний, 2006. – 664 с.

20. Partitioned Matrices and the Schur Complement [Электронный ресурс]. – URL: https://quickfem.com/wp-content/uploads/IFEM.AppP_.pdf (дата обращения: 19.08.2026).

21. Golub G.H. Matrix Computations / G.H. Golub, C.F. Van Loan. – 4th ed. – Baltimore: Johns Hopkins University Press, 2013. – 756 p.

22. Tian Y. Low-Latency QC-LDPC Encoder Design for 5G NR / Y. Tian, Y. Bai, D. Liu // Sensors. – 2021. – Vol. 21, No. 18. – P. 6266.

23. Petrović V.L. Flexible 5G New Radio LDPC Encoder Optimized for High Hardware Usage Efficiency / V.L. Petrović, D.M. El Mezeni, A. Radošević // Electronics. – 2021. – Vol. 10, No. 9. – P. 1106.

Кравченко Александр Николаевич
Канд. техн. наук, системный архитектор Ubiso GmbH,
Hermann-Schwer-Str. 3, VS-Villingen, Germany, 78048
Тел.: +49 (0) 7-721-944-76-77
Эл. почта: alexander.kravtchenko@ubiso.com

Покаместов Дмитрий Алексеевич
Канд. техн. наук, доцент каф. телекоммуникаций и основ радиотехники Томского государственного ун-та систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0002-9405-4573
Тел.: +7-952-809-43-75
Эл. почта: dmaltomsk@mail.ru

Поступила в редакцию: 06.04.2026.
Принята к публикации: 24.08.2026.

Kravtchenko A.N., Pokamestov D.A.

Encoding Methods for Random Low-Density Parity-Check Codes

***Relevance*.** The development and investigation of encoding methods for random low-density parity-check (LDPC) codes is an important task in the design of high-speed communication systems, since different codeword generation algorithms differ significantly in computational cost and implementation requirements. ***Purpose of the study*.** Comparative analysis of encoding methods for LDPC-codes defined by random sparse parity-check matrices, with an assessment of their computational complexity and practical performance. ***Methods*.** The study employs methods of linear algebra, analytical estimation of computational complexity, simulation modeling, and the Monte Carlo method. The encoding approaches under study include Neal's method, algorithms based on LU and LUP-decompositions, transformation of the parity-check matrix to row-echelon form, as well as lower and approximate lower triangulation methods. ***Scientific novelty*.** A comparative study of encoding methods for random LDPC codes was carried out, including a joint assessment of their analytical computational complexity and software implementation time under identical initial conditions. ***Results*.** Differences in the growth patterns of computational costs among the considered methods as the size of the parity-check matrix increases were identified. It was shown that Neal's method, LU decomposition, and the method based on transformation to row-echelon form require fewer arithmetic operations, whereas the results of software simulations depend significantly on the implementation approach and the optimization capabilities for matrix operations in MatLab. ***Practical significance*.** The obtained results can be used to select encoding algorithms for random LDPC-codes and developing software and hardware encoders for high-speed communication systems.

Keywords: LDPC, low-density parity-check codes, random matrices, error-correcting coding, computational complexity.
DOI: 10.21293/1818-0442-2026-29-2-64-73

References

1. Pokamestov D.A., Kryukov Y.V., Abenov R.R., Rogozhnikov E.V., Brovkin A.A., Shinkevich A.S., Shalin G.N. 6G communication systems: concept, trends, physical level technologies. *Radio Engineering and Electronics*, 2024, vol. 69, no. 1, pp. 3–33.

2. Gallager R.G. Low-density parity-check codes. *IRE Transactions on Information Theory*, 1962, vol. 8, no. 1, pp. 21–28.

3. Rowshan M., Qiu M., Xie Y., Gu X., Yuan J. Channel Coding Toward 6G: Technical Overview and Outlook. *IEEE Open Journal of the Communications Society*, 2024, vol. 5, pp. 2585–2685.

4. Trifonov P.V. *Osnovy pomekhoustoichivogo kodirovaniya: uchebnoe posobie* [Fundamentals of error-correcting coding: textbook]. Saint-Petersburg, ITMO University Publ., 2022, 231 p. (in Russ.).

5. Lin S. Structured Low-Density Parity-Check Codes: Algebraic Constructions. IEEE 802.3an Task Force Meeting, IEEE, Portland, OR, July 13–15, 2004.

6. Donovan D., Price A., Rao A., Üsküplü E., Yazıcı E.Ş. High-rate LDPC codes from partially balanced incomplete block designs. *Journal of Algebraic Combinatorics*, 2022, vol. 55, pp. 259–275.

7. MacKay D.J.C. Good error-correcting codes based on very sparse matrices. *IEEE Transactions on Information Theory*, 1999, vol. 45, no. 2, pp. 399–431.

8. Kravchenko A.N. [Reduced-complexity decoding LDPC code]. *Digital Signal Processing*, 2010, no. 2, pp. 35–41 (in Russ.).

9. Subhi M.I., Al-Doori Q., Alani O. Enhancing data communication performance: a comprehensive review and evaluation of LDPC decoder architectures. *Ingénierie des Systèmes d'Information*, 2023, vol. 28, no. 5, pp. 1113–1125.

10. MacKay D.J.C. Encyclopedia of Sparse Graph Codes. Available at: http://www.inference.org.uk/mackay/codes/data.html (accessed: 19 August 2026).

11. Neal R.M. Sparse matrix methods and probabilistic inference algorithms. *IMA Program on Codes, Systems and Graphical Models*, 1999, pp. 1–16.

12. Strang G. Introduction to Linear Algebra. 6th ed. Wellesley, MA, USA: Wellesley-Cambridge Press, 2023, 537 p.

13. Verzhbitskii V.M. *Vychislitelnaya lineinaya algebra* [Computational linear algebra]. Moscow, Vysshaya shkola Publ., 2009, 352 p. (in Russ.).

14. *LUP-razlozhenie kvadratnykh matrits* [LUP decomposition of square matrices]. Available at: https://study.sfukras.ru/DATA/docs/ProgramTheory/recurs/lup_expn.htm (accessed: 19 August 2026).

15. Harris S.L., Harris D. Digital design and computer architecture: RISC-V edition. Elsevier, Morgan Kaufmann, 2021, 592 p.

16. Lee D.-U., Luk W., Wang C., Jones C., Smith M., Villasenor J. A flexible hardware encoder for low-density parity-check codes. *Proc. 12th Annual IEEE Symposium on Field-Programmable Custom Computing Machines*, 2004, pp. 101–111.

17. Liao X., Guo J., Luo Z., Xu Y., Chu Y. Research and implementation of high-efficiency and low-complexity LDPC coding algorithm. *Electronics*, 2023, vol. 12, no. 17, pp. 1–20.

18. Richardson T.J., Urbanke R.L. Efficient encoding of low-density parity-check codes. *IEEE Transactions on Information Theory*, 2001, vol. 47, no. 2, pp. 638–656.

19. Watkins D.S. *Osnovy matrichnykh vychislenii* [Fundamentals of matrix computations]. Moscow, BINOM. Laboratoriya znanii Publ., 2006, 664 p. (in Russ.).

20. Partitioned matrices and the Schur complement. Available at: https://quickfem.com/wp-content/uploads/IFEM.AppP_.pdf (accessed: 19 August 2026).

21. Golub G.H., Van Loan C.F. Matrix computations. 4th ed. Baltimore, Johns Hopkins University Press, 2013, 756 p.

22. Tian Y., Bai Y., Liu D. Low-latency QC-LDPC encoder design for 5G NR. *Sensors*, 2021, vol. 21, no. 18, pp. 6266.

23. Petrović V.L., Mezeni D.M., Radošević A. Flexible 5G New Radio LDPC encoder optimized for high hardware usage efficiency. *Electronics*, 2021, vol. 10, no. 9, pp. 1106.

Alexander N. Kravchenko
PhD, System Architecture, Ubiso GmbH (retired)
3, Hermann-Schwer-Str., VS-Villingen, Germany, 78048
Phone: +49 (0) 7-721-944-76-77
Email: alexander.kravtchenko@ubiso.com

Dmitriy A. Pokamestov
Candidate of Engineering Sciences, Associate Professor,
Department of Telecommunications and Fundamentals
of Radio Engineering, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0002-9405-4573
Phone: +7-952-809-43-75
Email: dmaltomsk@mail.ru

Received: 06.04.2026.
Accepted: 24.08.2026.

УДК 621.396.96
Шифр специальности ВАК: 2.2.16

И.В. Безрукова, Э.В. Семенов

Моделирование и особенности нелинейного отклика малоразмерных электронных устройств при интермодуляционном методе сверхширокополосного зондирования

Актуальность. Применение сверхширокополосных сигналов в системах нелинейного зондирования ограничено тем, что полезный нелинейный отклик типичных малоразмерных электронных объектов на порядок слабее собственных искажений приемных трактов и оцифровщиков. Для преодоления этого барьера требуется увеличение относительного уровня нелинейного отклика. Перспективным решением выступает интермодуляционный метод зондирования, использующий двухкомпонентный сигнал, где низкочастотная часть влияет на степень нелинейности состояния объекта, а высокочастотная – на считывание этого состояния, что позволяет раздельно оптимизировать обе составляющие и добиться существенного роста отклика по сравнению с противополярными ступенчатыми импульсами. ***Цель исследования.*** Параметризация поведенческой модели нелинейного рассеяния применительно к интермодуляционному методу зондирования для прогнозирования нелинейного отклика малоразмерных электронных устройств в расширенном диапазоне напряженностей поля. ***Методы.*** В основе работы лежит поведенческая модель нелинейного рассеяния, описывающая преобразование формы тестового сигнала в нелинейный отклик через передаточную характеристику, вычисляемую в частотной области по экспериментально зарегистрированным парам «зондирующее поле – нелинейный отклик». Параметризация для интермодуляционного метода выполнена на основе экспериментальных данных: для двухкомпонентного сигнала рассчитана передаточная характеристика основного канала и подобрана нелинейная функция влияния, аппроксимирующая амплитудную зависимость отклика от напряженности поля во всем диапазоне. ***Научная новизна.*** Предложена передаточная характеристика для интермодуляционного метода, учитывающая совместный вклад нелинейностей первого и второго порядка. В структуру модели внедрена степенная функция четвертого порядка с амплитудным ограничением, что позволяет адекватно описывать нелинейный отклик в расширенной напряженности поля. ***Результаты.*** На основе экспериментальных данных показано, что интермодуляционный метод увеличивает нелинейный отклик объекта в четыре раза по сравнению с методом противополярных ступенчатых сигналов. Параметризованная поведенческая модель описывает этот отклик с погрешностью не более 7,5%, что является удовлетворительным результатом. ***Практическая значимость.*** Параметризованная модель может быть использована на этапе системного проектирования систем нелинейного зондирования для прогнозирования нелинейных откликов малоразмерных объектов.

Ключевые слова: нелинейное рассеяние, сверхширокополосный сигнал, интермодуляционный метод, поведенческое моделирование.
DOI: 10.21293/1818-0442-2026-29-2-74-81

Введение

Системы нелинейного зондирования, использующие сверхширокополосные (СШП) тестовые сигналы, обладают рядом преимуществ. Такие сигналы дают возможность увеличивать пиковую мощность зондирующего сигнала при фиксированной средней мощности. В условиях мощных воздействий объект оказывается в режиме более сильной нелинейности, что ведет к увеличению уровня нелинейного отклика. Кроме того, использование сверхширокополосных сигналов способствует улучшению разрешающей способности по дальности и проникающей способности за счет наличия низкочастотной части спектра [1–4]. Практическое применение данных преимуществ упирается в сложность селекции нелинейного отклика на фоне сплошного спектра линейного отклика.

На сегодняшний день в работе [5] рассмотрен метод селекции нелинейного отклика малоразмерного объекта на СШП-импульсы различной амплитуды и формы. Однако реализация указанного метода возможна лишь в цифровом виде, что подразумевает использование прямой оцифровки принимаемого сигнала. Собственные нелинейные искажения оцифровщиков составляют 0,2…0,3%, что значительно больше уровня относительного нелинейного отклика типичных малоразмерных электронных устройств (0,01% [6] и менее). Данная проблема, в частности, решаема за счет синтеза таких тестовых сигналов, которые увеличивают уровень нелинейного отклика от объектов (прежде всего относительно общего отклика).

Наши исследования показывают, что классический моноцикл и ступенчатый импульс вызывают весьма слабый нелинейный отклик от электронных объектов. Причина заключается в том, что в таких тестовых сигналах невозможно раздельно оптимизировать низкочастотные и высокочастотные составляющие сигнала. Необходимость этого обусловлена двумя ранее установленными обстоятельствами: низкочастотная область спектра сигнала (100…300 МГц) влияет на степень нелинейности состояния объекта, а высокочастотная область – на считывание этого состояния. Невозможность раздельного управления этими обстоятельствами и приводит к низкому уровню нелинейного отклика.

Для получения больших величин относительного нелинейного отклика нами предложен и экспериментально проверен интермодуляционный метод зондирования, подразумевающий воздействие на объект вместо одного импульса со сплошным спектром двумя: в области нижних и верхних частот соответственно (двухкомпонентный сигнал).

Помимо максимизации относительного нелинейного отклика интермодуляционный метод зондирования дает возможность уменьшить влияние нелинейности приемного устройства на результаты измерений, поскольку в этом случае высокочастотную часть двухкомпонентного сигнала можно использовать с одинаковой полярностью (фазой) в двух сравниваемых тестах. При этом при обработке нелинейные искажения приемником двух импульсов будут практически компенсировать друг друга.

Для описания нелинейных свойств объектов в рамках интермодуляционного метода зондирования необходима соответствующая поведенческая модель. Данная модель создала бы возможность эффективно проектировать на системном уровне системы зондирования, в которых в качестве информативного признака используется нелинейный отклик объекта на сверхширокополосный сигнал. Ранее нами была разработана модель нелинейного рассеяния СШП-сигналов при альтернативном методе зондирования (в качестве тестовых сигналов выступала пара ступенчатых импульсов близкой формы и противоположной полярности) [7]. Основные постулаты, исходя из которых синтезирована структура этой модели, сохраняются и для интермодуляционного метода зондирования, но новый режим воздействия требует заново синтезировать модель параметрически.

Целью данной работы является параметризация поведенческой модели [7] для интермодуляционного метода нелинейного сверхширокополосного зондирования.

Интермодуляционный метод нелинейного СШП-зондирования

В работе [5] метод нелинейного сверхширокополосного зондирования состоит в воздействии на малоразмерный объект парой ступенчатых сигналов близкой формы, но противоположной полярности (рис. 1, штриховая кривая). Нелинейные отклики получают в результате суммирования противополярных откликов объекта, что показывает разницу их форм. В интермодуляционном же методе на объект воздействует двухкомпонентная структура зондирующего сигнала, состоящая из высокочастотного вейвлета и низкочастотного импульса нелинейного воздействия (см. рис. 1, сплошная кривая).

Низкочастотный импульс, практически не создавая регистрируемого отражения, оказывает нелинейное воздействие на внутренние цепи объекта. Высокочастотный вейвлет, наоборот, формирует отклик, отражающий состояние объекта с учетом этого воздействия, но не способен изменить степень его нелинейности. Из этого следует, что во втором тестовом сигнале нет необходимости инвертировать низкочастотный импульс для получения воздействия противоположного знака, достаточно просто его отключить. Разность откликов объекта на вейвлет при наличии и отсутствии низкочастотного импульса позволяет выделить нелинейную составляющую.

Сама концепция интермодуляционного метода нелинейного зондирования известна из литературы [8]. Однако в [8] она реализована с использованием гармонического сигнала на фиксированной частоте 850 МГц в качестве импульса нелинейного воздействия. Практическая реализация такого подхода имеет существенные недостатки: применяемый гармонический сигнал обладает фиксированными параметрами, что делает его уязвимым к действию входных преселекторов объекта и исключает возможность поиска оптимальной формы импульса для максимизации нелинейного воздействия.

Предложенная выше двухкомпонентная структура позволяет отказаться от поиска компромисса между противоречивыми требованиями по уровню нижних и верхних частот в спектре используемого импульса и оптимизировать импульс нелинейного воздействия и считывающий импульс раздельно. Вторым преимуществом интермодуляционного метода является отсутствие надобности в чересперiодной инверсии импульсов, что дает возможность использовать существующие наработки по генераторам импульсов большой мощности. Достаточно лишь частоту повторения импульсов нелинейного воздействия выбрать вдвое меньше частоты считывающих импульсов.

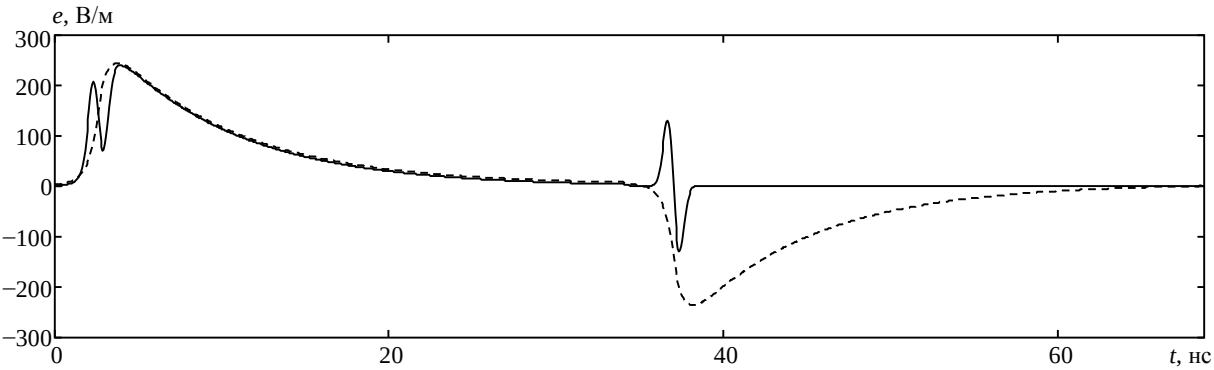


Рис. 1. Тестовые воздействия методов нелинейного зондирования: сплошная кривая – интермодуляционный метод, штриховая кривая – метод противополярных ступенчатых сигналов

Одинаковая полярность считывающих импульсов приводит также к тому, что пары этих импульсов (в присутствии импульса нелинейного воздействия и без) испытывают почти одинаковые нелинейные искажения в приемном устройстве. При обработке нелинейные искажения приемником двух импульсов практически компенсируют друг друга, что значительно уменьшает влияние нелинейности приемного устройства на результаты измерений.

Помимо этого, экспериментальная проверка показала, что нелинейный отклик объекта (FM-приемник Precision PS-298), зарегистрированный с применением интермодуляционного метода измерения, в четыре раза больше, чем отклик на простой ступенчатый импульс при примерно одинаковых амплитудах напряженностей поля низкочастотной составляющей сигнала (около 400 В/м) (рис. 2).

Структура использованной поведенческой модели нелинейного рассеяния

Существующие подходы к моделированию нелинейных динамических систем при сверхширокополосных воздействиях, такие как ряды Вольтерра [9–11], нелинейные фильтры [12] и метод расщепления [13], имеют недостатки: они либо предъявляют строгие требования к корреляционным свойствам тестовых сигналов, либо требуют детальной информации о внутренней нелинейной динамике объекта, либо основаны на трансверсальных фильтрах без обратных связей (что нерационально для устройств с экспоненциальными переходными характеристиками и не дает физически интерпретируемых характеристических функций). В работе [7] используется упрощенная модель, лишенная этих недостатков. Она строится на передаточной характеристике и функционале влияния, что позволяет обойтись без сложных математических описаний внутренней динамики.

Поведенческая модель нелинейного рассеяния сверхширокополосных сигналов, представленная на рис. 3, основана на том наблюдении, что при сплошном спектре тестового воздействия форму нелинейного отклика $\varepsilon(t)$ с точностью до амплитуды можно получать из тестового поля при помощи передаточной характеристики $H_n(\omega)$. Стоит отметить, что поскольку $H_n(\omega)$ связывает потенциал рассеянного поля с напряженностью падающего поля, то ее значения получаются с размерностью длины

$$H_n(\omega) = F[\varepsilon(t)]/F[e(t)], \quad (1)$$

F – прямое преобразования Фурье.

Амплитуда моделируется отдельно при помощи функционала влияния $A[e(t)]$ [8]:

$$\varepsilon(t) = A[e(t)]\, F^{-1}\{H_n(\omega)F[e(t)]\}, \quad (2)$$

где $A[e(t)]$ – функционал, учитывающий степень отличия текущего тестового воздействия от того, которое применялось при вычислении передаточной функции объекта $H_n(\omega)$, F и F^{-1} – прямое и обратное преобразования Фурье.

Правильное значение $A[e(t)]$ можно получить, если пропустить тестовый сигнал $e(t)$ через фильтр нижних частот $H_a(\omega)$. Упомянутая передаточная функция фильтра нижних частот моделирует инерционность нелинейных цепей объекта: низкочастотные составляющие тестового сигнала $e(t)$ больше влияют на степень проявления нелинейных свойств, чем высокочастотные [7]. Такая передаточная функция $H_a(\omega)$ позволяет описывать амплитуду нелинейного отклика для тестовых сигналов с различной формой и амплитудой.

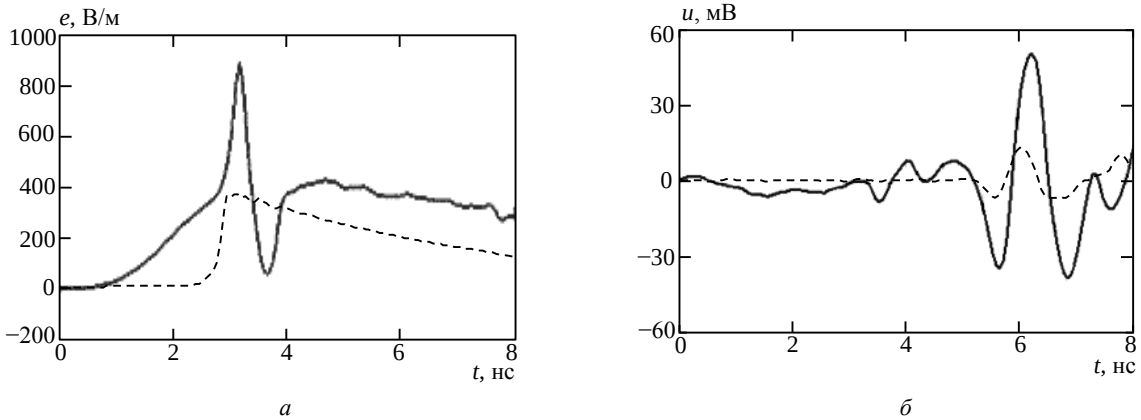
$$H_a(\omega) = 1/(1 + j\omega\tau), \quad (3)$$

τ – постоянная времени, определяющая частоту среза фильтра нижних частот.

Параметр τ выбран вариационным методом для нескольких пар сигналов разной формы с целью минимизации погрешности моделирования амплитуды нелинейного отклика [7].

Для выбранного объекта (FM-приемник Precision PS-298) значение получилось равным 0,57 нс для тестов противополярными импульсами.

В настоящей работе мы приняли то же значение τ, основываясь на том, что инерционность внутренних цепей объекта является его собственным свойством и не зависит от формы тестовых сигналов и метода тестирования.



а *б*

Рис. 2. Экспериментальные тестовые сигналы (*а*) и нелинейные отклики FM-приемника Precision PS-298 (*б*): сплошная кривая – для двухкомпонентного тестового сигнала, штриховая кривая – для ступенчатого сигнала

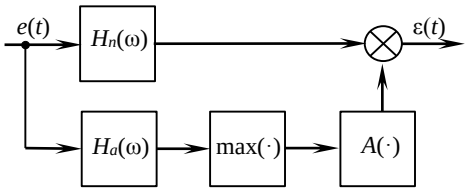


Рис. 3. Функциональная схема поведенческой модели нелинейного рассеяния сверхширокополосных сигналов

По амплитуде выходного сигнала фильтра с передаточной функцией $H_a(\omega)$ определяется значение функционала влияния при помощи некоторой нелинейной функции влияния $A(\cdot)$.

$$A[e(t)] = A[\max(F^{-1}(H_a(\omega)F[e(t)]))]. \qquad (4)$$

В прототипе модели [7] $A(\cdot)$ имеет вид простой степенной функции второго порядка.

В прототипе характеристика $H_n(\omega)$ (рис. 4, штрихпунктирная кривая) в наблюдаемом диапазоне частот (до 2 ГГц) имеет характер фильтра верхних частот. Поскольку при расчете $H_n(\omega)$ применены отклик и напряженность тестового поля, соответствующие методу зондирования с противополярными ступенчатыми сигналами, передаточная характеристика учитывает компоненты нелинейности практически только второго порядка. Вследствие этого максимум модуля передаточной функции смещен вверх относительно максимума амплитудного спектра тестового сигнала (см. рис. 4, сплошная кривая), эффективная полоса которого составляет 1,7 ГГц (от 0,3 до 2 ГГц), что соответствует критерию СШП (более 500 МГц). При переходе к интермодуляционному методу этот эффект не будет выражен в такой степени, поэтому модель-прототип, синтезированная под противополярное ступенчатое воздействие, прогнозирует нелинейный отклик со значительной погрешностью.

Существенный вклад в погрешность использованной модели [7] вносит также имплементированная кубическая амплитудная характеристика (рис. 5, штриховая кривая). Третий порядок образуется от того, что в основном канале поведенческой модели (передаточная характеристика $H_n(\omega)$) сигнал возрастает линейно и квадратично возрастает функционал влияния.

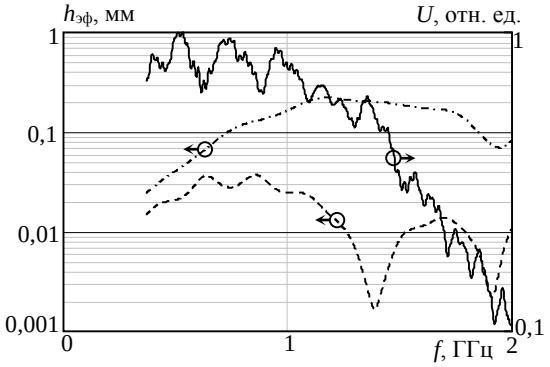


Рис. 4. Спектр тестового сигнала (сплошная кривая) и передаточные характеристики: штриховая кривая – для интермодуляционного метода, штрихпунктирная кривая – для метода с противополярными ступенчатыми сигналами

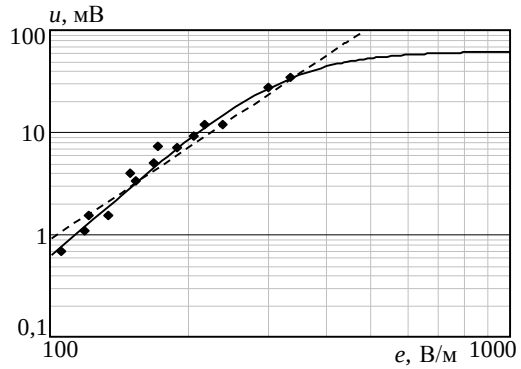


Рис. 5. Амплитудные характеристики потенциала нелинейного отклика: штриховая кривая – кубическая, сплошная кривая – четвертого порядка с амплитудным ограничением, ромбы – экспериментально измеренная

Такая кубическая амплитудная характеристика адекватна лишь в ограниченном диапазоне напряженностей поля, так как потенциал нелинейного отклика не может возрастать неограниченно (как это предсказывает степенная функция третьего порядка) [14]. Это подтверждается известными амплитудными характеристиками рассеивателей для гармонического режима [15], а также нашими предварительными оценками.

Параметризация поведенческой модели для интермодуляционного метода нелинейного зондирования

Как было отмечено ранее, основные постулаты, определяющие структуру поведенческой модели (см. рис. 3), сохраняют свою силу и для интермодуляционного метода зондирования, но двухкомпонентный режим воздействия требует заново синтезировать модель параметрически.

Функционал влияния $A[e(t)]$ в модели для интермодуляционного метода зондирования сохраняет структуру выражения (4). В данном выражении значение постоянной времени τ остается неизменным (0,57 нс) для передаточной функции фильтра нижних частот $H_a(\omega)$, позволяющей описывать амплитуду нелинейного отклика для тестовых воздействий различной формы. Таким образом, в выражении (4) параметризации подлежит только нелинейная функция влияния $A(\cdot)$.

Было выявлено, что в прототипе модели для интермодуляционного метода зондирования существенный вклад в погрешность моделирования нелинейного отклика вносит имплементация кубической амплитудной характеристики. Для устранения этой погрешности на основе серии измерений нелинейного рассеяния сверхширокополосных сигналов (см. рис. 5, ромбы) до напряженности поля, равной 390 В/м, нами уточнена амплитудная характеристика нелинейного рассеивателя в расширенном диапазоне напряженностей [14]. Она основана на четвертой степени полинома с амплитудным ограничением (см. рис. 5, сплошная кривая), при этом нелинейный отклик ε связан с амплитудой напряженности зондирующего поля e следующим соотношением:

$$\varepsilon(e) = \frac{ae^4 \cdot \varepsilon_{макс}}{ae^4 + \varepsilon_{макс}}, \quad (5)$$

где a – коэффициент, определяющий эффективность рассеяния, $\varepsilon_{макс}$ – предельное значение нелинейного отклика. Основываясь на измерениях, для того же самого объекта зондирования мы получили $a = 3{,}28\times10^{-12}$ м^4/В^3, $\varepsilon_{макс} = 62$ мВ.

Для удобства параметризации формулы (5) целесообразно коэффициент a заменить выражением $\varepsilon_{макс}/e_{макс}^4$, в котором $e_{макс}$ – значение напряженности поля, при котором полиномиальная часть достигает предельного значения $\varepsilon_{макс}$. В результате подстановки получаем

$$\varepsilon(e) = \varepsilon_{макс} \times \frac{e^4 / e_{макс}^4}{e^4 / e_{макс}^4 + 1}. \quad (6)$$

Параметр $e_{макс}$ можно рассчитать из уже полученных значений коэффициентов a и $\varepsilon_{макс}$. Имеем $e_{макс} = 371$ В/м.

Введение ограничения $\varepsilon_{макс}$ описывает уменьшение крутизны амплитудной характеристики при увеличении тестового поля, что, в свою очередь, позволяет избавиться от неограниченного роста потенциала нелинейного отклика (как это предсказывает степенная функция третьего порядка). Предельное значение $e_{макс}$ позволяет осуществить переход от монотонного увеличения напряженности поля к насыщению.

Имплементация указанной амплитудной характеристики в параметризованную модель достигается путем соответствующей параметризации нелинейной функции влияния $A(e)$. Функционально ее зависимость такая же, как и в (6), за исключением меньшей на единицу степени аргумента в числителе (так как сигнал в основном канале модели нарастает пропорционально первой степени аргумента):

$$A(e) = a_{макс} \times \frac{e^3 / e_{макс}^4}{e^4 / e_{макс}^4 + 1}, \quad (7)$$

где $a_{макс}$ – параметр, определяющий максимальное значение $A(e)$.

Нужно иметь в виду, что, в соответствии со структурой модели, аргументом $A(e)$ является напряженность поля, взвешенная передаточной функцией $H_a(\omega)$. Эта операция уменьшает амплитуду использованных ступенчатых сигналов (см. рис. 2, *а*, штриховая кривая) на 14,3%. Соответственно, параметр $e_{макс}$ становится равным 318 В/м.

Параметр $a_{макс}$ в формуле (7) определяется исходя из того, что функция влияния должна иметь единичное значение для взвешенной напряженности поля $e_н$, при которой определяется функция $H_n(\omega)$. В нашем случае $e_н = 150$ В/м. Приравняв (7) к единице и подставив $e_н$ вместо e в (7), получим

$$a_{макс} = \frac{e_н^4 / e_{макс}^4 + 1}{e_н^3 / e_{макс}^4}. \quad (8)$$

Подставив (8) в (7), после преобразований будем иметь

$$A(e) = \frac{e^3}{e_н^3} \times \frac{e_н^4 / e_{макс}^4 + 1}{e^4 / e_{макс}^4 + 1}. \quad (9)$$

Параметризация нелинейной функции влияния $A(e)$ по выражению (9) позволяет реализовать в модели амплитудную характеристику с ограничением по амплитуде (рис. 6, сплошная кривая). Максимальное значение этой функции составляет 5,5. Если не учитывать данную аппроксимацию, то функция влияния приняла бы вид, соответствующий штриховой кривой на рис. 6. Такая характеристика дает удовлетворительный результат моделирования амплитуды нелинейного отклика только в диапазоне напряженностей 150…300 В/м. При меньших и больших напряженностях результат моделирования оказался бы со значительной погрешностью.

Для завершения параметризации модели, представленной на рис. 3, необходимо определить передаточную характеристику в основном канале модели. Для расчета передаточной характеристики $H_n(\omega)$ по выражению (1) использованы экспериментальные отклик и тестовый сигнал, соответствующие интермодуляционному методу зондирования (см. рис. 2, сплошные кривые).

Стоит отметить, что в методе с противополярными ступенчатыми сигналами передаточная характеристика рассчитана для тестового сигнала с напряженностью поля 175 В/м [8] в целях обеспечения достаточной амплитуды нелинейного отклика. В экспериментах с интермодуляционным методом зондирования был использован тестовый сигнал с взвешенной пиковой напряженностью 504 В/м. Для корректного сравнения передаточных характеристик в двух различных методах зондирования $H_n(\omega)$ нужно разделить на значение функции влияния при напряженности 504 В/м (см. рис. 6, сплошная кривая). Полученная в результате передаточная функция $H_n(\omega)$ для интермодуляционного метода зондирования представлена на рис. 4 штриховой кривой.

Передаточная характеристика для интермодуляционного метода зондирования имеет важные особенности.

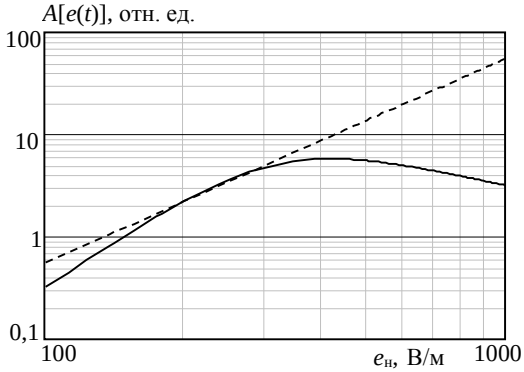


Рис. 6. Функция влияния третьего порядка с амплитудным ограничением (сплошная кривая) и квадратичная функция влияния (штриховая кривая)

Во-первых, ее максимум спектра располагается ниже по сравнению с передаточной характеристикой метода с противополярными ступенчатыми сигналами (см. рис. 4, штрихпунктирная кривая). Благодаря этому она захватывает спектральную область тестового сигнала (см. рис. 4, сплошная кривая) с большей энергией. Во-вторых, для интермодуляционного метода характеристика отражает вклад компоненты нелинейности как первого, так и второго порядка, тогда как характеристика для метода с противополярными ступенчатыми сигналами – только второго.

Погрешность модели, параметризованной для интермодуляционного метода, составляет 7,5% (рис. 7). Данное значение получено как относительное отклонение пикового значения модельного нелинейного отклика от экспериментального. Результат является удовлетворительным, поскольку в данном случае моделируются нелинейные искажения сигналов, которые сами являются формой погрешности.

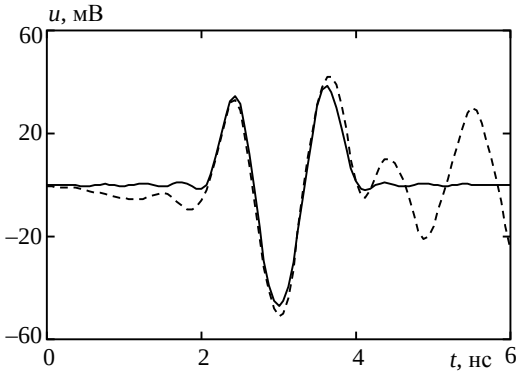


Рис. 7. Нелинейные отклики объекта, полученные интермодуляционным методом: сплошная кривая – модельный; штриховая кривая – экспериментальный

Заключение

Экспериментальная проверка интермодуляционного метода показала, что нелинейный отклик объекта при его использовании в четыре раза превышает отклик, полученный методом с противополярными ступенчатыми сигналами при одинаковой амплитуде напряженности поля. Это подтверждает практическую значимость данного метода и обосновывает целесообразность создания параметризованной модели для проектирования систем нелинейного зондирования.

В работе выполнена параметризация поведенческой модели нелинейного рассеяния сверхширокополосных сигналов применительно к интермодуляционному методу зондирования. Прототип поведенческой модели, синтезированной для метода с противополярными ступенчатыми сигналами, прогнозирует нелинейный отклик двухкомпонентного сигнала со значительной погрешностью. В отличие от прототипа параметризованная модель учитывает двухкомпонентную структуру зондирующего сигнала.

В параметризованной модели имплементирована амплитудная зависимость нелинейного отклика от напряженности поля в виде степенной функции четвертого порядка с амплитудным ограничением. Четвертый порядок образуется от того, что в основном канале поведенческой модели (передаточная характеристика $H_n(\omega)$) сигнал возрастает линейно и кубически возрастает функционал влияния. Это позволяет в режиме экстраполяции корректно давать оценку напряженности зондирующего поля, дающую максимальный нелинейный отклик при сверхширокополосном воздействии.

Передаточная характеристика для интермодуляционного зондирования, описывающая преобразование формы тестового сигнала в форму нелинейного отклика в наблюдаемом диапазоне частот (до 2 ГГц), имеет максимум спектра в низкочастотной области, где сосредоточена большая часть энергии тестового сигнала. Также она учитывает компоненты нелинейности как первого, так и второго порядка.

Поведенческая модель нелинейного рассеяния сверхширокополосных сигналов, параметризованная для интермодуляционного метода зондирования, обеспечивает погрешность моделирования не более 7,5%, что является удовлетворительным результатом.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (проект № FWRM-2024-0001).

Литература

1. Авдеев В.Б. Сверхкороткоимпульсная сверхширокополосная нелинейная радиолокация / В.Б. Авдеев, А.В. Бердышев, С.Н. Панычев // Телекоммуникации. – 2006. – № 8. – С. 23–27.

2. Якубов В.П. Использование сверхширокополосного излучения для диагностики нелинейностей / В.П. Якубов, Д.В. Лосев, А.И. Мальцев // Журнал радиоэлектроники. – 2000. – № 3. – С. 2.

3. Liu S. Research on ultra-wideband NLFM waveform synthesis and grating lobe suppression / S. Liu, Y. Jia, Y. Liu, X. Zhang // Sensors. – 2022. – Vol. 22. – No. 24. – P. 9829.

4. Mishra A. A review: Recent progress in the design and development of nonlinear radars / A. Mishra, C. Li // Remote Sensing. – 2021. – Vol. 13, No. 24. – P. 4982.

5. Modeling and experimental study of the nonlinear baseband pulse radar prototype / E.V. Semyonov, A.V. Fateev, M.A. Nazarov, A.A. Berezin // Radiation and Scattering of Electromagnetic Waves: proceedings of IEEE International conference. – Divnomorskoe, Russia: IEEE, 2023. – P. 428–431.

6. Щербаков Г.Н. Применение нелинейной радиолокации для дистанционного обнаружения малоразмерных объектов // Специальная техника. – 1999. – № 1. – С. 34–39.

7. Features and behavioral modeling of ultra-wideband signals nonlinear scattering by small-sized electronic devices / E.V. Semyonov, M.A. Nazarov, K.M. Poltorykhin, A.A. Berezin, A.V. Fateev // IEEE Transactions on Radar Systems. – 2025. – Vol. 3. – P. 843–851.

8. Шипилов С.Э. Нелинейные преобразования сигналов в импульсной радиотомографии / С.Э. Шипилов, В.П. Якубов // Известия высших учебных заведений. Физика. – 2020. – Т. 63, № 2. – С. 5–14.

9. Абас В. М. А. Моделирование нелинейных динамических и стационарных систем на основе интегрофункциональных рядов Вольтерры и различных классов квадратурных формул / В.М.А. Абас, Р.В. Арутюнян // Математическое моделирование и численные методы. – 2021. – № 2. – С. 68–85.

10. Воскобойников Ю.Е. Идентификация квадратичного ядра уравнения Вольтерра для моделирования нелинейных динамических систем / Ю.Е. Воскобойников, В.А. Боева // Системы анализа и обработки данных. – 2022. – Т.85, № 1. – С. 25–40.

11. Солодуша С. В. Об одном способе моделирования нелинейной динамики полиномами Вольтерра и цепными дробями / С.В. Солодуша, Ю.И. Коконова // Аналитические и численные методы моделирования естественно-научных и социальных проблем: сб. ст. по матер. XVII Всерос. с междунар. участием науч.-техн. конф. / под общ. ред. И.В. Бойкова. – Пенза: Пенз. гос. ун-т, 2022. – С. 145–149.

12. Pedro J.C. A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches / J.C. Pedro, S.A. Maas // IEEE Trans. on Microwave Theory and Techniques. – 2005. – Vol. 53, No. 4. – P. 1150–1163.

13. Ланнэ А.А. Синтез нелинейных систем // Электронное моделирование. – 1980. – № 1. – С. 60–68.

14. Безрукова И.В. Амплитудная характеристика нелинейного рассеивателя в режиме сверхширокополосного воздействия / И.В. Безрукова, Э.В. Семенов, К.М. Полторыхин, И.А. Градов // Науч.-техн. конф. Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио: сборник докладов. – СПб.: СПбГЭТУ «ЛЭТИ», 2026. – № 1 (81). – С. 40–43.

15. Горбачев А.А. Амплитудные характеристики нелинейных рассеивателей / А.А. Горбачев, С.В. Ларцов, С.П. Тараканков, Е.П. Чигин // Радиотехника и электроника. – 1996. – Т. 41, № 5. – С. 558–562.

Безрукова Ирина Владимировна
Техник Института сильноточной электроники
СО РАН (ИСЭ СО РАН)
Академический пр-т, 2/3, г. Томск, Россия, 634055;
инженер каф. радиоэлектроники и систем связи (РСС)
Томского государственного университета
систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0009-0007-1911-8550
Тел.: (382-2) 49-15-44
Эл. почта: bezrukova_i03@mail.ru

Семенов Эдуард Валерьевич
Д-р техн. наук, ст. науч. сотр. ИСЭ СО РАН
Академический пр-т, 2/3, г. Томск, Россия, 634055;
проф. каф. РСС ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0001-5470-1185
Тел.: (382-2) 49-15-44
Эл. почта: edwardsemyonov@narod.ru

Поступила в редакцию: 07.05.2026.
Принята к публикации: 01.09.2026.

Bezrukova I.V., Semyonov E.V.

Modeling and features of the nonlinear response of small electronic devices using the intermodulation method of ultra-wideband sensing

Relevance. The use of ultra-wideband signals in nonlinear sensing systems is limited due to the fact that the useful nonlinear response of typical small electronic objects is significantly weaker than the inherent distortions in the receiving paths and digitizers. To address this challenge, it is necessary to increase the relative level of the nonlinear response. A promising solution is an intermodulation sensing method using a two–component signal, where the low-frequency part affects the degree of nonlinearity of the object's state, and the high-frequency part affects the reading of this state, which makes it possible to optimize both components separately and achieve a significant increase in response compared to antipolar step pulses. ***Purpose of the study*** is to parameterize a behavioral model of nonlinear scattering in relation to the intermodulation sensing method to predict the nonlinear response of small electronic devices over an extended range of field strengths. ***Methods.*** The work is based on a behavioral model of nonlinear scattering, which describes the transformation of a test signal's shape into a nonlinear response via a transfer characteristic calculated from experimentally recorded pairs of «probing field – nonlinear response» in the frequency domain. Parameterization for the intermodulation technique is based on experimental data. For a two-component signal, the main channel's transfer characteristic is calculated and a nonlinear influence function that approximates the response amplitude's dependence on field strength across the entire range is selected. ***Novelty.*** A transfer characteristic is proposed for the intermodulation method, taking into account both first- and second-order nonlinear effects. A fourth-order power function, with an amplitude constraint, is incorporated into the model, allowing adequate description of the nonlinear response over a wide range of field strengths. ***Results.*** Based on experimental data, it is shown that the intermodulation method increases the nonlinear response of an object by four times compared to the method of opposite-polarity step signals. The parameterized behavioral model describes this response with an error not exceeding 7.5%, which is considered an acceptable result. ***Practical significance.*** The parameterized model can be utilized during the design phase of nonlinear sensing systems in order to forecast the nonlinear responses of small objects.

Keywords: nonlinear scattering, ultra-wideband signal, intermodulation method, behavioral modeling.
Funding: The work was carried out within the state assignment of the Ministry of Education and Science of Russia (project No. FWRM-2024-0001).
DOI: 10.21293/1818-0442-2026-29-2-74-81

References

1. Avdeev V.B., Panychev S.N. *Sverhkorotkoimpul'snaja sverhshirokopolosnaja nelinejnaja radiolokacija* [Ultra-Short-Pulse Ultra-Wideband Nonlinear Radar]. *Telekommunikatsii (Telecommunications)*, 2006, no. 8, pp. 23–27 (in Russ.).

2. Yakubov V.P., Losev D.V., Maltsev A.I. *Ispol'zovanie sverhshirokopolosnogo izlucheniya dlya diagnostiki nelinejnostej* [The Use of Ultra-Wideband Radiation for the Diagnosis of Nonlinearities]. *Zhurnal radioelektroniki (Journal of Radio Electronics)*, 2000, no. 3 (in Russ.).

3. Liu S., Jia Y., Liu Y., Zhang X. Research on ultra-wideband NLFM waveform synthesis and grating lobe suppression. *Sensors*, 2022, vol. 22, no. 24, p. 9829.

4. Mishra A., Li C. A review: Recent progress in the design and development of nonlinear radars. *Remote Sensing*, 2021, vol. 13, no. 24, p. 4982.

5. Semyonov E.V., Fateev A.V., Nazarov M.A., Berezin A.A. Modeling and experimental study of the nonlinear baseband pulse radar prototype. Radiation and Scattering of Electromagnetic Waves: proceedings of International IEEE Conference, Divnomorskoe, Russia, 2023, pp. 428–431.

6. Shcherbakov G.N. *Primenenie nelinejnoj radiolokacii dlya distancionnogo obnaruzheniya malorazmernyh obĕktov*

[Application of Nonlinear Radar for Remote Detection of Small Objects]. *Special'naya tekhnika*, 1999, no. 1, pp. 34–39 (in Russ.).

7. Semyonov E.V., Nazarov M.A., Poltorykhin K.M., Berezin A.A., Fateev A.V. Features and behavioral modeling of ultra-wideband signals nonlinear scattering by small-sized electronic devices. *IEEE Transactions on Radar Systems*, 2025, vol. 3, pp. 843–851.

8. Shipilov S.E., Yakubov V.P. *Nelinejnye preobrazovaniya signalov v impulsnoj radiotomografii* [Nonlinear signal transformations in pulse radiotomography]. *Izvestiya vysshix uchebnyx zavedenij. Fizika*, 2020, vol. 63, no 2, pp. 5–14 (in Russ.).

9. Abas W.M.A., Arutyunyan R.V. [Modeling of nonlinear dynamic and stationary systems based on Volterra integro-functional series and various classes of quadrature formulas]. *Mathematical Modeling and Computational Methods*, 2021, no 2, pp. 68–65 (in Russ).

10. Voskoboinikov Yu.E., Boeva V.A. [Identification of the quadratic kernel of the Volterra equation for modeling nonlinear dynamic systems]. *Sistemy analiza i obrabotki dannykh–Analysis and Data Processing Systems*, 2022, vol. 85, no. 1, pp. 25–40. (in Russ).

11. Solodusha S.V., Kokonova Yu.I. *Ob odnom sposobe modelirovaniya nelinejnoj dinamiki polinomami Vol'terra i cepnymi drobyami* [On one method for modeling nonlinear dynamics by Volterra polynomials and continued fraction]. Analiticheskie i chislennye metody modelirovaniya estestvenno-nauchnyh i social'nyh problem. [*Analytical and numerical methods of modelling of natural science and social problems*]. Proc. of the XVII All-Russian with Intern. Participation Sci.-Tech. Conf, Penza, Penza State University Publ., 2022, pp. 145–149 (in Russ).

12. Pedro J.C., Maas S.A. A Comparative Overview of Microwave and Wireless Power-Amplifier Behavioral Modeling Approaches. *IEEE Trans. on Microwave Theory and Techniques*, 2005, vol. 53, no. 4, pp. 1150–1163.

13. Lanne A.A. *Sintez nelinejnyx system* [Synthesis of nonlinear systems]. *Elektronnoe modelirovanie*, 1980, no. 1, pp. 60–68. (in Russ).

14. Bezrukova I.V., Semenov E.V., Poltorykhin K.M., Gradov I.A. *Amplitudnaya xarakteristika nelinejnogo rasseivatelya v rezhime sverxshirokopolosnogo vozdejstviya* [Amplitude characteristic of a nonlinear scatterer in the mode of ultra-wideband exposure]. Nauchno-texnicheskaya konferenciya sankt-peterburgskogo NTO RES im. A.S. Popova, posvyashhennaya dnyu radio. [*Scientific and Technical Conference of the St. Petersburg NTO RES named after A.S. Popov, dedicated to Radio Day*]. Proc. of St. Petersburg State Technical University «LETI», 2026, pp. 40–43 (in Russ.).

15. Gorbachev A.A., Lartsov S.V., Tarakankov S.P., Chigin E.P. *Amplitudnye xarakteristiki nelinejnyx rasseivatelej* [Amplitude characteristics of nonlinear scatterers]. *Radiotexnika i elektronika*, 1996, vol. 41, no 5, pp. 558–562 (in Russ.).

Irina V. Bezrukova
Technical Staff, Institute of High Current Electronics SB RAS (HCEI SB RAS)
2/3, Akademichesky Ave., Tomsk, Russia, 634055;
Engineer, Department of Radioelectronics and Communication Systems, Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0009-0007-1911-8550
Phone: (382-2) 49-15-44
Email: bezrukova_i03@mail.ru

Edward V. Semyonov
Doctor of Engineering Sciences, Senior Researcher, HCEI SB RAS
2/3, Akademichesky Ave., Tomsk, Russia, 634055;
Professor, Department of Radioelectronics and Communication Systems, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0001-5470-1185
Phone: (382-2) 49-15-44
Email: edwardsemyonov@narod.ru

Received: 07.05.2026.
Accepted: 01.09.2026.

УПРАВЛЕНИЕ, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАТИКА

В разделе статьи по следующим специальностям (технические науки)

1.2.2. Математическое моделирование, численные методы и комплексы программ.
2.3.1. Системный анализ, управление и обработка информации.
2.3.2. Вычислительные системы и их элементы.
2.3.3. Автоматизация и управление технологическими процессами и производствами.
2.3.4. Управление в организационных системах.
2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.
2.3.6. Методы и системы защиты информации, информационная безопасность.
2.3.8. Информатика и информационные процессы.

УДК 621.36.41
Шифр специальности ВАК: 1.2.2

Е.Е. Слядников, И.Ю. Турчановский

Математическая модель неравновесных структурных переходов в аморфных металлических сплавах при термомеханических воздействиях

Актуальность. Выявление физической природы и механизмов изменения внутренней структуры аморфных металлических сплавов при воздействии на них инфразвуком в настоящее время представляет значительный интерес и является актуальным направлением исследований. ***Цель исследования.*** Формулировка математической модели влияния термомеханического воздействия на неравновесные структурные переходы в аморфных металлических сплавах при воздействии инфразвуком. ***Методы.*** Для моделирования влияния термомеханического воздействия на структурные переходы использован метод кинетических уравнений для параметра превращения, температуры, механического напряжения. С помощью метода конечных разностей система кинетических уравнений дискретизирована и выполнены численные расчеты модельных задач. ***Научная новизна.*** В работе представлена новая модель влияния термомеханического воздействия на изменение внутренней структуры аморфных сплавов. ***Результаты.*** Полученные на основе представленной в работе математической модели результаты расчетов методом конечных разностей позволяют объяснить физические механизмы, а также причины и условия последовательности структурных переходов в рассматриваемой системе. Это, в свою очередь, позволяет объяснить экспериментально полученную закономерность изменения внутренней структуры материала при инфразвуковом воздействии. ***Практическая значимость.*** Полученные результаты формируют новые представления о физической природе процессов изменения внутренней структуры материалов при термомеханическом воздействии, которые могут быть использованы в электронике, материаловедении для совершенствования известных и создания новых приборов и материалов. Правила изменения состояний клеточного автомата могут быть использованы для развития метода клеточных автоматов на неравновесные структурные переходы в аморфных металлических сплавах.

Ключевые слова: математическая модель, численный расчет, аморфные металлические сплавы, изменения внутренней структуры, инфразвук.
DOI: 10.21293/1818-0442-2026-29-2-85-92

Введение

Современные аморфные материалы являются неоднородными на многих структурных уровнях, начиная с субнаноструктурного уровня и наноструктурного уровня [1–3]. Экспериментально установлено, что физические и механические свойства аморфных материалов зависят и контролируются субнано-, нано- и микроструктурой, формируемой во время протекания сложных неравновесных структурных переходов, стимулированных термомеханическими воздействиями [4–7].

В работах [4–6] образцы аморфной тонкой ленты подвергали 10-минутному воздействию инфразвука, обладающего частотой 20 Гц при амплитуде смещений 1 и 4 мкм в условиях поддержания постоянной температуры T_0 = 25 °C. При амплитуде смещения 1 мкм аморфная структура образца практически не изменяется, а при амплитуде смещения 4 мкм в аморфной матрице обнаружены кристаллические кластеры с размерами 3–5 нм [4–6]. Увеличение продолжительности воздействия до 2 ч увеличивает размер кристаллических нанокластеров, а воздействие в течение 4 ч приводит к возникновению микрокристаллической структуры, состоящей из зерен различных кристаллографических ориентаций [7]. Физическая природа и механизмы такого влияния инфразвуковых колебаний на изменение внутренней структуры аморфного металлического сплава (нанокристаллизацию и полимикрокристаллизацию) остаются неясными.

Поэтому для системного развития технологий создания новых аморфных металлических материалов, обладающих уникальными физическими, механическими свойствами и широко применяемых в электронике для создания высокоэффективных трансформаторов, миниатюрных дросселей, датчиков тока, а также легких и ударопрочных корпусов для мобильных устройств, необходимо понимание физики протекающих процессов и математическое моделирование кинетики сложных неравновесных структурных переходов.

В работах [8–10] была высказана гипотеза о физическом механизме нанокристаллизации, сформулированы феноменологическая кинетическая модель, физическая причина и условия неравновесного перехода первого рода в металлической пленке из аморфного состояния в наноструктуру, стимулированного инфразвуком.

Однако феноменологическая модель применима для описания процессов на микроскопическом и миллиметровом структурном уровне. А реальное возникновение нанокристаллов происходит на субнаноструктурном и наноструктурном уровне, поэтому для исследования этого процесса неравновесного перехода первого рода в наносистеме и верификации гипотезы, механизма и модели нанокристаллизации [8–10] необходимо применение теоретических методов исследования и численного моделирования.

Согласно классификации [11], компьютерное моделирование различных процессов в аморфных материалах содержит, в частности, статистические [12], детерминированные [13], кинетические [14] подходы. В этих моделях субнано-, нано- и микроструктура делается дискретной с помощью решетки ячеек, причем границы, разделяющие различные элементы субнано-, нано- и микроструктуры (например, субнанокластеры с кристаллическим ближним порядком, нанокластеры с кристаллическим порядком, микрозерна с кристаллическим порядком), являются резкими, что приводит к существенным вычислительным сложностям.

Для преодоления этих сложностей в данной работе предлагаются математическая и вычислительная модели кинетики сложного неравновесного структурного перехода с образованием новой субнано-, нано- и микроструктуры, стимулированного термомеханическим воздействием. Ключевая особенность модели неравновесных структурных превращений [8–10], описываемых параметрами превращения, состоит в том, что границы различных элементов нано- и микроструктуры (нанокластеров и микрозерен) являются размытыми (диффузными), т.е. гладкими, что приводит к существенному упрощению вычислений.

На основе предложенной вычислительной модели предлагается развить вычислительный метод клеточных автоматов [15] для исследования кинетики сложного неравновесного структурного перехода в аморфных металлических сплавах, стимулированного термомеханическим воздействием. Поскольку вычислительная модель описывает конкретную кинетику неравновесного структурного превращения аморфное – нанокристаллическое – полимикрокристаллическое состояние, то она позволяет сформулировать конкретные правила к модели клеточных автоматов, описывающей этот сложный процесс на различных пространственно-временных масштабах (субнано, нано, микро).

Математическая модель сложного неравновесного структурного перехода в аморфных металлических сплавах

Характерная особенность аморфных металлических сплавов (АМС), обладающих некристаллическим ближним порядком (БП), – первоначальная неоднородность на субнано-, наноструктурном уровнях. Существуют зоны (субнано- и нанокластеры и т.д.), обладающие избыточным свободным объёмом (ИСО). Изначально АМС находятся в неравновесном состоянии. Внешние термомеханические воздействия приводят к изменению как субнано-, наноструктуры, так и формированию микроструктуры. Материалы с новыми субнано-, нано-, и микроструктурами обладают уникальными физическими и механическими свойствами. Какой физический механизм формирования таких структур? Очевидно, что при термомеханическом воздействии в АМС протекают неравновесные структурные переходы (НСП), которые являются универсальным физическим механизмом изменения структуры и свойств материалов.

Неупругая, пластическая деформация АМС на различных стадиях, стимулированная термомеханическим воздействием, представляет собой упорядоченную по времени взаимосвязанную последовательность НСП первого рода [8–10], протекающих в системе зон, обладающих ИСО, на различных пространственно-временных уровнях.

Внесенная в систему энергия (работа внешних сил, переданное из внешней среды тепло) создает неустойчивость первоначального стационарного состояния системы зон с начальным ближним и средним порядком, относительно НСП первого рода в новое стационарное состояние системы зон с новым ближним и средним порядком, ориентированным в направлении механического поля. Новое стационарное состояние может иметь более низкое значение свободной энергии, чем первоначальное состояние, поэтому оно может быть глобально устойчивым. Однако новое стационарное состояние может иметь более высокое значение свободной энергии, быть неравновесным и существовать, пока есть поток энергии в систему из внешней среды. Неравновесный переход зон сопровождается неупругой деформацией зон (возникают зоны неупругой деформации) и поглощением скрытой энергии превращения первого рода. Согласно закону сохранения энергии, работа внешних сил, затраченная на неупругую деформацию зон, идет на увеличение кинетической энергии частиц (увеличение температуры) зон и на увеличение потенциальной энергии взаимодействия частиц (поглощение скрытой энергии превращения) зон. После прекращения внешнего воздействия начинается процесс релаксации системы зон из неравновесного стационарного состояния в состояние статистического равновесия.

При термомеханическом воздействии перестройка атомов БП, сопровождающаяся неупругой деформацией, происходит как путем локальной термической флуктуации, так и путем локальной механической активации. Механизмом НСП первого рода является взаимодействие перестроенных зон, которое возникает, когда их концентрация становится достаточно большой при достижении критического значения плотности энергии термомеханического воздействия. То есть неупругая деформация образца не может быть результатом независимых перестроек зон. Перестройки зон стремятся к корреляции. Когда происходит перестройка зоны в данном месте образца, вероятность следующего события с перестройкой зоны должна быть выше в месте, близком к месту первоначального события. Неупругая деформация образца – кооперативный процесс.

Рассмотрим простую макроскопическую систему – металлический стержень, находящийся в состоянии аморфного стекла, к которому прикладывается одноосное механическое напряжение и подводится (отводится) тепло. Мысленно разобьем макроскопическую систему на наноразмерные подсистемы, содержащие большое количество кластеров БП с ИСО. При НСП первого рода кластер БП с ИСО

испытывает неупругую перестройку, ориентированную в направлении одноосной деформации. Введем безразмерную величину, характеризующую структурное состояние наноразмерной подсистемы, которая называется параметром БП [8–10]. В начальном состоянии стекла величина η равна нулю, а в конечном нанокристаллическом состоянии величина η равна единице.

Начальное и конечное состояния подсистемы являются локально устойчивыми и разделены энергетическим барьером. Подсистема в состоянии аморфного стекла характеризуется температурой перехода первого рода T_c и температурой абсолютной неустойчивости начального состояния T_2, причем $T_2 > T_c$.

Чтобы начальное состояние, например, при начальной температуре T_0, стало абсолютно неустойчивым, подсистема должна получить критическое количество тепла, равное $Q_c = C_V(T_2 - T_0)$. При механическом воздействии критическое значение работы внешних сил, связанное с неупругой деформацией, при котором начальное состояние становится абсолютно неустойчивым, равно $A_c = \int_0^{\xi} \sigma(\varepsilon_{in}) d\varepsilon_{in}$, ξ – неупругая деформация насыщения. При совместном термомеханическом воздействии на изменение устойчивости начального состояния подсистемы влияет сумма работы внешних сил и количества тепла от внешнего источника $\int_0^{\varepsilon_{in}} \sigma(\varepsilon_{in}) d\varepsilon_{in} + \int_0^{t} \dot{Q} dt$.

Для такой простой модели свободной энергии системы F однородное уравнение Ландау–Халатникова для эволюции параметра БП примет вид

$$\frac{d\eta}{dt} = -\frac{1}{\tau_\eta E_Q^*}[(\partial F/\partial\eta) - \xi\sigma], \qquad (1)$$

$$\frac{\partial F/\partial\eta}{E_Q^*} = \eta(\eta - \eta_2)(\eta - 1),$$

$$\eta_2 = \frac{1}{2}\left[1 - \frac{\int_0^{\varepsilon_{in}} \sigma(\varepsilon_{in}) d\varepsilon_{in}}{A_c} - \frac{\int_0^{t} \dot{Q} dt}{Q_c}\right],$$

$$\tau_\eta(T,\sigma) = \tau_0 e^{-(E_\eta - \sigma\xi V_\eta)/kT} + \Theta(\sigma - \sigma_1)\cdot\tau_\sigma, \qquad (2)$$

где E_η – энергия активации тепловой флуктуации; σ – абсолютная величина механического напряжения; V_η – активационный объем кластера, испытавшего атомную перестройку; $\tau_0^{-1} = \omega_0 \approx 10^{13}$ Гц, t – время теплового воздействия; $\Theta(\sigma - \sigma_1)$ – тета-функция; σ_1 – предел упругости; τ_σ – характерное время механически активированной перестройки атомов БП; $\tau_\eta(T,\sigma)$ – время релаксации, которое содержит два члена: «термофлуктуационный» и «механически активированный» [8, 9]. Когда плотность энергии термомеханического воздействия достигает критического значения, начальное состояние становится абсолютно неустойчивым η, происходит конденсация зон деформации с кристаллическим БП, возникает нанокристаллическое состояние.

Однородное уравнение баланса тепла с учетом тепловыделения имеет вид

$$C_V\dot{T} = \dot{Q} + W = \dot{Q} - \left\{\frac{\partial F}{\partial\eta} - \nu\xi\sigma\right\}\dot{\eta},$$

$$W = -\left\{\frac{\partial F}{\partial\eta} - \nu\xi\sigma\right\}\dot{\eta} = \Gamma\left\{\frac{1}{E_Q^*}\left[\frac{\partial F}{\partial\eta} - \nu\xi\sigma\right]^2\right\}, \qquad (3)$$

где C_V – удельная теплоемкость, T – абсолютная температура; $\dot{Q}$ – мощность внешнего теплового источника. Сумма скорости изменения свободной энергии $-\frac{\partial F}{\partial\eta}\dot{\eta}$ при неравновесном процессе и доли скорости совершения работы неупругой деформации $\nu\zeta\sigma\dot{\eta} = \nu\sigma\dot{\varepsilon}_{in} > 0$ определяет скорость диссипации энергии W (3). Величина ν, изменяющаяся от нуля до единицы, характеризует долю работы неупругой деформации, которая превращается в тепло.

Сумма упругой и неупругой деформации [8–10] формирует полную деформацию одномерного образца ε:

$$\varepsilon = \varepsilon_{el} + \varepsilon_{in} = \varepsilon_{el} + \xi\eta, \quad \frac{d\varepsilon(t)}{dt} = \frac{1}{\mu}\frac{d\sigma(t)}{dt} + \xi\frac{d\eta(t)}{dt}. \qquad (4)$$

При $\varepsilon = \text{const}$, $\dot{\varepsilon} = 0$ уравнение (4) редуцируется в уравнение

$$\frac{d\sigma(t)}{dt} = \mu\left[\frac{d\varepsilon(t)}{dt} - \xi\frac{d\eta(t)}{dt}\right], \quad \varepsilon = \text{const},$$

$$\frac{d\sigma(t)}{dt} = -\mu\xi\frac{d\eta(t)}{dt}, \qquad \frac{d\sigma(\eta)}{d\eta} = -\mu\xi, \qquad (5)$$

$$\sigma(\eta) = \sigma(0) - \xi\mu\eta = \mu(\varepsilon_{el} - \xi\eta) = \mu(\varepsilon_{el} - \varepsilon_{in}), \qquad (6)$$

где ν – модуль упругости. Если НСП ($\eta = 0$) отсутствует, механическое напряжение увеличивается пропорционально деформации. При критическом значении термомеханического воздействия в точке ε_1 протекает структурное превращение ($\eta = 1$), возникает неупругая деформация $\varepsilon_2(\eta = 1) = \xi$ и происходит сброс механического напряжения на величину $-\xi\mu$.

При дальнейшем увеличении термомеханического воздействия эволюция структуры системы, испытывающей сложное неравновесное структурное превращение первого рода аморфное состояние – нанокристаллическое состояние – полимикрокристалл, может быть описана следующим образом. Рассмотрим простую макроскопическую систему – металлический стержень, находящийся в нанокристаллическом состоянии, к которому прикладывается одноосное механическое напряжение и подводится (отводится) тепло. Мысленно разобьем макроскопическую систему на микроразмерные подсистемы, содержащие некоторое количество критических крис-

таллических зародышей – критических нанокристаллов. Для описания НСП первого рода в микроразмерной подсистеме из начального состояния с несколькими критическими нанокристаллами в конечное полимикрокристаллическое (зернистое) состояние при термомеханическом воздействии введем набор параметров порядка $\eta_i(r,t) = \eta_i(r - r_i, t)$, i = 1, 2, 3, …, I, где I – число параметров порядка, равное числу критических зародышей. Пусть для определенности параметр порядка $\eta_i(r,t)$ описывает i-е кристаллическое зерно с i-й кристаллографической ориентацией. Эта величина может принимать действительные значения из интервала от 0 (нанокристаллическое состояние) до 1 (кристаллическое состояние всего микрозерна). Например, $\eta_1(r - r_1, 0) = 1$ означает, что в начальный момент времени кристаллический зародыш первого зерна в окрестности точки r_1 имеет кристаллографическую ориентацию 1, а $\eta_2(r - r_2, 0) = 1$ означает, что в начальный момент времени кристаллический зародыш второго зерна в окрестности точки r_2 имеет кристаллографическую ориентацию 2. В приграничной области между зерном 1 с ориентацией 1 и зерном 2 с ориентацией 2 соответствующие параметры порядка будут иметь значения между 1 и 0. Тогда функционал полной свободной энергии такой подсистемы, содержащей I зерен, имеет вид

$$F[\eta(r,t)] = \int_V \Big\{ f(\eta_1(r,t), \eta_2(r,t), ...\eta_I(r,t)) + \sum_{i=1}^{I} \frac{D_i}{2} \cdot [\nabla \eta_i(r,t)]^2 \Big\} dV , \quad (7)$$

где $F[\eta(r,t)]$ – функционал полной свободной энергии подсистемы, содержащей объемную и поверхностную (градиентную) энергию; $f(\eta_1(r,t), \eta_2(r,t), ..., \eta_I(r,t))$ – плотность объемной свободной энергии подсистемы; D_i – коэффициенты градиентной энергии подсистемы, V – объем подсистемы. Градиентный член имеет физический смысл энергетического штрафа на неоднородности поля параметра превращения, которая возникает на границе зерна. Поверхностная свободная энергия связана с наличием диффузной (размытой) границы зерна.

Релаксационная кинетика параметров порядка, а следовательно, и кинетика миграции межзеренных границ описывается системой уравнений Ландау–Халатникова

$$\frac{\partial \eta_i(r,t)}{\partial t} = -\gamma_i \frac{\delta F[\eta_1(r,t), \eta_2(r,t)...]}{\delta \eta_i(r,t)} = -\gamma_i \{ \frac{\partial f(\eta_1, \eta_2...)}{\partial \eta_i} - D_i \nabla^2 \eta_i \} , \quad i = 1, 2, 3, \ldots, I, \quad (8)$$

γ_i – коэффициент релаксации подсистемы, описывающий мобильность границ зерен. Для системы, содержащей I зерен и испытывающей сложное неравновесное структурное превращение первого рода, выберем плотность объемной свободной энергии в виде [8–10]

$$f(\eta_1, ...\eta_I) = \sum_{i=1}^{I} \left[\frac{1}{4}\eta_i^4 - \frac{1}{3}(1 + \eta_i^*)\eta_i^3 + \frac{1}{2}\eta_i^* \eta_i^2 - \frac{\xi\sigma\eta_i}{E_Q^*} \right] + \lambda \sum_{i=1}^{I} \sum_{j \neq i}^{I} \eta_i^2 \eta_j^2 , \quad (9)$$

где λ – константа ингибирующего взаимодействия различных параметров порядка, η^*_i – параметр неравновесности системы в окрестности точки r_i, E_Q^* – скрытая энергия структурного превращения, причем $\xi\sigma/E_Q^* << 1$ [10]. Значения феноменологических параметров и коэффициенты градиента энергии определяют энергию границ зерен (7) и толщину границ. Из (9) видно, что свободная энергия имеет $2I$ потенциальных минимумов, которым соответствуют стационарные состояния зерен. При фиксированных значениях феноменологических параметров можно определить как энергию границы раздела, так и ширину границы, которые регулируются путем изменения коэффициентов градиентной энергии. Чем больше коэффициент градиентной энергии, тем больше ширина границы. Когда коэффициент градиентной энергии стремится к нулю, толщина границы становится бесконечно тонкой. Используя выражения (8), (9) получим соответствующую систему кинетических уравнений Ландау–Халатникова, описывающих сложное неравновесное превращение первого рода:

$$\frac{\partial \eta_i(r,t)}{\partial t} = -\gamma_i \{ \eta_i(\eta_i - \eta_i^*)(\eta_i - 1) - \frac{\xi\sigma}{E_Q^*} + 2\lambda\eta_i \sum_{j \neq i}^{I} \eta_j^2 - D_i \nabla^2 \eta_i \} . \quad (10)$$

Для I = 1 одно уравнение эволюции описывает кинетику роста и миграции границы одного зерна с образованием мономикрокристалла, когда $I >> 1$, набор уравнений описывает кинетику роста и движения границ I зерен с образованием полимикрокристалла.

Численное моделирование НСП первого рода с образованием микроструктуры

Выполним численное моделирование НСП первого рода с образованием полимикрокристалла, описываемого системой уравнений (10); для набора параметров порядка необходимо сделать ряд шагов:

– определить физически адекватный набор параметров порядка для моделирования сложного неравновесного структурного перехода;

– сформулировать полную свободную энергию системы как функционал набора параметров порядка, которая правильно описывает ее термодинамическое поведение;

– выбрать значения феноменологических коэффициентов, входящих в кинетические уравнения, из экспериментальных данных или фундаментальных расчетов;

– ввести соответствующие начальные и граничные условия для численного решения кинетических уравнений.

Для численного решения системы кинетических уравнений (10) необходимо провести их дискретизацию в пространстве координат и на временной оси. Например, для одномерного образца оператор Лапласа может быть представлен в конечных разностях в виде

$$\frac{\partial^2}{\partial x^2}\eta(x) = \frac{1}{(\Delta x)^2}[\eta(x+\Delta x)+\eta(x-\Delta x)-2\eta(x)], \quad (11)$$

где Δx является длиной ячеек, на которые разбит образец. Для дискретизации времени запишем частную производную по времени в конечных разностях

$$\frac{\partial\eta(t)}{\partial t} = \frac{\eta(t+\Delta t)-\eta(t)}{\Delta t}, \quad (12)$$

где Δt – временной шаг.

Входящие в кинетические уравнения (10) константы γ_i и D_i имеют размерность $с^{-1}$ и $м^2$ соответственно. Значения этих величин зависят от длины одномерной подсистемы и длины критического зародыша. Выберем длину критического зародыша равной 10 нм, а величину $\sqrt{D_i} = 5$ нм равной половине длины критического зародыша. Размер ячейки Δx не может быть больше длины критического зародыша. Выберем длину ячейки $\Delta x = 10$ нм. Тогда длина подсистемы, разбитой на 50 ячеек, составит $L = 500$ нм.

Уравнения (10) можно сделать безразмерными, введя безразмерное время $\tilde{t} = \gamma_i t / 150$ и безразмерную координату $\tilde{x} = x / L$, а также безразмерный параметр $\tilde{D}_i = D / L^2 = 0{,}0001$. Для применимости схемы расчета должно выполняться условие $\Delta t / (\Delta x)^2 < 1$.

Для простоты и наглядности все численные результаты были получены для одномерной безразмерной системы уравнений (10), используя схему дискретизации (11), (12) с конкретными значениями феноменологических констант $\gamma_i = 150$ (независимость коэффициента от i – изотропная мобильность границ зерен); $D_i = 0{,}0001$ (независимость коэффициента от i – изотропная энергия границ зерен); $\lambda = 20$, $\eta_1^* = \eta_2^* = 0{,}25$, $\sigma_i = 0$. Одномерный образец длиной 1 разбивался на 50 ячеек длиной 0,02. Характерное время релаксации подсистемы $150 / \gamma_i$, выбранное за единицу, разбивалось на 2500 интервалов длиной 0,0004. Начальное условие для параметра порядка 1 выбиралось $\eta_1(x, 0) = 0$, и для параметра порядка 2 выбиралось $\eta_2(x, 0) = 0$. Граничное условие для параметра порядка 1 выбиралось $\eta_1(0, t) = 1$, а для параметра порядка 2 выбиралось $\eta_2(50, t) = 1$.

Выбор таких граничных условий физически означает, что на левой и правой границах образца возникли критические зародыши – нанокристаллы. Для перечисленных выше начальных, граничных условий и феноменологических коэффициентов после 2500 шагов по времени получились профили для параметра порядка 1 (рис. 1, *а*) и для параметра порядка 2 (см. рис. 1, *б*).

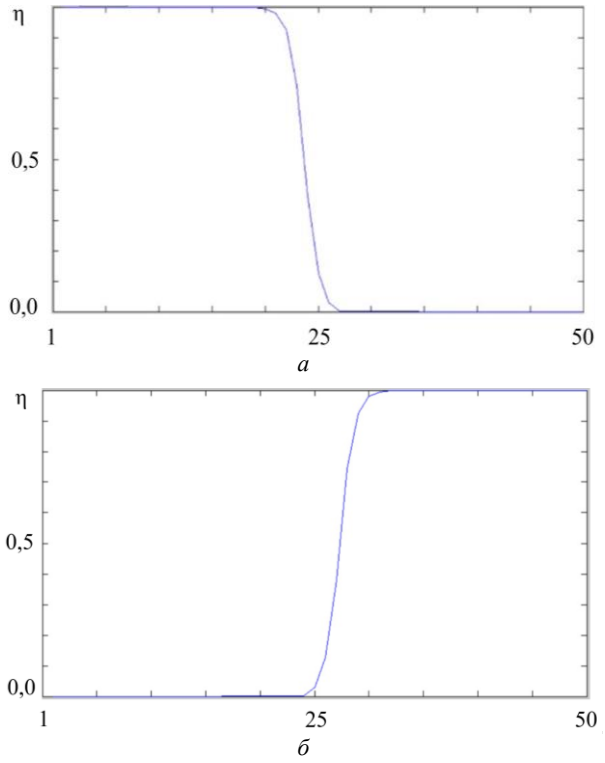


Рис. 1. Зависимость профиля параметра порядка от пространственной координаты, выраженной в ячейках разбиения, после 2500 шагов по времени: *а* – с ориентацией 1 зерна 1; *б* – с ориентацией 2 зерна 2

Для визуализации эволюции микроструктуры образца, состоящего из двух зерен с ориентацией 1 и 2, и описываемой двумя параметрами порядка, использовалась функция

$$\Psi(x,t) = \eta_1^2(x,t) + \eta_2^2(x,t). \quad (13)$$

Зависимость этой функции от пространственной координаты после 2500 шагов по времени показана на рис. 2.

Результаты численных расчетов позволяют сделать вывод, что возникшие в аморфной матрице критические нанокристаллы различных кристаллографических ориентаций сначала растут, а когда их границы соприкасаются, рост зерен прекращается и возникает полимикрокристаллическая структура.

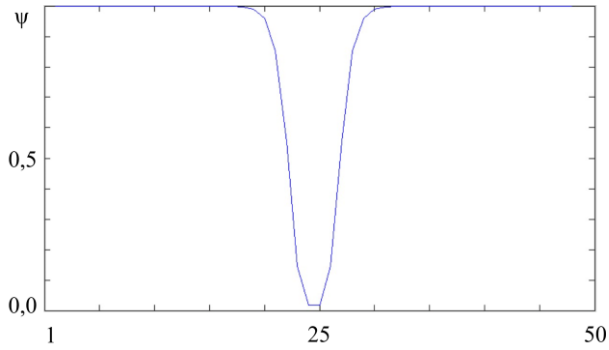


Рис. 2. Микроструктура образца $\Psi(x) = \eta_1^2(x) + \eta_2^2(x)$, состоящего из двух зерен с ориентацией 1 и 2, и описываемая двумя параметрами порядка, после 2500 шагов по времени

Вычислительный алгоритм к модели клеточных автоматов

Результаты дискретизации и численного исследования одномерной системы дифференциальных уравнений для физических переменных (на примере полей параметров порядка, температуры, напряжения), описывающих сложный неравновесный переход первого рода нанокристаллическое состояние – полимикрокристалл, позволяют сформулировать вычислительный алгоритм к модели клеточных автоматов, описывающей этот сложный процесс на различных пространственно-временных масштабах (субнано, нано, микро):

1. Задать поля температуры $T(t,x)$, параметров порядка $\eta_i(t,x)$, напряжения $\sigma(t)$, в которых координата автомата x меняется дискретно с шагом Δx, а время – с шагом Δt. Автомат, расположенный слева от автомата с координатой x, имеет координату $x-\Delta x$, а автомат, расположенный справа от автомата с координатой x, имеет координату $x+\Delta x$. В отличие от вычислительной модели (10)–(13) в модели клеточных автоматов каждая клетка автомата определяет ячейку пространства, в которой параметр порядка, температура, напряжение меняются дискретно.

2. Кроме полей температуры, параметров порядка, напряжения, в каждой ячейке клеточного автомата необходимо вычислить значения различных видов свободной энергии системы, параметра неравновесности $\eta_i^*(T,\sigma)$.

3. Используя дискретизацию одномерного дифференциального уравнения Ландау–Халатникова, задать правило изменения параметра порядка в клетке автомата со временем

$$\eta_i(x,t+\Delta t)-\eta_i(x,t) = -2\gamma_i\Delta t\Big\{\eta_i(x,t)(\eta_i(x,t)-\eta_i^*)\times$$
$$\times(\eta_i(x,t)-1)-\frac{\xi\sigma}{E_Q^*}+2\lambda\eta_i(x,t)\sum_{j\neq i}^{I}\eta_j^2-$$
$$-\frac{D_i}{(\Delta x)^2}[\eta_i(x+\Delta x,t)+\eta_i(x-\Delta x,t)-2\eta_i(x,t)]\Big\}. \quad (14)$$

4. Переход автомата в состояние с новым значением параметра порядка возможен при условиях, если хотя бы один из его соседей уже находится в кристаллическом состоянии, а параметр неравновесности превышает критическое значения.

5. Используя дискретизацию одномерного дифференциального уравнения теплопроводности, задать правило изменения температуры в клетке автомата со временем:

$$T(x,t+\Delta t)-T(x,t) =$$
$$=\frac{\alpha\cdot\Delta t}{(\Delta x)^2}[T(x+\Delta x,t)+T(x-\Delta x,t)-2T(x,t)]+$$
$$+\frac{\tau_\eta Q_T}{\Delta t}[\eta_i(x,t+\Delta t)-\eta_i(x,t)]^2. \quad (15)$$

Здесь α – коэффициент температуропроводности; Q_T – скрытая энергия превращения первого рода на единицу объема, выраженная в единицах температуры; $Q_T = Q/c\cdot\rho$, c – теплоемкость [Дж/кг·К]; ρ – плотность АМС [кг/м^3]; Q – скрытая энергия превращения первого рода на единицу объема [Дж/м^3].

6. Если задана (фиксирована) полная деформация одномерного образца $\varepsilon = \varepsilon_{упр}+\varepsilon_{неупр}$, то механическое напряжение зависит от времени и вычисляется по формуле

$$\sigma(t) = E\varepsilon_{упр}(t) = E[\varepsilon-\frac{\xi}{I}\sum_{i=1}^{I}\eta_i(t)]. \quad (16)$$

Заключение

Сформулирована физико-математическая модель для сложного неравновесного перехода первого рода аморфное состояние – нанокристаллическое состояние – полимикрокристалл (описываемого с помощью системы уравнений Ландау–Халатникова для набора различных параметров порядка). Основная черта этой модели состоит в учете конечной ширины границы между зернами, энергии межзеренных границ, температуры, механического напряжения, а также учете ингибирующего взаимодействия между соседними параметрами порядка, что приводит к самосогласованной эволюции микроструктуры аморфного сплава без отслеживания движения межзеренных границ. Проведена дискретизация модели и выполнены численные расчеты модельных задач.

В рамках предложенной модели изменение внутренней структуры материала при увеличении времени механического воздействия [5–7] происходит путем реализации упорядоченной последовательности структурных превращений первого рода, протекающих на различных пространственно-временных уровнях.

При малой продолжительности механического воздействия перестройка атомов БП, сопровождающаяся неупругой деформацией, происходит путем локальной механической активации [8–10].

Механизмом НСП первого рода является взаимодействие перестроенных зон, которое возникает, когда их концентрация становится достаточно большой. Если плотность энергии механического воздействия достигает первого критического значения, то происходит конденсация зон деформации с кристаллическим БП, возникает нанокристаллическое состояние.

При продолжении механического воздействия нанокристаллы растут, и некоторые из них достигают размеров критического зародыша. При дальнейшем увеличении длительности механического воздействия, когда плотность энергии механического воздействия достигает второго критического значения, критические нанокристаллы начинают нормальный рост, при этом выделяется скрытая теплота превращения первого рода, которая и обеспечивает самопроизвольное развитие процесса кристаллизации. В результате роста зерен с различной кристаллографической ориентацией формируется полимикрокристаллическая структура.

Таким образом, предложенная в работе математическая модель качественно описывает закономерное изменение внутренней структуры аморфных сплавов при увеличении длительности механического воздействия, наблюдаемое экспериментально [5–7].

Результаты численного исследования модели кинетики неравновесных переходов первого рода в аморфных металлических сплавах при термомеханических воздействиях позволили сформулировать вычислительный алгоритм к модели клеточных автоматов, описывающей этот сложный процесс на различных пространственно-временных масштабах (субнано, нано, микро).

Сложный неравновесный структурный переход в АМС (как упорядоченная последовательность простых неравновесных переходов на различных пространственно-временных масштабах) может описываться как процесс в связанной системе моделей клеточных автоматов, соответствующих субнано-, нано-, микропространственно-временным масштабам. Ячейки клеточных автоматов имеют размеры субнано-, нано-, микрометров. Универсальными параметрами ячеек, описывающими процессы на различных масштабных уровнях, являются соответствующие параметры порядка (ближнего порядка с ячейкой $0{,}3\cdot10^{-9}$ м, среднего порядка с ячейкой $10\cdot10^{-9}$ м, микроскопического порядка с ячейкой $500\cdot10^{-9}$ м и т.д.). В качестве правила изменения состояния клетки по параметру порядка разумно использовать дискретизированное уравнение Ландау–Халатникова, в которые входит фундаментальная величина – свободная энергия как функция соответствующих параметров порядка. На каждом структурном уровне свободная энергия включает энергию соответствующих структурных объектов и их взаимодействия. Наличие начальных структурных особенностей для каждого масштабного уровня (кластеры БП с ИСО, критические нанокристаллы, межзеренные границы и т.д.) определяют начальные и граничные условия к уравнениям Ландау–Халатникова.

Работа выполнена в рамках государственного задания Минобрнауки России для Федерального исследовательского центра информационных и вычислительных технологий.

Литература

1. Progress, applications, and challenges of amorphous alloys: A critical review / Z. Feng, H. Geng, Y. Zhuang, P. Li // Inorganics. – 2024. – Vol. 12, No. 9. – P. 232.

2. Nanoscale structural heterogeneity and magnetic properties of Fe-based amorphous alloys via Co and Ni additions / Z.Z. Yang, L. Zhu, S.S. Jiang, C. Zhu, Q.H. Xu, Y. Lin, F.G. Chen, Y.G. Wang // Journal of Alloys and Compounds. – 2022. – Vol. 904. – 25 May. – P. 164067.

3. Zhang G. Advances in amorphous nanomaterials: Synthesis, characterization and application / G. Zhang, H. Hu, J. Wu // Materials Today Nano. – 2025. – December. – Vol. 32. – Article: 100675.

4. Изменение радиуса первой координационной сферы в аморфных сплавах при деформировании / Г.Е. Абросимова, В.В. Астанин, Н.А. Волков, Д.В. Гундеров, Е.Ю. Постнова, А.С. Аронин // Физика металлов и металловедение. – 2023. – Т. 124, № 7. – С. 622–634.

5. Influence of low-frequency vibrations on the structure of amorphous Ti40.7Hf9.5Ni44.8Cu5 alloy / S. Belyaev, N. Resnina, V. Rubanik, A. Shelyakov, V. Niapomniashchay, E. Ubyivovk, I. Kasatkin // Materials letters. – 2017. – Vol. 209. – P. 231–234.

6. Crystallization of amorphous Ti40.7Hf9.5Ni41.8Cu8 alloy during the low-frequency mechanical vibrations at room temperature / S. Belyaev, V. Rubanik, N. Resnina, V. Rubanik, E. Ubyivovk, E. Demidova, A. Uzhekina, I. Kasatkin, A. Shelyakov // Materials Letters. – 2020. – Vol. 275. – P. 128084–128087.

7. Variation in the of the amorphous NiTi-based alloys during mechanical vibrations / S. Belyaev, V. Rubanik, N. Resnina, V. Rubanik, E. Ubyivovk, E. Demidova, A. Uzhekina, I. Kasatkin, A. Shelyakov //Journal of Non-Crystalline Solids. – 2020. – Vol. 542. – P. 120101.

8. Слядников Е.Е. О возможном механизме внешнего инфразвукового механического стимулирования процесса формирования нанокристаллов в аморфной металлической пленке / Е.Е. Слядников, И.Ю. Турчановский // Журнал технической физики. – 2021. – Т. 91, № 11. – С. 1662–1673.

9. Слядников Е.Е. Неоднородная пластическая деформация аморфных металлических сплавов под действием квазистатической механической нагрузки / Е.Е. Слядников, И.Ю. Турчановский // Журнал технической физики. – 2023. – Т. 93, № 1. – С. 104–116.

10. Слядников Е.Е. Модель электропластичности в аморфных металлических сплавах / Е.Е. Слядников, И.Ю. Турчановский // Журнал технической физики. – 2025. – Т. 95, № 10. – С. 2029–2037.

11. Application of Machine Learning in Amorphous Alloys / L. Zhang, H. Zhang, B. Ji, L. Liu, X. Liu, D. Chen // Materials. – 2025. – Vol. 18, No. 8. – Art. 1771.

12. On the reliability of using reverse Monte Carlo simulations to construct the atomic structure model of metallic glasses / Ch. Liu, Zh. Zhang, J. Ding, E. Ma // Scripta Materialia. – 2023. – Vol. 225. – Art. 115159.

13. Molecular simulation-derived features for machine learning predictions of metal glass forming ability / B.T. Afflerbach, L. Schultz, J.H. Perepezko, P.M. Voyles, I. Szlufarska, D. Morgan // Computational Materials Science. – 2021. – Vol. 199. – P. 110728.

14. Модификация кинетической модели Вильсона-Френкеля и атомистическое моделирование скорости плавления/кристаллизации металлов / В.И. Мажукин, А.В. Шапранов, О.Н. Королёва, А.В. Мажукин // Математическое моделирование. – 2023. – Т. 35, № 11. – С. 103–121.

15. Zhi Y. Review on Cellular Automata for Microstructure Simulation of Metallic Materials / Y. Zhi, Y. Jiang, D. Ke // Materials. – 2024. – Vol. 17(6). – P. 1370.

Слядников Евгений Евгеньевич
Проф. каф. электронных приборов
Томского государственного университета
систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 41-38-87
Эл. почта: eeslad1962@gmail.com
Ст. науч. сотр. Томского ф-ла фед. исследовательского центра
информационных и вычислительных технологий (ФИЦ ИВТ)
Академический пр-т, 10/4, г. Томск, Россия, 634055
Тел.: +7 (382-2) 49-17-74
Эл. почта: tomsk@ict.nsc.ru

Турчановский Игорь Юрьевич
Рук-ль Томского ф-ла ФИЦ ИВТ
Академический пр-т, 10/4, г. Томск, Россия, 634055
Тел.: +7 (382-2) 49-17-74
Эл. почта: tomsk@ict.nsc.ru

Поступила в редакцию: 19.05.2026.
Принята к публикации: 22.09.2026.

Slyadnikov E.E., Turchanovsky I.Yu.

Mathematical model of non-equilibrium structural transitions in amorphous metal alloys under thermomechanical loading

Relevance. The identification of the physical nature and mechanisms of internal structural changes in amorphous metal alloys under infrasound exposure is of considerable scientific interest. ***Purpose of the study***. The aim of this study is to formulate a mathematical model describing the effect of thermomechanical loading on nonequilibrium structural transitions in amorphous metal alloys subjected to infrasound exposure. ***Methods.*** To simulate the effect of thermomechanical loading on structural transitions, the method of kinetic equations was employed for the transformation parameter, temperature, and mechanical stress. The system of kinetic equations was discretized using the finite difference method, and numerical calculations were performed for benchmark model problems. ***Scientific novelty***. This work presents a novel model of the effect of thermomechanical loading on changes in the internal structure of amorphous alloys. ***Results.*** A mathematical model (a system of kinetic equations) and numerical calculations using the finite difference method reveal the physical mechanisms, causes, and conditions of the sequence of structural transitions and explain the experimentally obtained pattern of changes in the internal structure of a material under infrasound action. Rules for changing the states of a cellular automaton in a cellular automaton model. ***Practical significance.*** The findings contribute new insights into the physical nature of internal structural transformation processes in materials subjected to thermomechanical loading. These results can be used in electronics and materials science to improve existing devices and materials, as well as for the development of novel ones. The state transition rules of the cellular automaton can be extended to the study of nonequilibrium structural transitions in amorphous metallic alloys, thereby advancing the cellular automaton method in this domain.

Keywords: mathematical model, numerical calculation, amorphous metal alloys, internal structural changes, infrasound.

Funding: The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation for Federal Research Center for Information and Computational Technologies.

DOI: 10.21293/1818-0442-2026-29-2-85-92

References

1. Feng Z., Geng H., Zhuang Y., Li P. Progress, applications, and challenges of amorphous alloys: A critical review. *Inorganics*, 2024, vol. 12, no. 9, pp. 232.

2. Yang Z.Z., Zhu L., Jiang S.S., Zhu C., Xu Q.H., Lin Y., Chen F.G., Wang Y.G. Nanoscale structural heterogeneity and magnetic properties of Fe-based amorphous alloys via Co and Ni additions. *Journal of Alloys and Compounds*, 2022, vol. 904, art. 164067.

3. Zhang G., Hu H., Wu J. Advances in amorphous nanomaterials: Synthesis, characterization and application. *Materials Today Nano*, 2025, vol. 32, art. 100675.

4. Abrosimova G., Gunderov D., Postnova E., Aronin A. Changes in the Structure of Amorphous Alloys under Deformation by High-Pressure Torsion and Multiple Rolling. *Materials*, 2023, vol. 16, no. 3, art. 1321.

5. Belyaev S., Resnina N., Rubanik V., Shelyakov A., Niapomniashchay V., Ubyivovk E., Kasatkin I. Influence of low-frequency vibrations on the structure of amorphous $Ti_{40}.7Hf_9.5Ni_{44}.8Cu_5$ alloy. *Materials letters*, 2017, vol. 209, pp. 231–234.

6. Belyaev S., Rubanik, V., Resnina, N., Rubanik, V., Ubyivovk E., Demidova E., Uzhekina A., Kasatkin I., Shelyakov A. Crystallization of amorphous $Ti_{40}.7Hf_9.5Ni_{41}.8Cu_8$ alloy during the low-frequency mechanical vibrations at room temperature. *Materials Letters*, 2020, vol. 275, pp. 128084–128087.

7. Belyaev S., Rubanik V., Resnina N., Rubanik V., Ubyivovk E., Demidova E., Uzhekina A., Kasatkin I., Shelyakov A. Variation in the of the amorphous NiTi-based alloys during mechanical vibrations. *Journal of Non-Crystalline Solids*, 2020, vol. 542, art. 120101.

8. Slyadnikov E.E., Turchanovsky, I.Y. [On the Possible Mechanism of External Infrasonic Mechanical Stimulating the Process of Formation of Nanocrystals in an Amorphous Metal Film]. *Zhurnal Tekhnicheskoi Fiziki*, 2021, vol. 91, no. 11, pp. 1662–1673 (in Russ.).

9. Slyadnikov E.E., Turchanovsky I.Y. Non-Homogeneous Plastic Deformation of Amorphous Metallic Alloys under the Action of a Quasi-Static Mechanical Load. *Technical Physics*, 2023, vol. 68, no. 1, pp. 99–110.

10. Slyadnikov E.E., Turchanovsky I.Y. Electroplasticity model for amorphous metal alloys. *Technical Physics*, 2025, vol. 70, no. 10, pp. 1911–1918.

11. Zhang L., Zhang H., Ji B., Liu L., Liu X., Chen D. Application of Machine Learning in Amorphous Alloys. *Materials*, 2025, vol. 18, no. 8, art. 1771.

12. Liu Ch., Zhang Zh., Ding J., Ma E. On the reliability of using reverse Monte Carlo simulations to construct the atomic structure model of metallic glasses. *Scripta Materialia*, 2023, vol. 225, art. 115159.

13. Afflerbach B.T., Schultz L., Perepezko J.H., Voyles P.M., Szlufarska I., Morgan D. Molecular simulation-derived features for machine learning predictions of metal glass forming ability. *Computational Materials Science*, 2021, vol. 199, art. 110728.

14. Mazhukin V.I., Shapranov A.V., Koroleva O.N., Mazhukin A.V. Modification of the Wilson–Frankel kinetic model and atomistic simulation of the rate of melting/crystallization of metals. *Math. Models Comput. Simul.*, 2024, vol. 16, pp. 223–234.

15. Zhi Y., Jiang Y., Ke D. Review on Cellular Automata for Microstructure Simulation of Metallic Materials, *Materials*, 2024, vol. 17(6), art. 1370.

Evgeniy E. Slyadnikov
Professor, Department of Electronic Devices,
Tomsk State University of Control Systems
and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
Phone: +7 (382-2) 41-38-87
Email: eeslad1962@gmail.com
Sr. Researcher, Tomsk branch of the Federal Research Center
for Information and Computing Technologies (FRC ICT)
10/4, Academichesky Ave., Tomsk, Russia, 634055
Phone: +7 (382-2) 49-17-74
Email: tomsk@ict.nsc.ru

Igor Yu. Turchanovsky
Head of the Tomsk branch of FRC ICT
10/4, Academichesky Ave., Tomsk, Russia, 634055
Phone: +7 (382-2) 49-17-74
Email: tomsk@ict.nsc.ru

Received: 19.05.2026.
Accepted: 22.09.2026.

УДК 621.382
Шифр специальности ВАК: 1.2.2

Н.М. Богатов, В.С. Володин, Л.Р. Григорьян, М.С. Коваленко

Численное решение уравнения Пуассона в кремниевой *n–i–p*-наноструктуре с учетом поверхностных состояний

Актуальность. В последнее время достаточно широкое распространение получили *n–i–p*-наноструктуры в составе современных электронных приборов. Поверхностные электронные состояния влияют на распределение носителей заряда в объеме кремниевой *n–i–p*-наноструктуры с диффузионным распределением доноров и акцепторов, что изменяет свойства полупроводниковых приборов. Этот эффект изучен не полностью. ***Цель исследования.*** Анализ влияния плотности поверхностного заряда на распределения концентрации электронов, дырок и их дрейфовой скорости в *n–i–p*-наноструктуре. ***Методы.*** Для моделирования изменения концентрации носителей заряда, зависящей от потенциала электрического поля, использован метод численного решения уравнения Пуассона с учетом поверхностных состояний. ***Научная новизна.*** Получены принципиально новые знания о влиянии поверхностных состояний на область пространственного заряда кремниевой *n–i–p*-наноструктуры. ***Результаты.*** Показано, что поверхностные состояния создают область пространственного заряда, в которой уменьшаются разность электрического потенциала на границах структуры и концентрация основных носителей заряда. Поэтому электрическая проводимость поверхностной области оказывается меньше, чем в легированном слое. Дрейфовая скорость термолизованных носителей заряда меняет знак на противоположный в поверхностной области. Уменьшение плотности поверхностных состояний до 10^{16} м$^{-2}$ практически устраняет эти эффекты. ***Практическая значимость.*** Результаты исследования могут быть использованы в полупроводниковой наноэлектронике для совершенствования известных и создания новых приборов.

Ключевые слова: *n–i–p*-наноструктура, потенциал электрического поля, поверхностные состояния, электроны, дырки, кремний, дрейфовая скорость.
DOI: 10.21293/1818-0442-2026-29-2-93-98

Введение

В современной полупроводниковой наноэлектронике используются приборы *n–i–p*-типа (*p–i–n*), например, блоки памяти и защиты от статического электричества, диоды с перестраиваемыми параметрами и др. [1–3]. Одним из физических эффектов, снижающих их параметры, является вклад поверхностных состояний в рекомбинационные, электрические и кинетические свойства носителей заряда.

Расчеты электрофизических свойств классических полупроводниковых структур с *n–p*-переходом основаны на теории Шокли [4]. Модификациями этой теории являются модели резкого и плавного *n–p*-переходов [5, 6]. В расчетах свойств полупроводниковых структур с реальным распределением примесей и в реальных условиях эксплуатации используются численные модели [7–10].

Выпрямляющие свойства *n–p*-перехода определяются электрическим потенциальным барьером и областью пространственного заряда (ОПЗ). Распределение электрического потенциала, плотности заряда, концентрации электронов и дырок в ОПЗ имеет особенности [11], не содержащиеся в классических моделях [4–6]. Расчет структур, содержащих сильно несимметричные, а также сверхмелкие *n–p*-переходы, выполняется с помощью численных моделей.

Уменьшение размеров полупроводниковых структур до наноразмерных масштабов является одной из тенденций полупроводниковой техники. В работах [12, 13] построен численный метод решения уравнения Пуассона в *n–p*-переходе с учетом поверхностных состояний. Влияние плотности поверхностного заряда не ограничивается поверхностной областью. В [14] показано, что напряженность электрического поля в слое *i*-типа проводимости изменяется с уменьшением толщины наноструктуры.

Цель исследования – анализ влияния плотности поверхностного заряда на распределения концентрации электронов, дырок и их дрейфовой скорости в *n–i–p*-наноструктуре.

Методика решения задачи

Модель кремниевой *n–i–p*-наноструктуры включает четыре части (рис. 1): при $-w_n \le x < 0$ расположен диффузионный слой *n*-типа проводимости; при $0 \le x < w_p$ расположен высокоомный слой p_0-типа (база); при $w_p \le x < w_{p2}$ расположен диффузионный слой *p*-типа; при $w_{p2} \le x \le (w_p + L)$ $(L >> w_n, w_p, w_{p2})$ расположен низкоомный слой *p*-типа проводимости. Слой *n*-типа содержит примесные атомы фосфора. Слои *p*-типа содержат примесные атомы бора с различной концентрацией. Поверхностные состояния находятся в слое толщиной w_s.

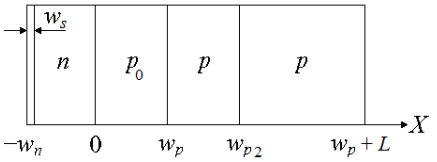


Рис. 1. Условная схема *n–i–p*-наноструктуры

Распределение примесей бора в базе $N_{B0} = \text{const}$, фосфора при $-w_n \le x$

$$N_P(x) = N_{PP}\,\text{erfc}\left(\frac{(x + w_n)}{x_{0P}}\right), \qquad (1)$$

где N_{PP} – поверхностная концентрация, x_{0P} – характеристическая диффузионная длина фосфора. Концентрация бора в p-слое при $w_p \le x < w_{p2}$

$$N(x) = N_B(x) + N_{B0}, \quad (2)$$

$$N_B(x) = (N_{BB} - N_{B0})\operatorname{erfc}\left(\frac{(w_{p2} - x)}{x_{0B}}\right), \quad (3)$$

где N_{BB} – концентрация бора при $x = w_{p2}$, x_{0B} – характеристическая диффузионная длина бора. $N(x) = N_{BB}$ в подложке при $w_{p2} \le x \le (w_p + L)$.

Фосфор и бор являются мелкими примесями в кремнии. Концентрация ионизованных примесей зависит от положения уровня Ферми F в запрещенной зоне.

$$N_P^{\text{ion}}(x) = \frac{N_P(x)}{1 + g_P e^{\frac{F_1(x) - E_g + E_P}{kT}}}, \quad (4)$$

$$N_B^{\text{ion}}(x) = -\frac{N_B(x)}{1 + g_B e^{\frac{E_B - F_1(x)}{kT}}}, \quad (5)$$

где $F_1(x) = F + q\phi(x)$, E_g – ширина запрещённой зоны; E_P и g_P – энергетический уровень в запрещенной зоне и фактор его вырождения для примеси фосфора; E_B и g_B – энергетический уровень в запрещенной зоне и фактор его вырождения для примеси бора; q – элементарный заряд; $\phi(x)$ – электрический потенциал; k – константа Больцмана; T – температура полупроводника. Начало отсчета энергии совпадает с вершиной валентной зоны в точке $x = (w_p + L)$. При $x = 0$ концентрации ионизованных доноров и акцепторов равны.

Заряженные поверхностные состояния расположены в слое толщиной w_s, сравнимой с периодом кристаллической решетки, при $-w_n \le x < (w_s - w_n)$ (см. рис. 1). Обозначим плотность поверхностных состояний N_{ss}, тогда их объемная концентрация равна $N_s(x)$ [12].

$$N_s(x) = \frac{N_{ss} kT}{w_s E_g} \ln \frac{\left(1 + e^{\frac{2F_1(x) - E_g}{2kT}}\right)\left(1 + e^{\frac{E_g - F_1(x)}{kT}}\right)}{\left(1 + e^{\frac{F_1(x)}{kT}}\right)\left(1 + e^{\frac{E_g - 2F_1(x)}{2kT}}\right)}. \quad (6)$$

Концентрация фосфора на поверхности $N_{PP} = 10^{26}$ м$^{-3}$, бора в базе $N_{B0} = 10^{17}$ м$^{-3}$, бора в подложке $N_{BB} = 10^{26}$ м$^{-3}$. При этих значениях для всех $-w_n \le x \le (w_p + L)$ $N_P^{\text{ion}}(x) < N_c$ эффективной плотности электронных состояний в окрестности дна зоны проводимости, $\left|N_B^{\text{ion}}(x)\right| < N_v$ эффективной плотности электронных состояний в окрестности вершины валентной зоны. Тогда в невырожденном состоянии концентрация электронов $n(x)$ и дырок $p(x)$ равна

$$n(x) = N_c e^{\frac{F - E_g + q\phi(x)}{kT}}, \quad p(x) = N_v e^{\frac{-F - q\phi(x)}{kT}}. \quad (7)$$

Электрический потенциал найдем, решая уравнение Пуассона численно, методом [12]. Имеем:

$$\frac{d^2}{dx^2}\phi(x) = -\frac{q}{\varepsilon\varepsilon_0}\left(p(x) - n(x) + N^{\text{ion}}(x) + N_s(x)\right). \quad (8)$$

Уравнение (8) записано в метрической системе СИ, $N^{\text{ion}}(x) = N_P^{\text{ion}} + N_B^{\text{ion}}$ – концентрация ионизованных дефектов с учетом знака.

В рассматриваемой модели выполняется условие электронейтральности

$$\int_{-w_n}^{w_p + L}\left(p(x) - n(x) + N^{\text{ion}}(x) + N_s(x)\right)dx = 0. \quad (9)$$

Из (8) и (9) следует

$$\frac{d}{dx}\phi(x)\Big|_{x=-w_n} = \frac{d}{dx}\phi(x)\Big|_{x=w_p+L} = 0. \quad (10)$$

Считаем, что $\phi(x) = 0$ при $x \ge (w_p + L)$, тогда из (10) следует $\phi(x) = \phi_0$ при $x \le -w_n$. Значение ϕ_0 определяется в процессе численного решения уравнения (8).

В расчётах толщина слоёв n–i–p-наноструктуры $w_n = w_p = (w_{p2} - w_p) = w = 20$ нм, толщина поверхностного слоя $w_s = 1$ нм, температура $T = 300$ К.

Результаты расчётов и их обсуждение

Влияние поверхностных состояний на электрическое поле, концентрацию носителей заряда и ионизованных объемных и поверхностных дефектов продемонстрировано на рис. 2, *а–в*.

Изменение поверхностного потенциального барьера мало в масштабе рис. 2, *а* ($N_{ss} = 10^{16}$ м$^{-2}$). В этом случае поверхностные состояния мало влияют на состав ОПЗ, занимающей весь объем n–i–p-наноструктуры. Отметим особенности строения ОПЗ [13], продемонстрированные на рис. 2, *а*, необходимые для обоснования результатов работы. Область положительного заряда $N_P^{\text{ion}}(x) > n(x)$ расположена при $x \in (-w_n, -8$ нм). Область отрицательного заряда $\left|N_B^{\text{ion}}(x)\right| > p(x)$ расположена при $x \in (26{,}4$ нм, 40 нм). Ионизованные примесные атомы дают основной вклад в значение напряженности электрического поля в базе $E_0 \approx 2 \cdot 10^7$ В/м. Существуют две области: обогащённая электронами при $x \in (-8$ нм, 7,2 нм) и обогащённая дырками при $x \in (13{,}6$ нм, 26,4 нм). Область, обеднённая электронами и дырками ($n, p < N_{B0}$), расположена при $x \in (7{,}2$ нм, 13,6 нм). Области, обогащённые и обеднённые носителями заряда, создают изменение напряжённости $\Delta E(x) << E_0$.

Таким образам, база имеет дипольную структуру. Значение E_0 превышает критическое, так что дрейфовая скорость электронов превышает тепловую [15]. ОПЗ распространяется на всю n–i–p-наноструктуру, её размер 60 нм. В работе [16] отмечается, что для реализации туннельного эффекта средний размер

ОПЗ составляет 10 нм, а при 40 нм активное туннелирование возможно при достаточно большом обратном смещении. Следовательно, вклад квантового туннелирования при толщине трёх слоёв по 20 нм пренебрежимо мал. Применимость данной модели ограничена толщиной слоев ~10 нм.

Плотность поверхностных состояний $N_{ss} = 10^{17}$ м$^{-2}$ на рис. 2, *б*, $N_{ss} = 10^{18}$ м$^{-2}$ на рис. 2, *в*, при этом $N^{\text{ion}}(x) + N_S(x) < 0$ при $x < -1{,}9 \cdot 10^{-8}$ м в поверхностном слое толщиной $w_s = 1$ нм. Особенность распределения концентрации заряженных частиц $Q(x) = p(x) - n(x) + N^{\text{ion}}(x) + N_S(x)$ в поверхностной области поясняет рис. 3.

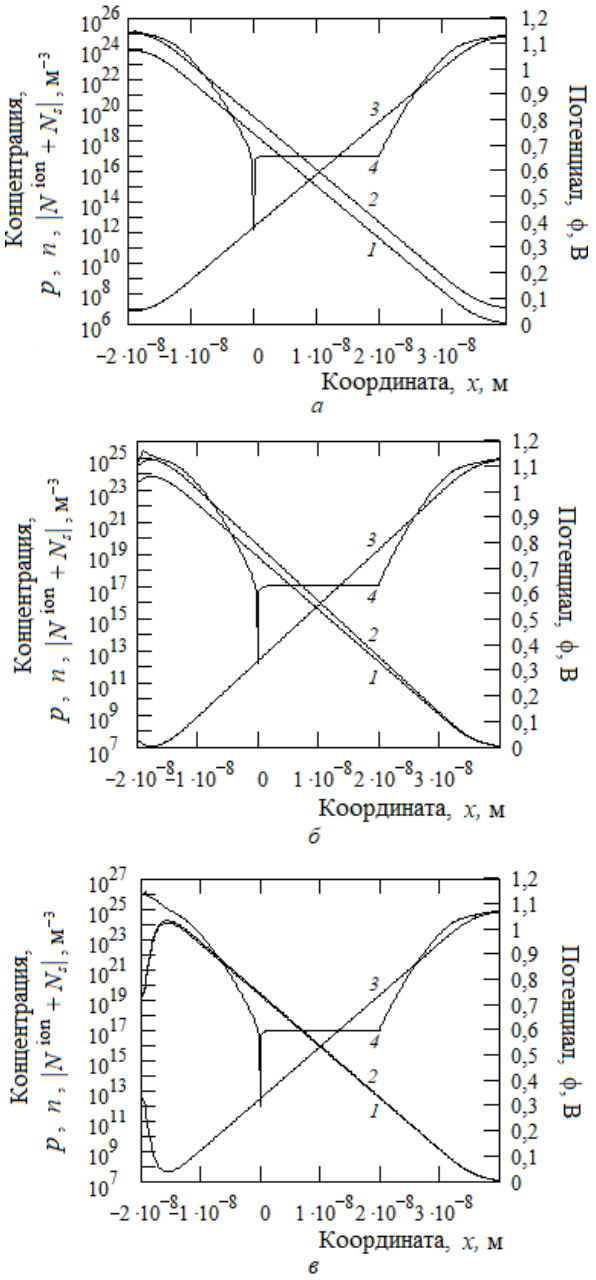


Рис. 2. Электрический потенциал и концентрация частиц:
1 – ϕ(x), *2* – $n(x)$, *3* – $p(x)$, *4* – $|N^{\text{ion}}(x) + N_S(x)|$;
а – $N_{ss} = 10^{16}$ м$^{-2}$, *б* – $N_{ss} = 10^{17}$ м$^{-2}$, *в* – $N_{ss} = 10^{18}$ м$^{-2}$

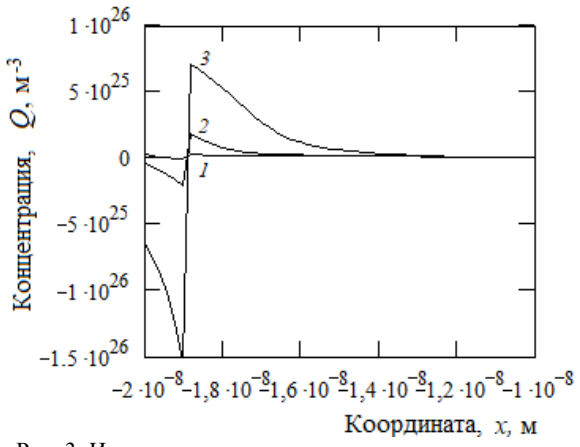


Рис. 3. Изменение концентрации заряженных частиц, обусловленное поверхностным зарядом:
1 – $N_{ss} = 10^{16}$ м$^{-2}$; *2* – $N_{ss} = 10^{17}$ м$^{-2}$; *3* – $N_{ss} = 10^{18}$ м$^{-2}$

В модели заполнения электронами поверхностных состояний [12] эти состояния заряжены отрицательно на материале *n*-типа. В этом случае отрицательный заряд экранирован положительным зарядом доноров (см. рис. 3), так как концентрации электронов меньше концентрации ионизованных доноров (см. рис. 2, *б*, *в*). Увеличение N_{ss} приводит к уменьшению электрического потенциального барьера *n*–*i*–*p*-структуры $\Delta\phi = \phi(-w_n) - \phi(w_p + L)$ и концентрации основных носителей заряда в поверхностной области (см. рис. 2, *б*, *в*). Значения электрического потенциального барьера: $\Delta\phi$ = 1,0737 В при $N_{ss} = 10^{16}$ м$^{-2}$, $\Delta\phi$ = 1,0414 В при $N_{ss} = 10^{17}$ м$^{-2}$, $\Delta\phi$ = 0,7343 В при $N_{ss} = 10^{18}$ м$^{-2}$. Таким образом, существует поверхностная ОПЗ, обусловленная поверхностными состояниями N_{ss}, в которой уменьшается электрический потенциал ϕ(x) (см. рис. 2, *в*).

В неравновесном состоянии возможны осцилляции $\Delta\phi$. Поверхностные или под точечными контактами инжекционные неустойчивости тока в Si наблюдались экспериментально [17].

Поверхностные электронные состояния в запрещённой зоне возникают вследствие нарушения периодичности кристаллической решётки (таммовские состояния, уровни Шокли). Кроме того, дают вклад примеси и другие кристаллические дефекты. Согласно литературным данным [18, 19], плотность поверхностных состояний в кремнии находится в диапазоне 10^{15}–10^{19} м$^{-2}$. Нарушенный поверхностный слой – более широкое понятие, включающее кроме поверхностных состояний область, содержащую механические дефекты. Современные технологические методы обработки поверхности позволяют минимизировать нарушенный слой так, что остаются только поверхностные дефекты в слое толщиной 1–2 нм [20, 21].

Дрейфовая скорость термолизованных носителей заряда является сложной функцией концентрации примесей и напряженности электрического поля [22]. Изменение дрейфовой скорости электронов $v_{dn}(x)$ при различных значениях N_{ss}, рассчитанное с помощью феноменологической формулы [22] и зна-

чений $|N^{ion}(x) + N_s(x)|$, $E(x) = -\,d\phi(x)/dx$ (см. рис. 2, *а–в*), показано на рис. 4. Рассчитанное значение дрейфовой длины электронов в базе $d_n = \tau_0 v_{dn} >> w_p$, где τ_0 – эффективное время жизни [22].

Если плотность поверхностных состояний $N_{ss} >> 10^{16}$ м$^{-2}$, то в поверхностной области напряженность электрического поля отрицательна $E \geq -2{\cdot}10^8$ В/м, поэтому дрейфовая скорость меняет знак на противоположный с 10^5 на -10^5 м/с.

Поверхностные состояния мало влияют на распределение носителей заряда и электрический потенциальный барьер при $N_{ss} \leq 10^{16}$ м$^{-2}$ (см. рис. 2, *а*).

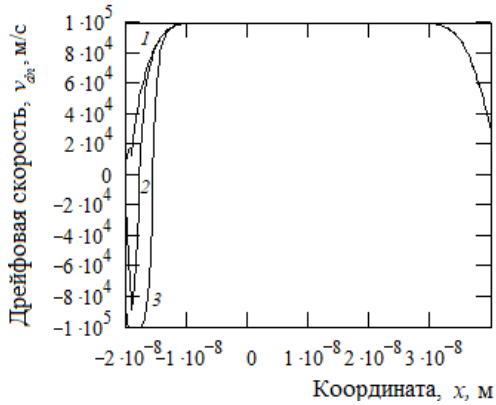


Рис. 4. Дрейфовая скорость термолизованных электронов в объёме наноразмерной *n–i–p*-структуры:
1 – $N_{ss} = 10^{16}$ м$^{-2}$; *2* – $N_{ss} = 10^{17}$ м$^{-2}$; *3* – $N_{ss} = 10^{18}$ м$^{-2}$

Заключение

Результаты численного моделирования показывают, что поверхностные состояния в зоне запрещённых энергий обусловливают существование ОПЗ в поверхностной области *n–i–p*-наноструктуры, в которой происходит уменьшение электрического потенциального барьера и концентрации основных носителей заряда. Заряд поверхностных состояний экранируется зарядом ионизованных примесей в поверхностной ОПЗ. Вследствие этого электрическая проводимость поверхностной области оказывается меньше, чем легированной, а переходное (контактное) сопротивление нелинейно. Дрейфовая скорость носителей заряда меняет знак на противоположный в ОПЗ поверхностного заряда, что препятствует их транспорту. Эти свойства поверхностной области отрицательно влияют на характеристики реальных полупроводниковых структур.

Уменьшение плотности поверхностных состояний до 10^{16} м$^{-2}$ практически устраняет поверхностную ОПЗ. Технологически это достигается пассивацией поверхности пленкой диэлектрика [23].

Увеличение напряжённости электрического поля базы *p–i–n*-структуры является важным фактором для повышения квантового выхода лавинных фотодиодов и детекторов ионизирующего излучения [24, 25]. Дипольный характер распределения электронов и дырок обусловливает резонансное поглощение и излучение электромагнитных волн ТГц диапазона [26]. Результаты расчетов распределения носителей заряда и напряженности электрического поля могут быть использованы для совершенствования этих приборов.

Литература

1. Характеристики *pin*-структуры с дискретно металлизированной поверхностью *i*-области / А.А. Даниленко, А.Д. Иванов, В.Л. Иванов, В.В. Марочкин, Н.И. Михайлов, В.В. Перепеловский // Известия вузов России. Радиоэлектроника. – 2020. – Т. 23, № 1. – С. 41–51.

2. Торина Е.М. Транзисторы для твердотельных СВЧ-переключателей (обзор) / Е.М. Торина, В.Н. Кочемасов, А.Р. Сафин // Известия вузов России. Радиоэлектроника. – 2023. – Т. 26, № 3. – С. 6–31.

3. A study on the high power microwave effects of *PIN* diode limiter based on deep learning algorithm / H. Chen, W. Gao, Y. Zhao, S. Wang, X. Yan, H. Zhou, S. Chen, H. Liu // Nanotechnology. – 2024. – Vol. 35, No. 26. – P. 265202.

4. Shockley W. The theory of *n–p* junctions in semiconductors and *n–p*-junction transistors // Bell Syst. Tech. J. – 1949. – Vol. 28, No. 7. − P. 435−439.

5. Sze S.M. Physics of Semiconductor Devices / S.M. Sze, Y. Li, K.K. Ng. – 4th ed. – Wiley, 2021. – 944 p.

6. Kazuya M. Elementary Semiconductor Device Physics / M. Kazuya, A. Shuhei. – Boca Raton: Taylor & Francis, 2024. – 362 p.

7. Numerical Simulation and Hole Transport Layers Optimization of a Lead Sulfide-Based Solar Cell with a Power Conversion Efficiency of Above 22% / E.L. Meyer, I. Vulindlela, A. Paca, M.A. Agoro, N. Rono // Coatings. – 2025. – Vol. 15, No. 3. – Art. 255.

8. Numerical Solution of Charge Transport Equations in Semiconductors / M. Benaichi, A. Karkri, A. Khriss, R. Zarrouk // EHEI Journal of Science and Technology. – 2024. – Vol. 04, No. 1. – P. 42–47.

9. Alahmadi A.N.M. Numerical Analysis and Design of Hole and Electron Transport Layers in Lead-Free $MASnIBr_2$ Perovskite Solar Cells // Eng. – 2025. – Vol. 6, No. 9. – Art. 222.

10. Моделирование солнечной батареи космического аппарата в процессе эксплуатации на орбите / Р.В. Козлов, Д.Д. Дубов, В.Г. Букреев, А.П. Пичугин // Доклады ТУСУР. – 2025. – Т. 28, № 3. – С. 163–170.

11. Богатов Н.М. Распределение заряда в резком несимметричном равновесном *n–p*-переходе // Экологический вестник научных центров Черноморского экономического сотрудничества. – 2016. – № 3. – С. 12–17.

12. Богатов Н.М. Численное решение уравнения Пуассона в *n–p*-переходе с учетом поверхностных состояний // Экологический вестник научных центров Черноморского экономического сотрудничества. – 2024. – Т. 21. – № 3. – С. 61–69.

13. Влияние поверхностных состояний на электрическое поле *n-p* перехода / Н.М. Богатов, В.С. Володин, Л.Р. Григорьян, М.С. Коваленко // Известия ЮФУ. Технические науки. – 2024. – № 3. – С. 266–275.

14. Моделирование электрического поля кремниевой *n–i–p*-наноструктуры / Н.М. Богатов, В.С. Володин, Л.Р. Григорьян, М.С. Коваленко // Известия ЮФУ. Технические науки. – 2025. – № 5. – С. 123–133.

15. Старосельский В.И. Физика полупроводниковых приборов микроэлектроники: учеб. пособие для вузов. – М.: Юрайт, 2022. – 463 с. [Электронный ресурс]. – URL: https://urait.ru/bcode/509181 (дата обращения: 15.12.2025).

16. Барыкин Д.А. Численное моделирование туннельного эффекта в гетероструктуре нитрида галлия на кремнии / Д.А. Барыкин, К.Ю. Шугуров, А.М. Можаров, И.С. Мухин // Научно-технические ведомости СПбГПУ. Физико-математические науки. – 2024. – Т. 17, № 3. – С. 46–56.

17. Зикриллаев Н.Ф. Автоколебательные процессы в кремнии, проблемы и перспективы исследования и применение их в электронике / Н.Ф. Зикриллаев, М.М. Шоабдурахимова, К.С. Аюпов, Ф.Э. Уракова, О.С. Нематов // Электронная обработка материалов. – 2023. – Т. 59, № 5. – С. 56–71.

18. Яфаров Р.К. Влияние встроенного поверхностного потенциала на ВАХ кремниевых МДП-структур // Микроэлектроника. – 2019. – Т. 48, № 2. – С. 155–159.

19. Александров О.В. Влияние интенсивности ионизирующего облучения на отклик МОП-структур // Физика и техника полупроводников. – 2021. – Т. 55, В. 2. – С. 152–158.

20. Волковский Ю.А. Исследование состояния приповерхностного слоя полированных кремниевых подложек методом рентгеновской рефлектометрии в зависимости от методов их очистки / Ю.А. Волковский, А.Ю. Серегин, М.С. Фоломешкин, П.А. Просекова, М.Д. Павлюк, Ю.В. Писаревский, А.Е. Благова, М.В. Ковальчук // Поверхность. Рентгеновские, синхротронные и нейтронные исследования. – 2021. – № 9. – С. 40–48.

21. Юров В.М. Толщина поверхностного слоя типичных полупроводников / В.М. Юров, Т. Жанабергенов, С.А. Гученко // The scientific heritage. – 2020. – № 43. – С. 20–23.

22. A review of some charge transport properties of silicon / C. Jacoboni, C. Canali, G. Otiaviani, A. Alberigi Quaranta // Solid-State Electronics. – 1977. – Vol. 20, No. 2. – P. 77–89.

23. Black silicon solar cells with interdigitated back-contacts achieve 22.1% efficiency / H. Savin, P. Repo, G. von Gastrow, P. Ortega, E. Calle, M. Garín, R. Alcubilla // Nature Nanotechnol. – 2015. – Vol. 10, No. 7. – P. 624–628.

24. Shobitha G.S. Potentiality of Avalanche Transit Time Devices for Biomedical Applications: A Comprehensive Review / G.S. Shobitha, G.C. Ghivela // Biomedical Materials & Devices. – 2026. – Vol. 4. – P. 4121–1437.

25. Wang B. High-speed Si-Ge avalanche photodiodes / B. Wang, J. Mu // PhotoniX. – 2022. – Vol. 3, No. 8. DOI: https://doi.org/10.1186/s43074-022-00052-6.

26. Андрианов А.В. Генерация терагерцевого излучения в полупроводниках / А.В. Андрианов // Физика твердого тела. – 2023. – Т. 65. – В. 10. – С. 1633–1671.

Богатов Николай Маркович
Д-р физ.-мат. наук, проф., зав. каф. физики
и информационных систем
Кубанского государственного университета (КубГУ)
Ставропольская ул., 149, г. Краснодар, Россия, 350040
ORCID: https://orcid.org/0000-0002-9301-4545
Тел.: +7-903-451-31-06
Эл. почта: bogatov@phys.kubsu.ru

Володин Владимир Сергеевич
Ст. преп. каф. физики и информационных систем КубГУ
Ставропольская ул., 149, г. Краснодар, Россия, 350040
ORCID: https://orcid.org/0000-0003-0639-1105
Тел.: +7-989-820-36-72
Эл. почта: volodinvs1995@mail.ru

Григорьян Леонтий Рустемович
Канд. физ.-мат. наук, доцент каф. физики
и информационных систем КубГУ
Ставропольская ул., 149, г. Краснодар, Россия, 350040
ORCID: https://orcid.org/0009-0002-4744-0356
Тел.: +7-918-168-17-19
Эл. почта: leonmezon@mail.ru

Коваленко Максим Сергеевич
Канд. физ.-мат. наук, доцент каф. физики
и информационных систем КубГУ
Ставропольская ул., 149, г. Краснодар, Россия, 350040
ORCID: https://orcid.org/0009-0006-5402-7366
Тел.: +7-918-469-49-54
Эл. почта: m.s.kovalenko@ya.ru

Поступила в редакцию: 16.03.2026.
Принята к публикации: 01.07.2026.

Bogatov N.M., Volodin V.S.,
Grigoryan L.R., Kovalenko M.S.

Numerical solution of the Poisson equation in a silicon *n–i–p*-nanostructure taking into account surface states

Relevance. Recently, *n–i–p*-nanostructures have become quite widespread in modern electronic devices. Surface electron states influence the distribution of charge carriers within the bulk of the silicon *n–i–p*-nanostructure, with a diffusion distribution of donors and acceptors, thereby altering the properties of semiconductor devices. However, this effect has not been fully studied. ***Purpose of the study.*** Analysis of the influence of surface charge density on the distribution of electron and hole concentrations and their drift velocity in an *n–i–p*-nanostructure. ***Methods.*** To simulate the change in the concentration of charge carriers depending on the electric field potential, a numerical method for solving the Poisson equation was employed, taking surface states into account. ***Scientific novelty.*** Fundamentally new knowledge is obtained about the influence of surface states on the space-charge region of a silicon *n–i–p*-nanostructure. ***Results.*** It is shown that the surface states create a space-charge region in which both the difference in electric potential at the structure boundaries and the concentration of the majority charge carriers decrease. Therefore, the electrical conductivity of the surface region is lower than that in the doped layer. The drift velocity of thermally generated charge carriers reverses in the surface region. Reducing the density of surface states to 10^{16} m^{-2} virtually eliminates these effects. ***Practical significance.*** The obtained results can be used in semiconductor nanoelectronics to improve existing devices and develop new ones.

Keywords: *n–i–p*-nanostructure, electric field potential, surface states, electrons, holes, silicon, drift velocity.
DOI: 10.21293/1818-0442-2026-29-2-93-98

References

1. Danilenko A.A., A.A. Ivanov A.A., Ivanov V.L., Marochkin V.V., Mikhailov N.I., Perepelovskiy V.V. [The Characteristics of the *pin*-Structure with a Discrete Metallic Surface i-Region]. *Journal of the Russian Universities. Radioelectronics*, 2020, vol. 23, no. 1, pp. 41–51 (in Russ.).

2. Torina E.M., Kochemasov V.N., Safin A.R. [Transistors for Solid-State Microwave Switches (A Review)]. *Journal of the Russian Universities. Radioelectronics*, 2023, vol. 26, no. 3, pp. 6–31 (in Russ.).

3. Chen H., Gao W., Zhao Y., Wang S., Yan X., Zhou H., Chen S., Liu H. A study on the high power microwave effects of *PIN* diode limiter based on deep learning algorithm. *Nanotechnology*, 2024, vol. 35, no. 26, pp. 265202.

4. Shockley W. The theory of *n–p* junctions in semiconductors and *n–p*-junction transistors. *Bell Syst. Tech. J.*, 1949, vol. 28, no. 7, pp. 435−439.

5. Sze S.M., Li Y., Ng K.K. Physics of Semiconductor Devices. 4th ed., Wiley, 2021, 944 p.

6. Kazuya M., Shuhei A. Elementary Semiconductor Device Physics, Boca Raton, Taylor & Francis, 2024, 362 p.

7. Meyer E.L., Vulindlela I., Paca A., Agoro M.A., Rono N. Numerical Simulation and Hole Transport Layers Optimization of a Lead Sulfide-Based Solar Cell with a Power Conversion Efficiency of Above 22%. *Coatings*, 2025, vol. 15, no. 3, art. 255.

8. Benaichi M., Karkri A., Khriss A., Zarrouk R. Numerical Solution of Charge Transport Equations in Semiconductors. *EHEI Journal of Science and Technology*, 2024, vol. 4, no. 1, pp. 42–47.

9. Alahmadi A.N.M. Numerical Analysis and Design of Hole and Electron Transport Layers in Lead-Free $MASnIBr_2$ Perovskite Solar Cells. *Eng*, 2025, vol. 6, no. 9, art. 222.

10. Kozlov R.V., Dubov D.D., Bukreev V.G., Pichugin A.P. [Modeling a spacecraft's solar array during on-orbit operation]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2025, vol. 28, no. 3, pp. 163–170 (in Russ.).

11. Bogatov N.M. [The charge distribution in sharp asymmetrical equilibrium *n–p*-junction]. *Ecological Bulletin of Research Centers of the Black Sea Economic Cooperation*, 2016, no. 3, pp. 12–17 (in Russ.).

12. Bogatov N.M. [Numerical solution of the Poisson equation in the n-p junction taking into account surface states] *Ecological Bulletin of Research Centers of the Black Sea Economic Cooperation*, 2024, vol. 21, no. 3, pp. 61–69 (in Russ.).

13. Bogatov N.M., Volodin V.S., Grigoryan L.R., Kovalenko M.S. [Influence of surface states on the electric field of the n-p junction]. *Izvestiya SFedU. Engineering Sciences*, 2024, no. 3, pp. 266–275 (in Russ.).

14. Bogatov N.M., Volodin V.S., Grigoryan L.R., Kovalenko M.S. [Modeling the electric field of a silicon *n–i–p*-nanostructure]. *Izvestiya SFedU. Engineering Sciences*, 2025, no. 5, pp. 123–133 (in Russ.).

15. Staroselsky V.I. *Fizika poluprovodnikovykh priborov mikroelektroniki* [Physics of semiconductor devices of microelectronics: a textbook for universities]. M.: Yurait, 2022, 463 p. Available: https://urait.ru/bcode/509181 (accessed: 15 December 2025) (in Russ.).

16. Barykin D.A., Shugurov K.Yu., Mozharov A.M., Mukhin I.S. [Numerical simulation of the tunnel effect in the gallium nitride heterostructure on silicon]. *St. Petersburg State Polytechnical University Journal. Physics and Mathematics*, 2024, vol. 17, no. 3, pp. 46–56 (in Russ.).

17. Zikrillaev N.F., Shoabdurakhimova M.M., Ayupov K.S., Urakova F.E., Nematov O.S. *Avtokolebatel'nyye protsessy v kremnii, problemy i perspektivy issledovaniya i primeneniye ikh v elektronike* [Self-oscillating processes in silicon: problems and prospects of research and their application in electronics]. *Elektronnaya Obrabotka Materialov/Electronic Processing of Materials*, 2023, vol. 59, no. 5, pp. 56–71 (in Russ.).

18. Yafarov R.K. Effect of the built-in surface potential on the *I–V* characteristics of silicon MIS-structures. *Russian microelectronics*, 2019, vol. 48, pp. 127–130.

19. Aleksandrov O.V. The Effect of the Ionizing Radiation Intensity on the Response of MOS-Structures. *Semiconductors*, 2021, vol. 55, no. 2, pp. 207–213.

20. Volkovsky Yu.A., Seregin A.Yu., Folomeshkin M.S., Prosekov P. A., Pavlyuk M.D., Pisarevsky Yu.V., Blagov A.E., Kovalchuk M.V. X-Ray Reflectometry Study of the State of the Surface Layer of Polished Silicon Substrates Depending on the Methods of Their Cleaning. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*, 2021, vol. 15, no. 5, pp. 927–933.

21. Yurov V., Zhanabergenov T., Guchenko S. [Surface layer typical semiconductors]. *The scientific heritage*, 2020, no. 43, pp. 20–23 (in Russ.).

22. Jacoboni C., Canali C., Otiaviani G., Alberigi Quaranta A. A review of some charge transport properties of silicon. *Solid-State Electronics*, 1977, vol. 20, no. 2, pp. 77– 89.

23. Savin H., Repo P., von Gastrow G., Ortega P., Calle E., Garín M., Alcubilla R. Black silicon solar cells with interdigitated back-contacts achieve 22.1% efficiency. *Nature Nanotechnol*, 2015, vol. 10, no. 7, pp. 624–628.

24. Shobitha G.S., Ghivela G.C. Potentiality of Avalanche Transit Time Devices for Biomedical Applications: A Comprehensive Review. *Biomedical Materials & Devices*, 2026, vol. 4, pp. 4121–1437.

25. Wang B., Mu J. [High-speed Si-Ge avalanche photodiodes]. *PhotoniX*, 2022, vol. 3, no. 8, https://doi.org/10.1186/s43074-022-00052-6.

26. Andrianov A.V. Generation of terahertz radiation in semiconductors. *Physics of the Solid State*, 2023, vol. 65, no. 10, pp. 1563–1598.

Nikolay M. Bogatov
Doctor of Physical and Mathematical Sciences, Professor,
Head of the Department of Physics and Information Systems,
Kuban State University
149, Stavropolskaya St., Krasnodar, Russia, 350040
ORCID: https://orcid.org/0000-0002-9301-4545
Phone: +7-903-451-31-06
Email: bogatov@phys.kubsu.ru

Vladimir S. Volodin
Senior Lecturer, Department Physics and Information Systems, Kuban State University
149, Stavropolskaya St., Krasnodar, Russia, 350040
ORCID: https://orcid.org/0000-0003-0639-1105
Phone: +7-989-820-36-72
Email: volodinvs1995@mail.ru

Leonty R. Grigoryan
Ph.D. (Phys. and Mathematics), Assoc. Prof.,Dept. Physics and Information Systems, Kuban State University
149, Stavropolskaya St., Krasnodar, Russia, 350040
ORCID: https://orcid.org/0009-0002-4744-0356
Phone: +7-918-168-17-19
Email: leonmezon@mail.ru

Maksim S. Kovalenko
Ph.D. (Phys. and Mathematics), Assoc. Prof., Dept. Physics and Information Systems, Kuban State University
149, Stavropolskaya St., Krasnodar, Russia, 350040
ORCID: https://orcid.org/0009-0006-5402-7366
Phone: +7-918-469-49-54
Email: m.s.kovalenko@ya.ru

Received: 16.03.2026.
Accepted: 01.07.2026.

УДК 004.94
Шифр специальности ВАК: 2.3.1, 2.3.3

А.О. Сапунов, М.М. Немирович-Данченко

Сравнительный анализ применения информационного моделирования в отечественной и зарубежной практике управления данными жизненного цикла объекта

Актуальность. Сопоставление международной практики информационного моделирования и ее российской институциональной реализации осложняется сохраняющейся в профессиональной и научной среде терминологической неопределенностью, что создает барьер для адаптации международного опыта и интеграции отечественных проектов в глобальные процессы. ***Цель исследования.*** Разработка рекомендаций по интеграции эффективных международных решений в российскую систему ТИМ (технологии информационного моделирования), а также формирование структуры автоматизированного комплекса для обеспечения функциональной совместимости ТИМ и BIM (Building Information Modeling). ***Методы.*** Исследование основано на сравнительном анализе по системе критериев, охватывающей нормативно-правовое регулирование рассматриваемых практик, их технологический стек, а также организационно-процессуальную модель. ***Научная новизна.*** Впервые проведено комплексное сравнение зарубежной и отечественной практик управления данными жизненного цикла. На основе этого сравнения определены принципы совмещения технологий BIM и ТИМ, а также обоснована необходимость их интеграции для развития строительной отрасли. ***Результаты.*** Доказано, что ТИМ – это самостоятельная практика информационного моделирования (ИМ), а не прямой синоним BIM. Разработаны практические рекомендации по адаптации эффективных элементов международной практики для повышения конкурентоспособности российской строительной отрасли. Сформирована структура автоматизированного комплекса, обеспечивающая совместимость ТИМ и BIM. ***Практическая значимость.*** Полученные рекомендации и структура автоматизированного комплекса могут стать основой для дальнейших исследований по выбору и объединению лучших практик информационного моделирования в российских проектах.

Ключевые слова: информационное моделирование, цифровая модель, проектирование, строительство, BIM, ТИМ, CAD-системы.
DOI: 10.21293/1818-0442-2026-29-2-99-106

Введение

Проектирование строительных объектов – это процесс разработки технической, архитектурной и инженерной документации, необходимой для возведения, реконструкции или капитального ремонта зданий и сооружений. К этапам проектирования принято относить: предпроектные работы, разработку концепции, проектирование, подготовку рабочей документации, авторский надзор и сопровождение строительства [1]. На каждом этапе требуется формирование, интеграция и управление цифровыми моделями и данными. В РФ основным подходом к проектированию длительное время являлся «Building Information Modeling» (BIM), однако с внесением изменений в Градостроительный кодекс в 2018 г. в РФ был официально введен термин «технологии информационного моделирования» (ТИМ). Согласно международному стандарту управления информацией ISO 19650, BIM – подход, основывающийся на использовании общей цифровой модели объекта для поддержки проектных, строительных и эксплуатационных процессов, формирующих достоверную основу для принятия решений. Согласно национальному стандарту Российской Федерации ГОСТ Р 10.00.00.01 от 2025 г., ТИМ – совокупность систематизированных знаний, методов и средств ИМ, обеспечивающих возможность достижения целей ИМ на всех этапах ее жизненного цикла, а также соответствующих вкладов в их достижение на отдельных стадиях. В современной научной и профессио-нальной литературе отсутствует единое мнение относительно соотношения BIM и ТИМ. Существующие разночтения затрудняют практическое внедрение технологий и адаптацию международного опыта. В связи с этим целью работы является выявление ключевых различий, влияющих на управление данными жизненного цикла, разработка рекомендаций по интеграции эффективных международных решений в российскую систему ТИМ, а также формирование структуры автоматизированного комплекса, которая позволит совместить лучшие функциональные возможности ТИМ и BIM.

Анализ подходов к соотношению BIM и ТИМ в научной литературе

Проведенный анализ литературных источников позволяет не просто констатировать наличие разночтений, но и выявить их системные причины, обусловленные разным уровнем рассмотрения данных понятий. Сформировавшиеся точки зрения могут быть структурированы в три основные группы. Первая группа трактует ТИМ как терминологическую адаптацию BIM. Ряд исследователей [2–5] рассматривают ТИМ исключительно как терминологическую адаптацию международной концепции BIM, обусловленную требованиями нормативно-правовой базы РФ. В рамках данного подхода подчеркивается общность базовых принципов, технологий и инструментария, а различия сводятся к лингвистическим и юридическим аспектам. Крайним проявлением этой позиции является упрощение BIM до инструментов трёхмерного моделирования [6] при одновременном

утверждении о тождественности BIM и ТИМ как понятий. Вторая группа акцентирует внимание на содержательных и функциональных различиях. В качестве критериев различия называются: структурная организация модели (единая в BIM и разделенная на цифровую и информационную составляющие в ТИМ [7]), степень комплексности подхода – ТИМ трактуется как более широкая методология по сравнению с BIM [8]. Данная группа смещает фокус сравнения с вопросов наименования к анализу архитектуры данных и процессов. Третья группа рассматривает BIM и ТИМ как самостоятельные, но сопряженные системы.

Позиция, представленная в работах [9, 10], исходит из того, что BIM и ТИМ эволюционировали в разных государственных средах, что привело к формированию самостоятельных, хотя и технологически сопряженных систем. В данном контексте ТИМ рассматривается не как синоним или адаптация, а как развитие и закрепление принципов BIM применительно к специфике российской нормативной базы, системы госзаказа и требований к отчетности [10]. Различия здесь видятся на уровне экосистемы подходов. Выявленный спектр мнений наглядно отображает глубину проблемы. Разногласия существуют не только относительно выводов, но и относительно самого предмета сравнения: сравниваются ли термины, технологии, методологии или целые системы. Именно эта неопределенность предмета сравнения и обусловливает невозможность достижения общего мнения в рамках сопоставления определений. Следовательно, для сравнения BIM и ТИМ требуется разработка структурированной системы критериев, позволяющей сравнивать их в контексте нормативного, технологического и организационного аспектов.

Методологическая основа сравнительного анализа: система критериев

Как показал анализ литературных источников, основным препятствием для корректного сопоставления BIM и ТИМ является неопределённость уровня и предмета сравнения. Для её преодоления и перехода от терминологических дискуссий к содержательному анализу необходима чёткая аналитическая рамка. В качестве такой рамки в данном исследовании используется система критериев, обеспечивающая комплексную оценку сравниваемых подходов как практик, реализуемых в конкретных технологических средах. Выбор критериев обусловлен необходимостью охватить три взаимосвязанных аспекта, определяющих жизненный цикл информационной модели от регламентации требований до практической реализации проекта [11].

Нормативное регулирование. Состоит из следующих параметров: происхождение стандартов, модуль управления данными, роль государственного регулирования. Данный критерий позволяет оценить, каким образом стандарты и законодательные требования формируют правила работы с моделью строительного объекта и задают границы применения технологий. Анализ по этому аспекту позволяет выявить степень гибкости системы и роль государственного регулирования, что критически важно для проектов, связанных с соблюдением отраслевых стандартов безопасности и проходящих государственную экспертизу.

Технологический стек. Состоит из следующих параметров: стратегическая ориентация, принцип разработки, архитектура и сервисная модель. Этот критерий фокусируется на инструментальной реализации. Он позволяет сравнить программное обеспечение, его функциональность и способность обеспечивать сквозную работу с данными, что напрямую влияет на возможности проектировщиков.

Организационно-процессуальная модель. Состоит из следующих параметров: модель взаимодействия, распределение ответственности и конечный результат. Критерий направлен на анализ того, как выстроены коммуникации и как распределены роли и ответственность между участниками проекта. Он отражает культуру подхода к формированию модели и точки принятия ключевых решений.

Последовательное применение данных критериев формирует объективную основу для сопоставления BIM и ТИМ, переходя от анализа терминов к анализу их внутренних процессов. Такой подход минимизирует субъективность и позволяет выявить не только различия, но и схожести в различных аспектах практик, обеспечивая полноту и баланс анализа.

Сравнение по критерию «нормативно-регламентирующая основа»

В рамках сравнения по первому критерию была определена и проанализирована нормативная база, регламентирующая каждую из практик. Для международной практики BIM в анализ были включены как актуальные глобальные стандарты, так и один из исторически значимых национальных документов, что позволяет проследить возникновение подходов и их потенциальную связь с российской практикой.

Ключевыми документами, определяющими нормативную основу BIM, являются:

– ISO 19650 – глобальный стандарт управления информацией на протяжении жизненного цикла строительного объекта;

– ISO 16739-1:2024 – глобальный стандарт, определяющий открытую схему данных для обмена информацией в области строительства;

– PAS 1192-2:2013 – Британская спецификация, регламентирующая принципы и требования к информационным моделям зданий. В настоящий момент спецификация заменена на ISO 19650-2:2018 в рамках глобального перехода на международные стандарты [12].

Нормативную базу ТИМ в РФ составляют:

– СП 333.1325800.2020 – свод правил, определяющий правила формирования и управления информационной моделью на различных стадиях жизненного цикла объекта строительства;

– СП 328.1325800.2020 – свод правил, определяющий правила описания компонентов информационной модели;

– Постановление правительства РФ № 331 от 5 марта 2022 г. – документ, вводящий обязательное

применение ТИМ для объектов государственного заказа.

Для обоих подходов характерен ряд общих принципов, закрепленных в нормативных документах: *чёткие требования к процессу, ориентация на открытые форматы обмена данными (стандарт Industry Foundation Classes), а также обязательность для государственного сектора* **(**как историческая практика Великобритании, так и практика РФ делают применение технологий обязательным для проектов с государственным финансированием).

Однако, как показано в табл. 1, по *ключевым параметрам* нормативные системы демонстрируют *принципиальные различия*, определяющие их институциональную природу.

Из табл. 1 видно, что несмотря на общую направленность на цифровизацию строительства, нормативные системы BIM и ТИМ различаются степенью унификации, гибкостью и ролью государственного регулирования. Потенциально, это может оказывать влияние на скорость внедрения ТИМ, особенно в проектах с международным участием, где требуется согласование стандартов.

Сравнение по критерию «Технологический стек»

Вторым критерием анализа является технологический стек – набор программных решений, обеспечивающих практическую реализацию принципов ИМ. Для обеспечения репрезентативности сравнения были отобраны по пять наиболее распространённых программных продуктов, представляющих международный рынок BIM-решений и отечественный рынок ПО для ТИМ. Данный выбор охватывает инструменты для ключевых процессов: от создания моделей и чертежей до управления данными и совместной работы над проектом.

Технологический стек в BIM представлен специализированными САПР-платформами, ориентированными на сложные проекты: «AutoCAD», «Revit», «ArchiCAD», «Bentley Architecture» (входит в состав широкого комплекса «Bentley Systems»). Также к технологическому стеку относится «Autodesk BIM 360» – облачная платформа для управления строительными проектами, а также объединения команды и данных в единое рабочее пространство.

По части программного обеспечения ТИМ опирается на международные BIM-решения, однако акцент делается на адаптацию ПО к требованиям российских стандартов. Наряду с этим развиваются отечественные САПР, такие как: «NanoCAD», «Renga», «Model Studio CS» (комплексная система ИМ, предназначенная для разработки моделей на базе «NanoCAD») и Компас 3D. Также важно упомянуть «CADLib» – семейство программных продуктов, которые отвечают за хранение данных, совместную работу над проектами, проверку моделей, а также обеспечивают поддержку жизненного цикла объекта.

Ключевая схожесть ПО в BIM и ТИМ – это их назначение. Все перечисленные инструменты применяются для создания и управления цифровыми моделями зданий и сооружений, включая архитектурные, конструкторские и инженерные решения.

Следующая схожесть заключается в функциональности продуктов, их интегрируемости и обеспечении совместной работы. Как в BIM, так и в ТИМ присутствует поддержка 2D-чертежей и 3D-моделей. Также большинство САПР-платформ может быть интегрировано в среду общих данных, что, в свою очередь, обеспечивает централизованное хранение данных, поддержку жизненного цикла проекта и возможность одновременно работать над одним проектом множеству специалистов.

Тем не менее, несмотря на общую функциональную направленность, анализ по заданным параметрам выявляет системные различия в архитектуре и стратегии развития технологических стеков (табл. 2).

Таблица 2 подтверждает выявленную тенденцию: технологический стек BIM эволюционирует как *интегральная часть глобальной цифровой экосистемы строительной отрасли*, где ключевыми ценностями являются масштабируемость, совместимость и облачное взаимодействие. В свою очередь, технологический стек ТИМ формируется как адаптированная под РФ *инструментальная среда*, где приоритет отдается гарантированному соблюдению норм, работе в условиях локальной инфраструктуры и поддержке типовых проектных решений. Данное различие создаёт технологический барьер для бесшовного международного взаимодействия, требуя дополнительных усилий по согласованию данных и процессов.

Таблица 1

Различия в нормативном регулировании

Параметр критерия	Практика BIM	Практика ТИМ	Различие
Происхождение стандартов	**Глобальные** стандарты, носящие **рекомендательный** характер и допускающие адаптацию	**Обязательные национальные стандарты и строительные нормы, регламентированные законодательством**	Гибкость и адаптивность в BIM, жёсткая регламентация в ТИМ
Модель управления данными	**Децентрализованная** модель через среду общих данных с **распределением** ответственности между участниками	**Централизованная модель** с контролем через государственные информационные системы	Сетевое взаимодействие в BIM, иерархический контроль в ТИМ
Роль государственного регулирования	**Государство** выступает как **заказчик**, продвигающий стандарты через технические требования	Государство является **законодательным инициатором, регулятором и главным контролёром** процесса	Государство как заказчик в BIM, государство как административный регулятор в ТИМ

Таблица 2

Различия в технологическом стеке

Параметр критерия	Практика BIM	Практика ТИМ	Различие
Стратегическая ориентация	**Глобальная экосистема**, ориентированная на международные проекты любого масштаба и сложности	**Национально-ориентированная среда**, сфокусированная на типовых задачах и специфике российского проектирования	Универсальность в BIM, национальная специализация в ТИМ
Принцип разработки и адаптивность	Разрабатывается под **глобальные рыночные практики**, предполагает **высокую гибкость**	Разрабатывается с **учётом национальных стандартов**, что ограничивает гибкость, но гарантирует нормативное соответствие	Гибкость конфигурации в BIM, нормативная гарантия в ТИМ
Архитектура и сервисная модель	Стандартом является **облачно-ориентированная экосистема**	Преобладает **локальная или сетевая установка** с акцентом на интеграцию в корпоративную структуру	Облачный сервис в BIM, локальное ПО в ТИМ

Сравнение по критерию «организационно-процессуальная модель»

Третьим критерием анализа выступает организационно-процессуальная модель, которая определяет структуру взаимодействия, распределение ролей и механизмы принятия решений в рамках практик ИМ. Эффективность внедрения технологий ИМ в значительной степени зависит от отлаженности данной модели.

В **международной практике BIM** закрепилось три роли: BIM-моделлер, BIM-координатор и BIM-менеджер [13]. BIM-моделлер отвечает за создание информационной модели, подготовку чертежей и документации на основе технического задания. BIM-координатор осуществляет координацию между дисциплинами и контроль модели. BIM-менеджер управляет стандартами проекта, организует работу участников и обучает их.

В системе ТИМ профессиональные роли регламентированы на национальном уровне Приказом Минтруда РФ от 16 ноября 2020 г. № 787н и также образуют трёхуровневую структуру, состоящую из таких ролей, как ТИМ-проектировщик, ТИМ-координатор и ТИМ-менеджер. *ТИМ-проектировщик* анализирует исходные данные и формирует модель. *ТИМ-координатор* выполняет анализ ТЗ и исходных данных, определяет необходимые ресурсы, составляет и согласовывает график проекта, а также контролирует его выполнение. *ТИМ-менеджер* управляет коммуникацией, документооборотом и выполняет сдачу итоговой модели.

Ролевые модели показывают два ключевых совпадения, таких как универсальная трёхуровневая иерархия и схожесть функций участников процесса. Однако за формальным сходством ролей скрываются принципиальные различия в реализации процессуальной модели, представленные в табл. 3.

Таблица 3 демонстрирует, как выявленные ранее нормативные и технологические различия проявляются в практике управления. Модель BIM формализуется как *сетевая система*, где цифровая модель является живым инструментом совместной работы и основой для оперативного принятия решений. Модель ТИМ формализуется как *иерархическая, документоориентированная процедура*, где цифровая модель, прежде всего, выступает в качестве строго регламентированного объекта контроля и основания для административного утверждения.

Таблица 3

Различия в организационно-процессуальной модели

Параметр критерия	Практика BIM	Практика ТИМ	**Различие**
Модель взаимодействия	**Совместная сетевая модель, работающая на облачных платформах и позволяющая участникам одновременно взаимодействовать с общими данными**	**Последовательная иерархическая модель, построенная на основе локальных информационных систем, которая предполагает поэтапное создание, проверку и утверждение как самой модели, так и сопутствующей документации**	Параллельная интеграция в BIM, последовательная верификация в ТИМ
Распределение ответственности и координация	Ответственность за данные несут исполнители под контролем BIM-менеджера	Ответственность смещена к **заказчику и надзорным органам**. Ключевые решения принимаются **заказчиком по итогам проверки** подготовленного пакета	Распределённая ответственность в BIM, централизованное утверждение в ТИМ
Конечный результат	Подготовка единой информационной модели	Подготовка **отчётного пакета документов** для государственных органов	Ориентация на целостную модель в BIM, ориентация на выпускаемую документацию в ТИМ

Это различие определяет не только операционную культуру, но и итоговую эффективность использования информационной модели в достижении стратегических целей проекта.

Опыт практического применения ТИМ и BIM

Выявленные различия между практиками находят своё отражение в наблюдаемой картине их применения. Рассмотрение реальных проектов, в которых использовалась та или иная практика, позволяет соотнести теоретические выводы с существующей практикой.

Так, BIM позиционируется как стандартный подход для реализации масштабных и сложных проектов. Примерами таких проектов, где использование BIM освещено в отраслевой литературе, являются:

– Шанхайская башня, расположенная в Китае, проектирование которой велось с использованием программных продуктов «Autodesk», что характерно для глобального технологического стека BIM [14]. Проект иллюстрирует применение BIM в задачах уникального архитектурного и инженерного проектирования (рис. 1).

– Линия метрополитена Crossrail, построенная в Великобритании, является хорошим примером применения BIM [15] для управления крупными государственными программами с жёсткими требованиями к координации и контролю затрат (рис. 2).

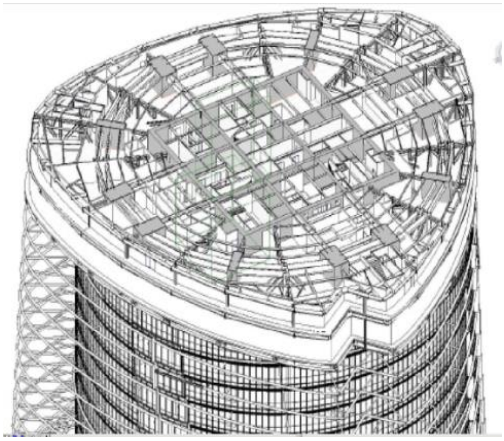


Рис. 1. BIM-модель шанхайской башни

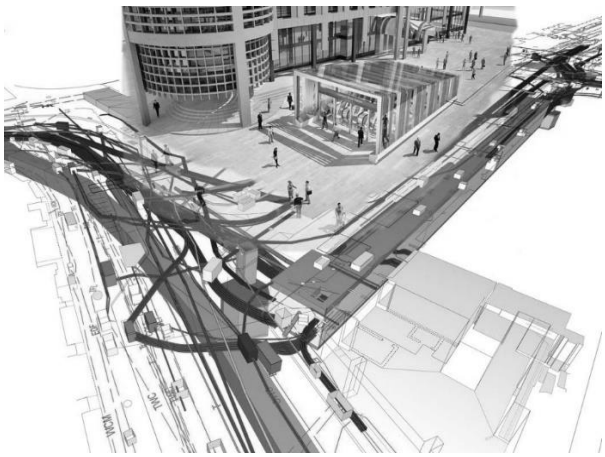


Рис. 2. Визуализация BIM-модели участка Crossrail

Внедрение ТИМ в РФ же, инициированное законодательно, определяет её государственную сферу применения: от типовых строительных проектов до крупных государственных инфраструктур. К примерам проектов с применением ТИМ относятся:

– Бирюлевская линия Московского метрополитена (рис. 3): отражает ориентацию системы на критически важные объекты государственного заказа, где обязательное соблюдение национальных норм и процедур является первостепенным. Согласно официальным источникам, проектирование данной линии проходит с применением технологий ТИМ.

– Дворец бракосочетаний, г. Бобров: типовой проект социальной инфраструктуры, реализованный при помощи ПО «Renga» (рис. 4). Он демонстрирует развитие ТИМ как инструмента стандартизации и оптимизации массового проектирования в рамках национальных нормативных требований.

Практика показывает, что BIM – выбор международных проектов, а ТИМ – национального госзаказа. Причина в разнице приоритетов: глобальные инициативы ценят гибкость, а госзаказ – соответствие нормам. На этом фоне особенно остро встаёт вопрос адаптации зарубежного опыта.



Рис. 3. Бирюлевская линия Московского метрополитена



Рис. 4. Трехмерная модель Дворца бракосочетаний

Формирование структуры автоматизированного комплекса

На основании результатов сравнения подходов к проектированию в зарубежной и отечественной практиках составлен перечень рекомендаций по интегрированию международного опыта в ТИМ, в котором предлагается рассмотреть возможность следующих действий:

1. Введение допуска на использование международных стандартов в проектах, подразумевающих работу с иностранными организациями, что обеспечит гибкость в работе над межнациональными проектами.

2. Внедрение облачных технологий работы с моделью, что позволит вести работу над проектом не только последовательно, но и параллельно. Это может ускорить введение проекта в эксплуатацию, а также обеспечить консолидированный доступ к данным для всех участников процесса.

3. Расширение перечня ролей для участников процесса проектирования. В международной практике аналогично укрепились три роли, однако они не закреплены законодательно, что дает возможность их изменения. В случае увеличения количества ролей можно более сбалансированно распределить обязанности между ними и увеличить их производительность.

Представленный перечень рекомендаций послужил основой для концептуальной структуры автоматизированного комплекса информационного моделирования, состоящего из трех взаимосвязанных подсистем:

1. Подсистема согласования данных – автоматизированный модуль конвертации данных между CAD-платформами, использующими международные спецификации, и платформами, в основе которых лежат структуры отечественного классификатора строительной информации ТИМ. Данный модуль необходим для того, чтобы проект мог пройти государственную экспертизу. Базисом модуля должна стать возможность решения задачи семантического сопоставления атрибутивных параметров модели строительства.

2. Подсистема управления процессами на базе облачных технологий – модуль управления средой общих данных. Данный модуль должен предоставлять непрерывный асинхронный доступ к данным проекта, чтобы обеспечить параллельное проектирование в режиме реального времени. Благодаря ему, независимо от типа архитектуры CAD-платформы и других средств проектирования, будет возможно организовать работу над проектом с привлечением специалистов из разных стран.

3. Подсистема динамического распределения ролей – модуль администрирования прав доступа и управления задачами. Должен давать возможность управлять ролями членов команды и распределять задачи в зависимости от компетенций участников команды. Модуль позволит адаптировать любое ролевое устройство под специфику межнационального взаимодействия.

Схематическое представление структуры представлено на рис. 5.

Согласно данной схеме, представленный комплекс объединяет разрозненные ранее процессы в единый автоматизированный цикл, а функциональная связка подсистем позволит эффективно применять международный опыт проектирования внутри отечественной экосистемы ТИМ.

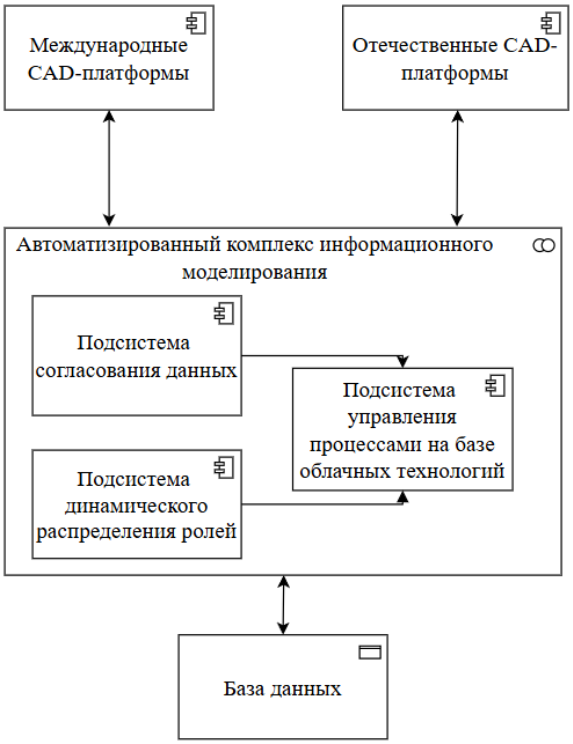


Рис. 5. Структурная архитектура автоматизированного комплекса информационного моделирования

Заключение

Анализ показал важные различия между международной практикой BIM и российской системой ТИМ, что указывает на их разную институциональную природу.

ТИМ нельзя считать точным аналогом BIM. Это самостоятельная система, которая строится на строгих правилах, подходит для проектов, финансируемых государством, где главное – соответствовать нормам. BIM же – это мировая методология, основанная на общих стандартах, гибкости и сотрудничестве, которая подходит для принципиально новых и международных проектов.

На основании результатов сравнения была сформирована структура автоматизированного комплекса информационного моделирования, обеспечивающего функциональную совместимость BIM и ТИМ. Структура включает в себя три подсистемы:

1. Подсистему согласования данных.

2. Подсистему управления процессами на базе облачных технологий.

3. Подсистему динамического распределения ролей.

Полученные результаты могут стать основой для дальнейших исследований по выбору и объединению лучших практик информационного моделирования в российских проектах.

Литература

1. Соловьева Е.В. Основные этапы внедрения технологии информационного моделирования (BIM) в строительных организациях / Е.В. Соловьева, М.А. Сельвиан // Электронный сетевой политематический журнал «Научные труды КубГТУ». – 2016. – № 11. – С. 110–119.

2. Забегина А.Р. Роль и место технологий моделирования (ТИМ) в современном архитектурном проектировании // Вестник Моск. информационно-технологического ун-та – Моск. архитектурно-строительного ин-та. – 2022. – № 4. – С. 42–48.

3. Хугаев А.З. Реинжиниринг бизнес-процессов в проектно-изыскательских организациях, связанных с цифровизацией отрасли и внедрением BIM-моделирования / А.З. Хугаев, Е.Н. Лейман // Экономика: вчера, сегодня, завтра. – 2024. – Т. 14, № 8А. – С. 132–142.

4. Тен М.Г. Решение педагогических проблем обучения студентов строительного вуза в условиях перехода на BIM (ТИМ) технологии / М.Г. Тен, С.В. Максимова // Инновационные технологии в инженерной графике: проблемы и перспективы: сб. тр. междунар. науч.-практ. конф., Новосибирск, 24 апреля 2024 г. – Брест: БрГТУ, – 2024. – С. 121–123.

5. Кирилюк С.А. Анализ текущих тенденций развития технологий информационного моделирования ТИМ // Вестник науки. – 2024. – Т. 2, № 4 (73). – С. 646–652.

6. Бикбова З.М. Применение BIM технологий // Цифровой регион: опыт, компетенции, проекты: сб. статей VI Междунар. науч.-практ. конф., посвященной 25-летнему юбилею Инженерно-экономического института, проводимой в рамках 10-летия науки и технологий в России. – Брянск: БГИТУ, 2024. – С. 101–104.

7. Гусарова А.А. План реализации проекта с ТИМ // Международный журнал гуманитарных и естественных наук. – 2023. – № 5-2 (80). – С. 79–83.

8. Шаповалов А.Г. Правовое обеспечение безопасности в сфере информационного моделирования в строительстве (BIM) // Образование и право. – 2024. – № 12. – С. 191–193.

9. Пименов С.И. Анализ современных программных комплексов для виртуального строительства (4D-моделирования) // Научный журнал строительства и архитектуры. – 2022. – № 3 (67). – С. 92–104.

10. Скуркан Е. BIM и ТИМ проектирование в РФ / Е. Скуркан, Д.А. Бредихин, Л.В. Чайковская // Технологии, машины, и оборудование для проектирования, строительства объектов АПК: сб. науч. статей 2-й Междунар. науч.-техн. конф. молодых ученых, аспирантов, магистров и бакалавров, 14 марта 2024 года. – Курск: Университетская книга, 2024. – С. 294–296.

11. Сапунов А.О. Автоматизация BIM-процессов на всех стадиях жизненного цикла объектов капитального строительства с помощью модуля актуализации сведений в информационных моделях // Технологии обустройства нефтяных, газовых и газоконденсатных месторождений: сб. тезисов VII науч.-техн. конф., Томск, 24–25 сентября 2024 г. – Томск: АО «Томский научно-исследовательский и проектный институт нефти и газа», 2024. – С. 235–236.

12. Malla V. Perceptions of built-environment professionnals on using ISO 19650 standards for information management / V. Malla, M. Tummalapudi, V.S.K. Delhi // Journal of Legal Affairs and Dispute Resolution in Engineering and Construction. – 2024. – Vol. 16, No. 1. DOI: 10.1061/JLADAH.LADR-1076.

13. Gatto C. et al. Alignment between standards and job market demand for BIM careers // Buildings. – 2025. – Vol. 15, No. 13. – P. 2323.

14. Jin S. Integrating BIM to enhance stakeholder collaboration and economic efficiency in architecture and construction projects // E3S Web of Conferences – EDP Sciences. – 2024. – Vol. 565. – P. 03014.

15. Hussain O.A.I. et al. Minimizing cost overrun in rail projects through 5D-BIM: a conceptual governance framework // Buildings. – 2024. – Vol. 14, No. 2. – P. 478.

Сапунов Анатолий Олегович
Аспирант каф. комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС) Томского государственного ун-та систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050;
ТИМ-мастер ООО «РН-Проектирования Добыча»
Мира пр-т, 72, г. Томск, Россия, 634027
ORCID: 0009-0009-1686-7876
Тел.: +7-913-119-12-27
Эл. почта: sanovta@yandex.ru

Немирович-Данченко Михаил Михайлович
Д-р физ.-мат. наук, проф. каф. КИБЭВС ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
Тел.: +7-913-877-90-03
Эл. почта: nmm@fb.tusur.ru

Поступила в редакцию: 30.04.2026.
Принята к публикации: 21.07.2026.

Sapunov A.O., Nemirovich-Danchenko M.M

Comparative analysis of the information modeling application in domestic and international practice of data lifecycle management

Relevance*.** The comparison of international information modeling practices with their institutional implementation in Russia is complicated by persistent terminological ambiguity within the professional and academic communities. This ambiguity creates a barrier to the effective adaptation of international experience and the integration of domestic projects into global processes. ***Purpose of the study*.** The study aims to develop recommendations for integrating effective international solutions into the Russian TIM (Information Modeling Technologies) framework and to develop the structure of an automated system ensuring the functional interoperability of TIM and BIM (Building Information Modeling). ***Methods*.** The analysis is based on a system of criteria covering the legal regulation of the practices under consideration, their technological stack, as well as their organizational and procedural model. ***Scientific novelty. For the first time, a comprehensive comparison of international and domestic practices for lifecycle data management has been conducted. Based on this comparison, principles for integrating BIM and TIM technologies have been identified, and the need for their integration to support the development of the construction industry has been substantiated. ***Results.*** It is concluded that TIM is an independent practice of information modeling, and not a direct synonym for BIM. Practical recommendations have been developed for adapting effective elements of international practice to enhance the competitiveness of the Russian construction industry. The structure of an automated system ensuring interoperability between TIM and BIM has been developed. ***Practical Significance.*** The resulting recommendations and the structure of the automated system can form the basis for further research on the selection and integration of best practices in information modeling in Russian projects.

Keywords: information modeling, digital model, design, construction, BIM, TIM, CAD systems.
DOI: 10.21293/1818-0442-2026-29-2-99-106

References

1. Solovieva E.V., Selvian M.A. [The main stages of the implementation of information modeling (BIM) technology in construction organizations]. *Electronic online polythematic*

*journal Scientific Works of KubSTU,"*2016, no. 11, pp. 110–119 (in Russ.).

2. Zabegina A.R. [The role and place of building information modeling (BIM) in modern architectural design]. *Moscow information and technology university- Moscow architecture and construction institute review*, 2022, no. 4, pp. 42–48 (in Russ.).

3. Khugaev A.Z., Leiman E.N. [Reengineering business processes in design and survey organizations related to industry digitalization and BIM implementation]. *Economics: yesterday, today, tomorrow*, 2024. vol. 14, no. 8A, pp. 132–142 (in Russ.)

4. Ten M.G., Maksimova S.V. *Resheniye pedagogicheskikh problem obucheniya studentov stroitel'nogo vuza v usloviyakh perekhoda na BIM (TIM) tekhnologii* [Solution of pedagogical problems of teaching students of a construction university in the context of the transition to BIM technologies]. Innovatsionnyye tekhnologii v inzhenernoy grafike: problemy i perspektivy [*Innovative technologies in engineering graphics: problems and prospects*] *Proc. of the International scientific and practical*, Novosibirsk, Brest, BrSTU, 2024, pp. 121–123 (in Russ.).

5. Kirilyuk S.A. [Analysis of current trends in development of information modeling technologies TIM]. *Science Bulletin*, 2024, vol. 2, no. 4 (73), pp. 646–652 (in Russ.)

6. Bikbova Z.M. *Primeneniye BIM tekhnologiy* [Application of BIM technologies]. Tsifrovoy region: opyt, kompetentsii, proyekty [*Digital region: experience, competencies, projects*] Proc. of International scientific and practical conference dedicated to the 25th anniversary of the Engineering and Economics Institute, held within the framework of the Decade of Science and Technology in Russia, Bryansk, BGITU, 2024, pp. 101–104 (in Russ.).

7. Gusarova A.A. [Project implementation plan with TIM]. *International Journal of Humanities and Natural Sciences*, 2023, no. 5-2 (80), pp. 79–83 (in Russ.).

8. Shapovalov A.G. [Legal support for security in the field of building information modeling (BIM)]. *Education and Law*, 2024, no. 12, pp. 191–193 (in Russ.).

9. Pimenov S.I. [Analysis of modern software for virtual construction (4D modeling)]. *Russian journal of building construction and architecture*, 2022, no. 3 (67), pp. 92–104 (in Russ.).

10. Skurkan E., Bredikhin D.A., Chaikovskaya L.V. *BIM i TIM proyektirovaniye v RF* [BIM and TIM design in the Russian Federation]. Tekhnologii, mashiny, i oborudovaniye dlya proyektirovaniya, stroitel'stva ob'yektov APK [*Technologies, machines, and equipment for the design and construction of agricultural facilities*]: Proc. of the 2nd International Scientific and Technical Conference of Young Scientists, Postgraduates, Masters, and Bachelors, March 14, 2024, Kursk, "University Book" Publ. house, 2024, pp. 294–296 (in Russ.).

11. Sapunov, A. O. *Avtomatizatsiya BIM-protsessov na vsekh stadiyakh zhiznennogo tsikla ob'yektov kapital'nogo stroitel'stva s pomoshch'yu modulya aktualizatsii svedeniy v informatsionnykh modelyakh* [Automation of BIM processes at all stages of the life cycle of capital construction projects using a module for updating information in information models]. Tekhnologii obustroystva neftyanykh, gazovykh i gazokondensatnykh mestorozhdeniy [*Technologies for the development of oil, gas and gas condensate fields*] *Proc. of the VII scientific and technical* conf., Tomsk, JSC Tomsk Research and Design Institute of Oil and Gas, 2024, pp. 235–236 (in Russ.).

12. Malla V., Tummalapudi M., Delhi V.S.K. Perceptions of built-environment professionals on using ISO 19650 standards for information management. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 2024, vol. 16, no. 1. DOI: 10.1061/JLADAH.LADR-1076.

13. Gatto C. et al. Alignment between standards and job market demand for BIM careers. *Buildings*, 2025, vol. 15, no. 13, pp. 2323.

14. Jin S. Integrating BIM to enhance stakeholder collaboration and economic efficiency in architecture and construction projects. *E3S Web of Conferences – EDP Sciences*, 2024, vol. 565, pp. 03014.

15. Hussain O.A.I. et al. Minimizing cost overrun in rail projects through 5D-BIM: a conceptual governance framework. *Buildings,* 2024, vol. 14, no. 2, pp. 478.

Anatoly O. Sapunov
Postgraduate, Department of Complex Information Security of Computer Systems (KIBEVS), Tomsk University of Control Systems and Radioelectronics (TUSUR)
40, Lenin pr., Tomsk, Russia, 634050
BIM-master, RN-Upstream Design
72, Mira pr., Tomsk, Russia, 634027
Phone: +7-913-119-12-27
Email: sanovta@yandex.ru

Mikhail M. Nemirovich-Danchenko
Doctor of Science in Physics and Mathematics, Professor, Department of KIBEVS TUSUR
40, Lenin pr., Tomsk, Russia, 634050
Phone: +7-913-877-90-03
Email: nmm@fb.tusur.ru

Received: 30.04.2026.
Accepted: 21.07.2026.

УДК 004.932.72
Шифр специальности ВАК: 2.3.8

С.И. Бабаев, Б.В. Костров, Е.Р. Муратов, А.С. Тарасов, В.Ю. Тарасова

Алгоритм уточнения совмещения разнородных изображений на основе ключевых точек

Актуальность. Рассматривается алгоритм уточнения совмещения разнородных изображений. Совмещение изображений осуществляется на основе нейросетевого отбора кандидатных областей. Совмещение разнородных изображений может служить основой для систем определения положения беспилотной воздушной техники, что является актуальной задачей в условиях отсутствия или вызванного помехами искажения сигнала глобальных систем навигации. ***Цель исследования.*** Разработка и экспериментальная оценка алгоритма уточнения совмещения изображений, сочетающего быстрый нейросетевой отбор кандидатных областей с точным позиционированием на основе сопоставления и геометрической верификации ключевых точек. Особое внимание уделяется работе в условиях, когда кадр, полученный с сенсора ЛА, пересекает границы нескольких спутниковых тайлов. ***Научная новизна*** заключается в определении положения ЛА при отсутствии сигналов глобальных навигационных систем на основе совмещения разнородных изображений подстилающей поверхности для заданного района полёта. На первом этапе производится отбор кандидатных областей с помощью сиамской свёрточной нейронной сети. Сеть анализирует пары, состоящие из текущего бортового изображения и изображений подстилающей поверхности, вычисляя для каждой пары численную оценку визуального сходства, что позволяет сократить зону поиска. На втором этапе выполняется уточнение координат на основе сопоставления ключевых точек. ***Методы.*** Рассмотрены методы обнаружения, описания и сопоставления локальных признаков. В результате применения алгоритма уточнения на тестовой выборке средняя ошибка совмещения снизилась с 300 пикселей после нейросетевого отбора до 20 пикселей. Наиболее точные результаты получены для случаев, когда кадр ЛА полностью располагается в пределах одного спутникового тайла. ***Практическая значимость.*** Разработанный алгоритм обеспечивает определение местоположения ЛА с точностью 10–20 м при времени обработки ≤ 207 мс на основе разнородных изображений подстилающей поверхности.

Ключевые слова: разнородные изображения, совмещение изображений, сиамская архитектура нейронной сети, ключевые точки, визуальное сходство.
DOI: 10.21293/1818-0442-2026-29-2-107-115

Введение

В данной работе рассматривается уточнение совмещения разнородных изображений (на основе небольшого фрагмента, соответствующего кадру, снятому строго вертикально) после первичного, возможно неточного, нейросетевого совмещения. Задача совмещения разнородных изображений, полученных с сенсора ЛА и из спутниковых данных, является частным случаем визуальной навигации в условиях отсутствия сигналов глобальных навигационных систем. Данная задача решается большим количеством исследователей как в Российской Федерации, так и за рубежом. Совмещением разнородных снимков в отечественной научной школе занимались, в частности, А.И. Новиков и А.И. Ефимов; их работы охватывают вопросы сопоставления разнородных изображений, отличающихся по ракурсу и разрешению [1–3].

В работах Ю.В. Визильтера и С.Ю. Желтова предложены методы корреляционно-экстремального совмещения изображений, применимые к данной задаче [4–6].

Для снижения погрешности совмещения пары разнородных изображений предлагается использовать нейросетевые технологии, в частности, декомпозированную сиамскую архитектуру свёрточной нейронной сети.

Так как спутниковые изображения имеют высокое разрешение, их необходимо предварительно обработать и разделить на части. Тайл – это небольшой фрагмент (участок) большой карты или изображения подстилающей поверхности. Каждый тайл записывается в базу данных изображений подстилающей поверхности (БД) и имеет уникальный идентификатор и метаданные.

Задача совмещения разнородных изображений осуществляется на основе пары изображений: одно получено в текущий момент времени с ЛА, второе – из БД изображений. Также известен предполагаемый маршрут ЛА. На предварительном этапе с использованием нейросетевого алгоритма осуществляется отбор тайла-кандидата для совмещения.

Алгоритм уточнения совмещения разнородных изображений на основе ключевых точек состоит из двух этапов: предварительный отбор тайлов-кандидатов в окрестности заданного маршрута на основе сиамской нейросетевой архитектуры и непосредственно уточнение, где с использованием ключевых точек осуществляется визуальная локализация внутри этих областей [7–9].

Компанией SandboxAQ разработана система AQNav для геомагнитной навигации в условиях отсутствия или ненадежности сигналов систем глобальной навигации [10]. Данная система использует дорогостоящее оборудование (квантованные сенсоры) и позволяет осуществлять навигацию с погрешностью до ~1,8 км [11].

Астрономическая навигация для БПЛА рассматривается в работе [12], погрешность навигации

составляет до 4 км. Для астронавигации [12] необходимы специализированное оборудование (оптика), а также чистое звёздное небо. Существенным недостатком данной системы является невозможность работы в условиях слабой видимости (туман, облачность, городская засветка), а также днем.

С помощью программного комплекса «Навигация» решается задача уточнения положения БПЛА вблизи известной точки [13]. Для экспериментальных исследований [13] проводились имитационное моделирование видеосъемки, полученной с сенсора БПЛА, подготовка исходных данных и полетных заданий. Предложенный подход ориентирован на решение более общей задачи – определение визуальной локализации «с нуля» после полной потери ориентировки в окрестностях заданного маршрута. На первом этапе осуществляется отсеивание зоны неопределенности («разброс»), на втором этапе – позиционирование.

Таким образом, целью данной работы является разработка и исследование алгоритма связывания ключевых точек между изображениями пары «спутник–ЛА» для уточнения позиции ЛА. Известны ориентировочная высота и углы ориентации (крен, тангаж, азимут), что позволяет прогнозировать и компенсировать ожидаемые геометрические искажения. В отличие от подходов, ориентированных на непосредственное совмещение изображений либо на уточнение положения вблизи заранее известной точки [4–6, 13–16], предлагаемый алгоритм использует двухэтапную схему: нейросетевое сужение области поиска и последующее локальное уточнение по ключевым точкам с геометрической верификацией. Отличительной особенностью является выполнение уточнения в кандидатной области, включающей несколько соседних тайлов, что позволяет обрабатывать случаи пересечения кадром ЛА границ тайлов. Новизна работы состоит прежде всего в указанной схеме совместного использования нейросетевого отбора и локального геометрического уточнения, а не в отдельных базовых операциях ORB, SIFT или RANSAC.

Анализ снимков с ЛА

Ключевой задачей является определение координат ЛА в условиях отсутствия сигналов спутниковой навигации. Решение данной задачи сводится к сопоставлению данных, полученных с сенсора ЛА, и спутникового изображения. Сложность представляет сопоставление разнородных изображений: текущего кадра с борта ЛА и заранее полученного снимка подстилающей поверхности, что обусловлено следующими факторами:

– изменение условий съемки (сезон, время суток, погода);

– различия в ракурсе и проекции;

– наличие шумов и искажений;

– ограниченная вычислительная мощность бортового оборудования.

Для повышения устойчивости алгоритма к изменяющимся условиям съёмки предлагается использовать подавление низкочастотных составляющих с применением спектрального преобразования Уолша, которое вычисляется следующим образом:

$$\mathrm{Wal}(k,n) = \sum_{j=0}^{m-1} (\mathrm{Rad}_j(n))^{k_j(r)},$$

где $\mathrm{Wal}(k,n)$ – функция Уолша порядка k в точке n; $N = 2^m$ – длина сигнала; $k_j(r)$ – j-й бит в коде Грея числа r, который вычисляется как

$$k_j(r) = k_j \oplus k_{j+1},$$

где k_j – j-й бит двоичного представления k; $\mathrm{Rad}_j(n)$ – функция Радемахера порядка j:

$$\mathrm{Rad}_j(n) = (-1)^{n/2_j}.$$

Такая обработка направлена на уменьшение влияния плавных изменений яркости и крупномасштабных различий освещения между бортовым и спутниковым изображениями. Фильтрация применяется к обоим изображениям до извлечения локальных признаков. Её вклад оценивался отдельным экспериментом на той же тестовой выборке: без фильтрации средняя ошибка (Validation Accuracy) составила 0,7300; с фильтрацией – 0,8081. Количество удаляемых коэффициентов преобразования Уолша составляет > 50%.

Для уточнения совмещения применяются методы, основанные на локальных признаках и ключевых точках, вычислительная сложность которых допускает реализацию на бортовом оборудовании с ограниченными вычислительными ресурсами.

Совмещение разнородных изображений внутри кандидатной области

Предложенный алгоритм осуществляет уточнение внутри кандидатной области, которая может включать несколько тайлов. Уточнение осуществляется с использованием ключевых точек.

На рис. 1 приведены различные варианты расположения снимка с ЛА относительно тайлов местности в окрестности заданного маршрута. На рис. 1 в рамке, выделенной пунктиром (затемнённый участок), отмечен кадр с ЛА, полученный в текущий момент времени. Точка-треугольник в центре этой прямоугольной рамки (FRAME UAV) – центр кадра с ЛА. Также выделены границы тайлов, на которые разбит спутниковый снимок, и отмечены номера тайлов.

Возможны следующие варианты расположения снимка с ЛА относительно тайлов местности: кадр с ЛА полностью перекрывает тайл спутникового снимка (см. рис. 1, *а*), перекрытие нескольких тайлов спутникового снимка (см. рис. 1, *б–г*).

Ключевые точки

Для уточнения совмещения предлагается использовать ключевые точки. Они, как правило, располагаются на участках с высокой информативностью. Для нахождения ключевых точек используются детекторы, а для их описания – дескрипторы. Дескриптор формирует для точки вектор фиксированной размерности.

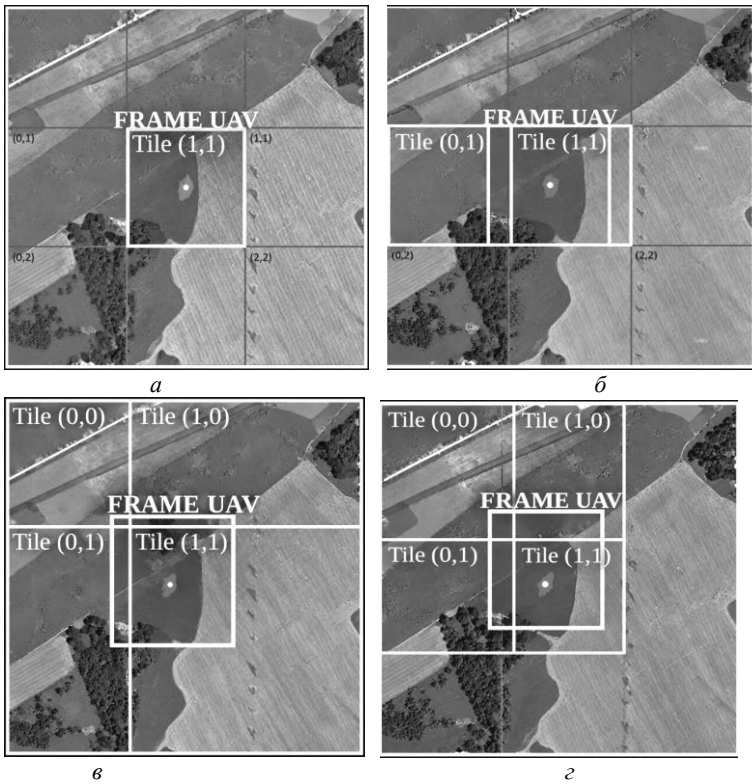


Рис. 1. Визуализация расположения ЛА относительно тайлов местности

Благодаря дескрипторам становится возможным идентифицировать и сравнивать одни и те же локальные особенности на разных изображениях, несмотря на изменения масштаба, условий освещения и перспективы.

В данной работе рассматривались методы обнаружения, описания и сопоставления локальных признаков, приведённые в табл. 1 [17–22]. Следует отметить, что LoFTR относится к детектор-свободным методам сопоставления, тогда как Deep Features и TIMM Features используют признаки, извлекаемые предобученными свёрточными нейронными сетями. Поэтому далее эти подходы сопоставляются как альтернативные способы получения соответствий между парой изображений, а не только как детекторы ключевых точек [23, 24].

В табл. 1 приведены характеристики найденных соответствий.

Таблица 1

Сравнение методов локальных признаков

Метод	Время обработки, мс	Найдено ключевых точек	Количество совпадающих точек	Доля совпадений, %
BRISK	583,93	100	7	7,0
LoFTR	2 800,15	228	22	9,6
ORB	48,76	100	10	10,0
AKAZE	154,41	100	11	11,0
Deep Features	1 491,8	20	6	30,0
ORB+SIFT	1 104,18	21	8	38,1
SIFT	254,02	11	5	45,5
TIMM Features	1 112,4	14	9	64,3

Метод Deep Features использует активации свёрточных слоёв предобученной CNN (например, ResNet50) в качестве дескрипторов [18]. TIMM Features использует модели из библиотеки timm, содержащей широкий набор предобученных архитектур, что позволяет выбирать глубину и тип извлекаемых признаков.

В рамках данной задачи требуется высокая надёжность каждого соответствия при приемлемой скорости сопоставления пар изображений. Рассмотрим нейросетевые методы Deep/TIMM Features, SIFT и комбинацию ORB+SIFT. В табл. 2 приведено сравнение их вычислительных характеристик.

Исходя из данных табл. 1 и 2, можно сделать вывод, что нейросетевые методы TIMM и Deep Features демонстрируют высокую долю совпадений (до 64,3%), однако требуют дополнительного времени и памяти на инициализацию модели. LoFTR формирует большое число соответствий, но в проведённом эксперименте имел низкую долю принятых совпадений (9,6%) и наибольшее время обработки. Связка ORB+SIFT выбрана как компромиссный вариант для бортового применения: в измерениях она обеспечила 29,1 пары/с при объёме памяти 38 МБ, т.е. несколько более высокую скорость и меньший объём памяти, чем SIFT (27,6 пары/с и 45 МБ).

Таблица 2

Сравнение методов локальных признаков

Метод	Скорость, пар/с	Память, МБ	Инициализация, мс
Deep Features	14,0	450	1 200
SIFT	27,6	45	0
ORB+SIFT	29,1	38	0
TIMM Features	79,5	280	860

При этом окончательное решение о соответствии принимается после фильтра Lowe и геометрической верификации RANSAC, поэтому доля исходных совпадений из табл. 1 не является единственным критерием выбора.

Алгоритм уточнения совмещения разнородных изображений

Алгоритм уточнения совмещения разнородных изображений состоит из следующих этапов:

Этап 1. Предварительная подготовка.

1. Фильтрация низких частот.

2. Подготовка БД спутниковых снимков. Разделение изображений на тайлы.

3. Нейросетевой отбор тайлов. Осуществляется анализ тайлов в окрестности заданного маршрута. Тайл-кандидат выбирается по значению косинусного сходства признаков, рассчитываемого по следующей формуле:

$$\text{cosine}_{\text{similar}}(A,B)=\frac{(\mathbf{A},\mathbf{B})}{\|\mathbf{A}\|\cdot\|\mathbf{B}\|},$$

где (A, B) – скалярное произведение векторов-признаков **A** и **B**; $\|\mathbf{A}\|$ и $\|\mathbf{B}\|$ – их нормы; i – индекс суммирования; m – размерность вектора признаков. Здесь обозначение m относится только к размерности вектора признаков и не связано с параметром m, используемым выше при описании преобразования Уолша.

На рис. 2 приведена зависимость средней ошибки совмещения пары изображений (в пикселях) от размера кандидатной области тайлов.

На рис. 3 приведена зависимость времени поиска и средней ошибки в пикселях от размера кандидатной области.

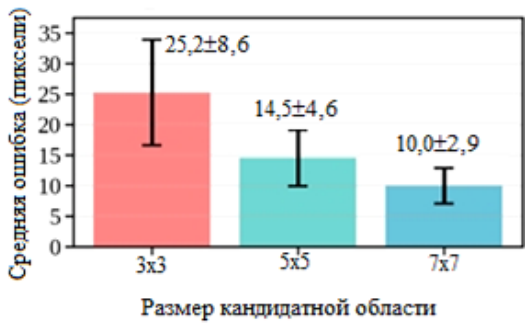


Рис. 2. Зависимость средней ошибки совмещения пары изображений (в пикселях) от размера кандидатной области тайлов

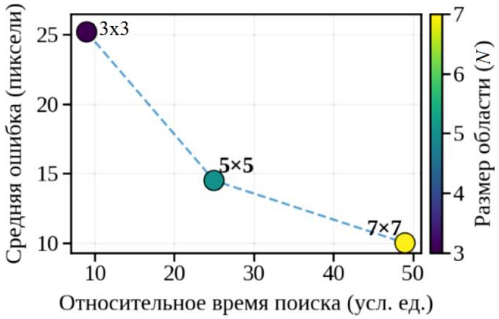


Рис. 3. Зависимость времени поиска и средней ошибки в пикселях от размера области поиска

Этап 2. Детектирование и описание ключевых точек.

1. Детектирование точек ORB-детектором. Для каждого изображения с ЛА и каждого кандидатного тайла применяется ORB-детектор, позволяющий устойчиво находить ключевые точки на разных масштабах.

2. Вычисление дескрипторов SIFT. Для ключевых точек, обнаруженных детектором ORB, вычисляются SIFT-дескрипторы.

Этап 3. Сопоставление и верификация.

1. Поиск соответствий. Для поиска соответствий используется алгоритм FLANN. Для отсечения ложных совпадений применяется критерий отношения Лоу (Lowe ratio test): сохраняются только пары точек, для которых отношение расстояния до ближайшего соседа к расстоянию до второго ближайшего соседа меньше 0,8 [20].

2. Геометрическая верификация. Для каждого кандидатного тайла с достаточным количеством совпадений необходимо вычислить гомографию с помощью RANSAC [25]. Сохраняются только внутренние точки, соответствующие найденной модели.

Этап 4. Уточнение совмещения изображений.

1. Производится комбинирование метрик: количества совпадений, доли внутренних точек после RANSAC и оценки нейросетевого алгоритма. Перед объединением показатели приводятся к сопоставимому диапазону. Выбирается тайл с максимальной итоговой оценкой.

2. Центр кадра с ЛА проектируется на выбранный тайл. Локальные координаты тайла преобразуются в глобальные координаты. В результате определяются глобальные координаты точки, соответствующей центру кадра; при принятой модели съёмки и известной ориентации ЛА эта точка используется для уточнения оценки его положения.

Для перевода в географические координаты используется модель билинейной интерполяции, рассчитываемой по формуле

$$\text{lat}_{\text{est}}=\text{lat}_{\text{TL}}-\frac{y}{H}\cdot(\text{lat}_{\text{TL}}-\text{lat}_{\text{BR}}),$$

$$\text{lon}_{\text{est}}=\text{lon}_{\text{TL}}+\frac{x}{W}\cdot(\text{lon}_{\text{BR}}-\text{lon}_{\text{TL}}),$$

где $(\text{lat}_{\text{TL}}, \text{lon}_{\text{TL}})$ – координаты верхнего левого угла тайла; $(\text{lat}_{\text{BR}}, \text{lon}_{\text{BR}})$ – координаты нижнего правого угла тайла; W, H – ширина и высота тайла в пикселях; $(\text{lat}_{\text{est}}, \text{lon}_{\text{est}})$ – расчётные географические координаты точки.

Этап 5. Оценка достоверности. Точность результата оценивается на основе количества геометрически согласованных совпадений. Для верификации выполняется визуализация итогового совмещения.

Подготовка эксперимента

Для тестирования механизма сопоставления ключевых точек был сформирован набор данных, состоящий из 2000 пар разнородных изображений, полученных с разных локаций: городская застройка, песчаная пустыня с застройкой, чистый лес и мегаполис. Все пары изображений были предварительно выровнены по высоте и углу съёмки и имели извест-

ное эталонное взаимное положение. Исследовалось 11 различных сценариев. Такая структура данных позволяет оценивать остаточную ошибку совмещения относительно известного эталонного положения.

Для каждой пары изображений ошибка совмещения определяется как евклидово расстояние между оценённым и эталонным положением центра кадра в общей пиксельной системе координат. В качестве основных агрегированных метрик используются средняя ошибка, медиана, RMSE и среднеквадратическое отклонение. Для оценки корректности отдельных соответствий введён порог 10 пикселей: соответствие считается допустимым, если расстояние между эталонной и найденной точками не превышает этого значения. Такое значение соответствует 10 м на местности, что является физически обоснованным допуском.

Экспериментальные результаты

В данной работе была проведена экспериментальная оценка работоспособности предложенного алгоритма уточнения и его сравнение с базовыми подходами. Для сравнения выбраны: прямое сопоставление SIFT как классический локальный метод; нейросетевой отбор тайла как результат первого этапа предлагаемого конвейера; VGG16+L2 как базовый вариант глобального CNN-дескриптора с расстоянием L2; CVM-Net как специализированная нейросетевая архитектура для сопоставления изображений разных ракурсов [26]. Сравнительные характеристики приведены в табл. 3.

Точность совмещения до применения алгоритма уточнения с использованием ключевых точек составляла 300 пикселей, после – до 20 пикселей. Предложенный алгоритм демонстрирует высокую точность на тестовой выборке на участках с выраженными текстурами (городская застройка, дороги, изгиб реки). Однако его эффективность снижается на однородных поверхностях (водная гладь, поля), где количество ключевых точек недостаточно для точного сопоставления.

Наилучший результат у сценария «Мегаполис», наихудший результат – «Обрабатываемое поле».

Результаты, представленные в табл. 3, получены на одном тестовом наборе данных путем повторных запусков исследуемых методов.

На рис. 4 приведен результат совмещения. Линией выделена траектория полета ЛА, квадратами выбранные тайлы.

На рис. 5 приведены оценка ошибки первого и второго рода для различных наборов сценариев.

Таблица 3

Сравнительные характеристики методов

Метод	Средняя ошибка, пикс.	Время, мс
SIFT (прямое)	>500	180
Нейросетевой отбор тайла	300	106
VGG16 + L2	150	300
CVM-Net	80	450
Алгоритм уточнения на основе ключевых точек	20	207

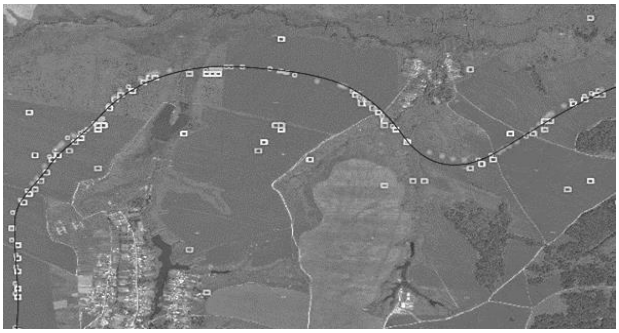


Рис. 4. Траектория полета ЛА
и выбранные тайлы для совмещения

Выводы

В рамках данной работы разработан и экспериментально исследован алгоритм уточнения совмещения изображений, полученных с разных сенсоров. Подход включает первоначальный нейросетевой отбор и последующее позиционирование за счёт сопоставления локальных признаков [27, 28].

Экспериментальная проверка показала, что применение связки ORB+SIFT с последующей фильтрацией соответствий по критерию Лоу и геометрической верификацией методом RANSAC позволяет снизить среднюю ошибку совмещения с 300 до 20 пикселей относительно результата после этапа нейросетевого поиска.

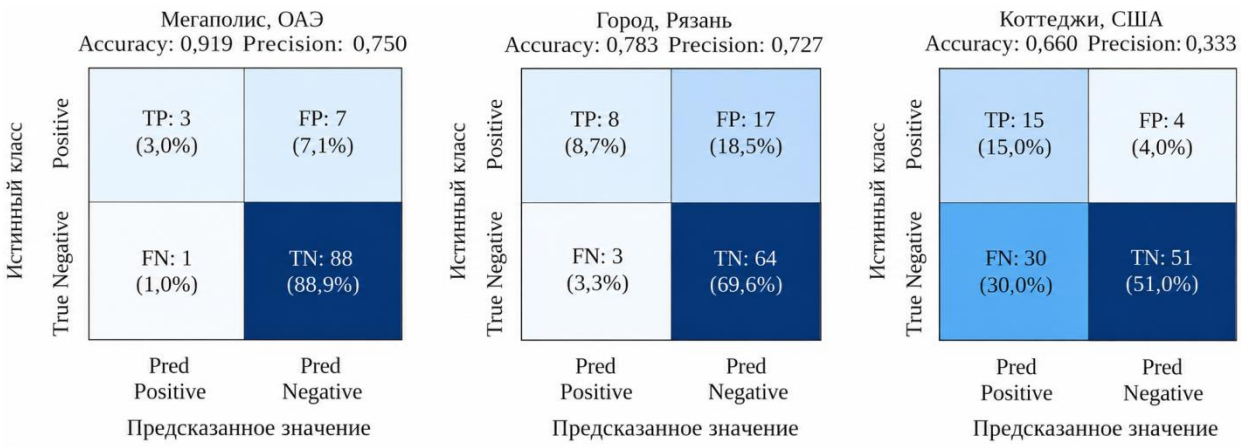


Рис. 5 (начало)

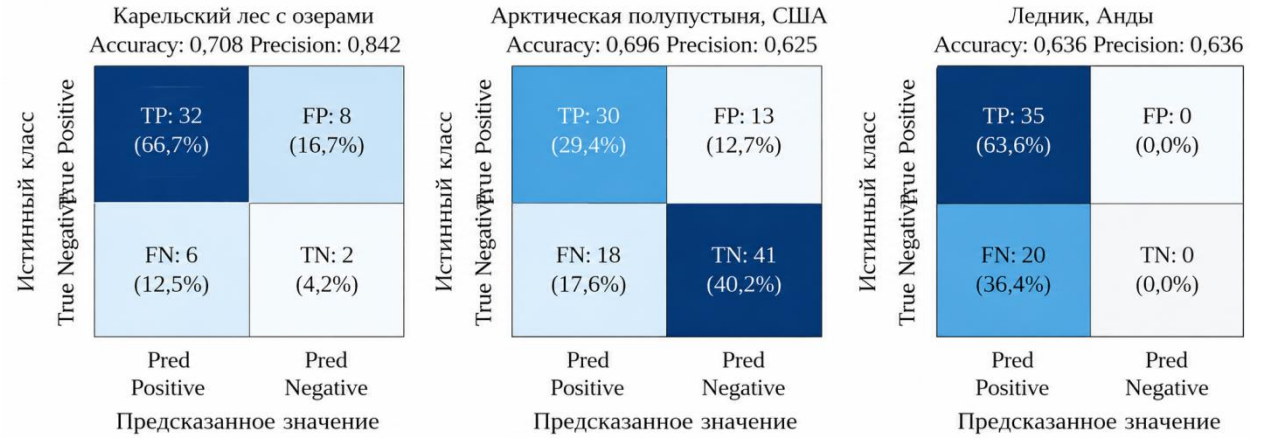


Рис. 5 (окончание). Ошибки первого и второго рода для различных сценариев

Среднее время выполнения полного уточнения составляет 207 мс, что соответствует частоте около 4,8 Гц. При скорости ЛА 50 м/с за один такой интервал аппарат перемещается примерно на 10,4 м.

Алгоритм сохраняет работоспособность в сценариях, при которых кадр ЛА пересекает границы нескольких тайлов спутниковых изображений.

Наиболее высокая точность достигается на участках с хорошо выраженными и устойчивыми текстурными деталями. На однородных поверхностях число геометрически согласованных соответствий уменьшается, вследствие чего возрастает ошибка и снижается доля успешных уточнений. Дальнейшие исследования целесообразно направить на повышение устойчивости алгоритма на слабоструктурированных участках и на сокращение времени обработки на бортовой вычислительной платформе.

Литература

1. Новиков А.И. Алгоритмы автоматической идентификации объектов на разнородных изображениях и совмещения изображений / А.И. Новиков, А.И. Ефимов, Д.А. Колчаев // Сборник трудов IV Междунар. конф. и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ–2018). – Самара: Новая техника, 2018. – С. 650–663.

2. Ефимов А.И. Комбинированный алгоритм совмещения облаков точек / А.И. Ефимов, Т.Н. Крючкова, А.И. Ярославцева // GraphiCon – 2023: труды 33-й Междунар. конф. по компьютерной графике и машинному зрению. – М.: Институт прикладной математики им. М.В. Келдыша РАН, 2023. – С. 539–548.

3. Sablina V.A. Combined approach to object contour superimposition in heterogeneous images / V.A. Sablina, A.I. Efimov, A.I. Novikov // 2017 6th Mediterranean Conference on Embedded Computing (MECO). – Bar, Montenegro: IEEE, 2017. DOI: 10.1109/MECO.2017.7977175.

4. Алгоритм автоматического совмещения сенсорной и синтезируемой видеоинформации для авиационной системы комбинированного видения / М.А. Лебедев, М.А. Бондаренко, Д.В. Комаров, Д.Г. Степаньянц, О.В. Выголов, Ю.В. Визильтер, С.Ю. Желтов // Вестник компьютерных и информационных технологий. – 2014. – № 7 (121). – С. 8–13.

5. Сравнение изображений по форме с использованием диффузной морфологии и диффузной корреляции / Ю.В. Визильтер, В.С. Горбацевич, А.Ю. Рубис, О.В. Выголов // Компьютерная оптика. – 2015. – Т. 39, № 2. – С. 265–274.

6. Визильтер Ю.В. Современные технологии искусственного интеллекта и их применение в авиационных комплексах / Ю.В. Визильтер, Б.В. Вишняков, С.Ю. Желтов // XVI Всерос. мультиконф. по проблемам управления (МКПУ–2023): матер. мультиконф. – Волгоград: Волгоград. гос. техн. уни-т, 2023. – С. 13–16.

7. Кочетуров В.В. Сиамские свёрточные нейронные сети для распознавания лиц / В.В. Кочетуров, О.И. Федяев // Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2024): сборник матер. и докл. V Междунар. науч.-практ. конф. – Донецк: Дон. нац. техн. ун-т, 2024. – С. 178–184.

8. Ладанова Е.О. Разработка методики и новой структуры сиамской нейронной сети для анализа геопространственных данных // Научно-технический вестник Поволжья. – 2024. – № 12. – С. 144–147.

9. Li Y. A Survey on Siamese Network: Methodologies, Applications, and Opportunities / Y. Li, C.L.Ph. Chen, T. Zhang // IEEE Transactions on Artificial Intelligence. – 2022. – Vol. 3, No. 6. – P. 994–1014. DOI: 10.1109/TAI.2022.3207112.

10. Broughton G. Robust Image Alignment for Outdoor Teach-and-Repeat Navigation / G. Broughton, P. Linder, T. Rouček, T. Vintr, T. Krajník // Proceedings of the 2021 European Conference on Mobile Robots (ECMR). – Bonn, Germany: IEEE, 2021. – P. 1–6. DOI: 10.1109/ECMR50962.2021.9568832.

11. Rozsypalek Z. Semi-Supervised Learning for Image Alignment in Teach and Repeat Navigation / Z. Rozsypalek, G. Broughton, P. Linder, T. Roucek, K. Kusumam, T. Krajnik // Proceedings of 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22). – New York, USA: ACM, 2022. DOI: 10.1145/3477314.3507061.

12. Teague S. An Algorithm for Affordable Vision-Based GNSS-Denied Strapdown Celestial Navigation / S. Teague, J. Chahl // Drones. – 2024. – Vol. 8, No. 11. – Art. 652. DOI: 10.3390/drones8110652.

13. Программный комплекс для автоматизации научных исследований в области создания перспективных систем навигации беспилотных летательных аппаратов по данным видеонаблюдения / Б.А. Алпатов, П.В. Бабаян, Ю.С. Коблов, В.С. Муравьев, В.В. Стротов, А.Б. Фельдман // Вестник Рязан. гос. радиотехнического ун-та. – 2012. – № 40. – С. 9–15.

14. Оптико-электронная корреляционно-экстремальная навигационная система малого беспилотного летательного аппарата / А.А. Шейников, В.А. Малкин, Э.А. Смоленский, Л.А. Иваницкий // Системный анализ и прикладная

информатика. – 2024. – № 2. – С. 30–40. DOI: 10.17212/2307-5981-2024-2-30-40.

15. ТТС-ориентированная визуальная навигация систем БПЛА / Е.А. Конников, Д.А. Сергеев, Г.И. Голиков, Д.Г. Родионов, П.А. Поляков, А.Ф. Тихомиров // Программные системы и вычислительные методы. – 2025. – № 4. – С. 1–26. DOI: 10.7256/2454-0714.2025.4.75830.

16. Обнаружение объектов в аэронавигации с использованием вейвлет-преобразования и сверточных нейронных сетей: первый подход / Х. Фортуна-Сервантес, М. Рамирес-Торрес, Х. Мартинес-Карранса, Х. Мургуия-Ибарра, М. Мехия-Карлос // Труды Института системного программирования РАН. – 2021. – Т. 33, № 2. – С. 149–162. DOI: 10.15514/ISPRAS-2020-33(2)-9.

17. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision. – 2004. – Vol. 60. – P. 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94.

18. Денисов А.А. Анализ методов детектирования, описания и сопоставления ключевых точек изображений / А.А. Денисов, А.И. Новиков // Вестник Рязан. гос. радиотехн. ун-та. – 2024. – № 89. – С. 104–116.

19. Bay H. SURF: Speeded Up Robust Features / H. Bay, T. Tinne, V.G. Luc // Proc. of. Computer Vision – ECCV 2006. – Lecture Notes in Computer Science. – 2006. – Vol. 3951. – Berlin, Heidelberg: Springer, 2006. https://doi.org/10.1007/11744023_3.

20. Mikolajczyk K. A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – P. 1615–1630. DOI: 10.1109/TPAMI.2005.188.

21. Дышлюк В.О. Исследование показателей качества и быстродействия поиска опорных точек на изображениях методом SURF // Молодой ученый. – 2018. – № 27. – С. 23–26.

22. DeTone D., Malisiewicz T., Rabinovich A. SuperPoint: Self-Supervised Interest Point Detection and Description // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – Salt Lake City, UT, USA: IEEE. DOI: 10.1109/CVPRW.2018.00060.

23. Костров Б.В. Разработка гибридной архитектуры свёрточной нейронной сети с использованием спектральных преобразований / Б.В. Костров, С.И. Бабаев, А.И. Ефимов, В.Ю. Тарасова // Известия ЮФУ. Технические науки. – 2026. – № 1 (249). – С. 235–249. DOI: 10.18522/2311-3103-2026-1-235-249.

24. Костров Б.В Подходы к формированию набора данных и его предварительной обработки в задачах обучения свёрточных сиамских нейронных сетей для коррекции положения летательного аппарата по аэрофотосъемке / Б.В. Костров, С.И. Бабаев, А.И. Ефимов, В.Ю. Тарасова // XXI век: итоги прошлого и проблемы настоящего плюс. – 2025. – Т. 14, № 3 (71). – С. 70–77.

25. Li J. Point Cloud Registration Based on One-Point RANSAC and Scale-Annealing Biweight Estimation / J. Li, Q. Hu, M. Ai // IEEE Transactions on Geoscience and Remote Sensing. – 2021. –Vol. 59, No. 11. – P. 9716–9729. DOI: 10.1109/TGRS.2020.3048026.

26. Hu S. CVM-Net: Cross-View Matching Network for Image-Based Ground-to-Aerial Geo-Localization / S. Hu, M. Feng, R.M.H. Nguyen, G.H. Lee // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018. Salt Lake City, UT, USA: IEEE, 2018. – P. 7258–7267. DOI: 10.1109/CVPR.2018.00758.

27. Св-во о гос. регистрации программы для ЭВМ № 2025691610 / Костров Б.В., Устюков Д.И., Тарасова В.Ю. // Свёрточная нейронная сеть с гибридными слоями для определения визуального сходства спутниковых снимков. Дата поступления 17 ноября 2025 г. Зарегистрировано в Реестре программ для ЭВМ 17 ноября 2025 г.

28. Никифоров М.Б. Метод определения местоположения съемки фотографии на основе визуального поиска / М.Б. Никифоров, В.Ю. Тарасова // XXI век: итоги прошлого и проблемы настоящего плюс. – 2023. – Т. 12, № 4 (64). – С. 103–110.

Бабаев Сергей Игоревич
Канд. техн. наук, доцент каф. электронных вычислительных машин (ЭВМ) Рязанского государственного радиотехнического университета им. В.Ф. Уткина» (РГРТУ)
Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0001-5829-8223
Тел.: +7 (491-2) 72-03-50
Эл. почта: babaev.s.i@gmail.com

Костров Борис Васильевич
Д-р техн. наук, проф., зав. каф. ЭВМ РГРТУ
Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0002-2984-4480
Тел.: +7 (491-2) 72-03-50
Эл. почта: kostrov.b.v@evm.rsreu.ru

Муратов Евгений Рашитович
Канд. техн. наук, доцент каф. ЭВМ РГРТУ
Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0002-1664-3954
Тел.: +7 (491-2) 72-03-50
Эл. почта: myratov_er@mail.ru

Тарасов Андрей Сергеевич
Канд. техн. наук, доцент каф. ЭВМ РГРТУ
Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0002-2812-3489
Тел.: +7 (491-2) 72-03-50
Эл. почта: vb2005@yandex.ru

Тарасова Валентина Юрьевна
Ст. преподаватель каф. ЭВМ РГРТУ
Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0009-0007-4358-7736
Тел.: +7-903-837-37-69
Эл. почта: valentina2008.91@mail.ru

Поступила в редакцию: 15.04.2026.
Принята к публикации: 31.08.2026.

Babaev S.I., Kostrov B.V., Muratov E.R.,
Tarasov A.S., Tarasova V.Yu.

Keypoint-based algorithm for registration of heterogeneous images

This paper discusses an algorithm for refining heterogeneous images registration. Image registration is performed using neural network-based selection of candidate regions. ***Relevance.*** The registration of heterogeneous images can serve as a basis for unmanned aerial vehicle (UAV) positioning systems, which is a relevant task when signals from global navigation satellite systems are unavailable or distorted due to interference.

Purpose of the study. The purpose of this study is to develop and experimentally evaluate an image registration refinement algorithm that combines fast neural network-based selection of candidate regions with precise positioning based on matching and geometric verification of keypoints. Particular attention is paid to scenarios in which a frame acquired by an onboard UAV sensor crosses the boundaries of multiple satellite imagery tiles. ***Scientific novelty.*** The scientific novelty lies in the position of a UAV in the absence of signals from global navigation satellite systems by registering heterogeneous images of the underlying surface within a specified flight area. At the first stage, candidate regions are selected using a Siamese convolutional neural network. The network analyzes pairs consisting of a current onboard image and images of the underlying surface, calculating a numerical visual similarity score for each pair, thereby reducing the search area. At the second stage, coordinates are refined based on keypoint matching. ***Methods.*** Methods for detecting, describing, and matching local features are considered. As a result of applying the refinement algorithm to the test sample, the average alignment error decreased from 300 pixels after neural network selection to 20 pixels. The most accurate results were obtained for cases in which the UAV frame was entirely located within a single satellite imagery tile. ***Practical significance.*** The developed algorithm enables UAV positioning with an accuracy of 10–20 m and a processing time of ≤ 207 ms based on heterogeneous images of the underlying surface.

Keywords: heterogeneous images, image registration, Siamese neural network architecture, keypoints, visual similarity.
DOI: 10.21293/1818-0442-2026-29-2-107-115

References

1. Novikov A.I., Efimov A.I., Kolchaev D.A. [Algorithms for automatic identification of objects on heterogeneous images and images superimposition]. Informatsionnye tekhnologii i nanotekhnologii. [*Information Technologies and Nanotechnologies*] Proc. of IV International Conference and Youth School ITNT–2018, Samara, Novaya tekhnika, 2018, pp. 650–663 (in Russ.).

2. Efimov A.I., Kryuchkova T.N., Yaroslavtseva A.I. [Combined Algorithm for Combining Point Clouds]. Proc. of 33rd International Conference on Computer Graphics and Vision GraphiCon 2023, 2023, no. 33, pp. 539–548 (in Russ.).

3. Sablina V.A., Efimov A.I., Novikov A.I. Combined approach to object contour superimposition in heterogeneous images. 2017 6th Mediterranean Conference on Embedded Computing (MECO), 2017, pp. 7977175.

4. Lebedev M.A., Bondarenko M.A., Komarov D.V., Stepaniants D.G., Vygolov O.V., Vizilter Yu.V., Zheltov S.Yu. [An algorithm for real and synthetic images matching with application to aviation combined vision system]. *Vestnik komp'iuternykh i informatsionnykh tekhnologii*, 2014, no. 7 (121), pp. 8–13 (in Russ.).

5. Vizilter Yu.V., Gorbatsevich V.S., Rubis A.Yu., Vygolov O.V. Image shape matching using diffusion morphology and diffusion correlation. *Computer optics*, 2015, vol. 39, no. 2, pp. 265–274 (in Russ.).

6. Vizilter Yu.V., Vishnyakov B.V., Zheltov S.Yu. *Sovremennye tekhnologii iskusstvennogo intellekta i ikh primenenie v aviatsionnykh kompleksakh* [Modern artificial intelligence technologies and their application in aviation systems]. XVI Vserossiyskaya mul'tikonferentsiya po problemam upravleniya (MKPU-2023) [*Proc. of XVI All-Russian Multiconference on Control Problems*], Volgograd, Volgograd State Technical University, 2023, pp. 13–16 (in Russ.).

7. Kocheturov V.V., Fedyaev O.I. *Siameskie svertochnye neyronnye seti dlya raspoznavaniya lits*. [Siamese convolutional neural networks for face recognition] Programmnaya inzheneriya: metody i tekhnologii razrabotki informatsionnovychislitel'nykh sistem (PIIVS-2024) [*Software Engineering: Methods and Technologies for Developing Information and Computing Systems*]. Proc. of the 5th International Scientific and Practical Conference, Donetsk, Donetsk National Technical University, 2024, pp. 178–184 (in Russ.).

8. Ladanova E.O. [Development of a methodology and a new structure of the Siamese neural network for analysis of geospatial data]. *Scientific and Technical Volga region Bulletin*, 2024, no. 12, pp. 144–147 (in Russ.).

9. Li Y., Chen C.L.Ph., Zhang T. A Survey on Siamese Network: Methodologies, Applications, and Opportunities. *IEEE Transactions on Artificial Intelligence*, 2022, vol. 3, no. 6, pp. 994–1014. DOI: 10.1109/TAI.2022.3207112.

10. Broughton G., Linder P., Rouček T., Vintr T., Krajník T. Robust Image Alignment for Outdoor Teach-and-Repeat Navigation. *Proceedings of the 2021 European Conference on Mobile Robots (ECMR)*, Bonn, Germany, IEEE, 2021, pp. 1–6. DOI: 10.1109/ECMR50962.2021.9568832.

11. Rozsypalek Z., Broughton G., Linder P., Roucek T., Kusumam K., Krajnik T. Semi-Supervised Learning for Image Alignment in Teach and Repeat Navigation. Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22), New York, USA, ACM, art. 4. DOI: 10.1145/3477314.3507045

12. Teague S., Chahl J. An Algorithm for Affordable Vision-Based GNSS-Denied Strapdown Celestial Navigation. *Drones*, 2024, vol. 8, no. 11, art. 652. DOI: 10.3390/drones8110652.

13. Alpatov B.A., Babayan P.V., Koblov Yu.S., Muravyev V.S., Strotov V.V., Feldman A.B. [Software for automatization of scientific investigations aimed at development of perspective vision-based navigation systems for unmanned aerial vehicles]. *Vestnik of Ryazan state radio engineering university*, 2012, no. 40, pp. 9–15 (in Russ.).

14. Sheinikov A.A., Malkin V.A., Smolensky E.A., Ivanitsky L.A. [Optical-electronic correlation-extreme navigation system of small unmanned aerial vehicle]. *System analysis and applied information science*, 2024, no. 2, pp. 30–40. DOI: 10.17212/2307-5981-2024-2-30-40 (in Russ.).

15. Konnikov E.A., Sergeev D.A., Golikov G.I., Rodionov D.G., Polyakov P.A., Tikhomirov A.F. [TTC-oriented visual navigation systems for UAVS]. *Software systems and computational methods*, 2025, no. 4, pp. 1–26. DOI: 10.7256/2454-0714.2025.4.75830 (in Russ.).

16. Fortuna-Cervantes J., Ramírez-Torres M., Martinez-Carranza J., Murguía-Barra J., Mejía-Carlos M. Object Detection in Aerial Navigation using Wavelet Transform and Convolutional Neural Networks: A first Approach. *Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS)*, 2021, vol. 33, no. 2, pp. 149–162. DOI: 10.15514/ISPRAS-2020-33(2)-9 (in Russ.).

17. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 2004, vol. 60, pp. 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94.

18. Denisov A.A., Novikov A.I. [Analysis of image key point detection, description and matching methods]. *Vestnik of Ryazan state radio engineering university*, 2024, no. 89, pp. 104–116 (in Russ.).

19. Bay H., Tinne T., Luc V.G. SURF: Speeded Up Robust Features. Proc. of. Computer Vision – ECCV 2006. ECCV 2006, *Lecture Notes in Computer Science*, vol. 3951. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11744023_3.

20. Mikolajczyk K., Schmid C.A. Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Ana-*

lysis and Machine Intelligence, 2005, pp. 1615–1630. DOI: 10.1109/TPAMI.2005.188.

21. Dyshlyuk V.O. *Issledovanie pokazateley kachestva i bystrodeystviya poiska opornykh tochek na izobrazheniyakh metodom SURF* [Investigation of quality and speed metrics for keypoint detection in images using the SURF method]. *Molodoy uchenyi*, 2018, no. 27, pp. 23–26 (in Russ.).

22. DeTone D., Malisiewicz T., Rabinovich A. SuperPoint: Self-Supervised Interest Point Detection and Description. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, IEEE. DOI: 10.1109/CVPRW.2018.00060.

23. Kostrov B.V., Babaev S.I., Efimov A.I., Tarasova V.Yu. [Convolutional neural network hybrid architecture development using spectral transformations]. *Izvestiya SFedU. Engineering Sciences*, 2026, no. 1 (249), pp. 235–249. DOI: 10.18522/2311-3103-2026-1-235-249 (in Russ.).

24. Kostrov B.V., Babaev S.I., Tarasova V.Yu. [Issues to the data set formation and preprocessing in the training convolutional Siamese neural networks tasks for an aircraft position correcting based on aerial photography]. *XXI century: resumes of the past and challenges of the present plus*, 2025, vol. 14, no. 3 (71), pp. 70–77 (in Russ.).

25. Li J., Hu Q., Ai M. Point Cloud Registration Based on One-Point RANSAC and Scale-Annealing Biweight Estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, vol. 59, no. 11, pp. 9716–9729. DOI: 10.1109/TGRS.2020.3048026.

26. Hu S., Feng M., Nguyen R.M.H., Lee G.H. CVM-Net: Cross-View Matching Network for Image-Based Ground-to-Aerial Geo-Localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, Salt Lake City, UT, USA: IEEE, pp. 7258–7267. DOI: 10.1109/CVPR.2018.00758.

27. Kostrov B.V., Ustyukov D.I., Tarasova V.Yu. *Svertochnaya neyronnaya set' s gibridnymi sloyami dlya opredeleniya vizual'nogo skhodstva sputnikovykh snimkov* [Convolutional neural network with hybrid layers for determining the visual similarity of satellite images]. Certificate of State Registration of a Computer Program No. 2025691610, 2025 (in Russ.).

28. Nikiforov M.B., Tarasova V.Yu. [Method for determining photo shooting location based on visual search]. *XXI century: resumes of the past and challenges of the present plus*, 2023, vol. 12, no. 4 (64), pp. 103–110 (in Russ.).

Sergey I. Babaev
Candidate of Technical Sciences, Associate Professor,
Department of Electronic Computers, Ryazan State Radio Engineering University named after V.F. Utkin (RSREU)
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0001-5829-8223
Phone: +7 (491-2) 72-03-50
Email: babaev.s.i@gmail.com

Boris V. Kostrov
Doctor of Engineering Sciences, Professor, Head
Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0002-2984-4480
Phone: +7 (491-2) 72-03-50
Email: kostrov.b.v@evm.rsreu.ru

Yevgeniy R. Muratov
Candidate of Technical Sciences, Associate Professor
Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0002-1664-3954
Phone: +7 (491-2) 72-03-50
Email: myratov_er@mail.ru

Andrey S. Tarasov
Candidate of Technical Sciences, Associate Professor,
Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0002-2812-3489
Phone: +7 (491-2) 72-03-50
Email: vb2005@yandex.ru

Valentina Yu. Tarasova
Senior Lecturer, Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0009-0007-4358-7736
Phone: +7-903-837-37-69
Email: valentina2008.91@mail.ru

Received: 15.04.2026.
Accepted: 31.08.2026.

УДК 004.85
Шифр специальности ВАК: 2.3.1

В.В. Коломиец

Автоматическое индексирование баз данных с использованием обучения с подкреплением

Актуальность. Индексы существенно влияют на производительность систем управления базами данных, их автоматический выбор при изменяющейся нагрузке остается сложной многокритериальной задачей. ***Цель исследования.*** Разработать запрос-ориентированный метод автоматического управления индексами на основе обучения с подкреплением для работы в высоконагруженной среде. ***Методы.*** Инструменты и подходы, которые применял автор: постановка задачи в виде марковского процесса принятия решений, сравнительный анализ современных подходов к индексированию, оценочный механизм оптимизатора запросов, эксплуатационная телеметрия СУБД, проверка результата на нагрузках TPC-H. ***Научная новизна.*** Предложен метод автоматической индексации с использованием обучения с подкреплением, в котором индексное действие выбирается в локальном контексте оптимизируемого запроса, а суррогатная награда оптимизатора отделена от фактической награды, вычисляемой по статистике выполненных запросов. ***Результаты.*** В последовательном, конкурентном и синтетическом сценариях стратегия обучения с подкреплением превзошла базовую индексацию по времени выполнения и интегральным метрикам производительности. ***Практическая значимость.*** Метод может применяться как компонент безопасного контура автоматической настройки индексов в промышленных СУБД.

Ключевые слова: обучение с подкреплением, автоматическое индексирование, запрос-ориентированный выбор индексов, динамическая настройка индексов, физическое проектирование баз данных, статистика выполнения запросов, оценка стоимости запросов.
DOI: 10.21293/1818-0442-2026-29-2-116-128

Введение

Индексы являются одним из базовых механизмов физического проектирования баз данных и непосредственно определяют методы доступа к данным, объем чтений и время выполнения запросов [1–3]. Наличие подходящего индекса может на несколько порядков снижать объем обрабатываемых данных. Вместе с тем построение и сопровождение индексов сопряжены с дополнительными затратами вычислительных и дисковых ресурсов, а также с возможным возникновением блокировок на объекты базы данных. Вследствие этого возникает риск снижения доступности системы. По данной причине задача выбора индексов должна рассматриваться как многокритериальная.

Классические средства автоматического проектирования индексов, развивавшиеся со времен AutoAdmin, основаны, главным образом, на анализе фиксированной рабочей нагрузки и оценке полезности кандидатов в индексы через оценочный интерфейс оптимизатора запросов [1, 2]. Такие методы доказали практическую ценность и остаются важным инструментом администратора баз данных (англ. *Database Administrator*, DBA), однако имеют ряд ограничений. Во-первых, они ориентированы преимущественно на отложенную настройку по статическому снимку рабочей нагрузки. Во-вторых, они плохо отражают ситуацию, когда нагрузка дрейфует, набор запросов постепенно меняется, а цена ошибки оказывается накопительной: один неудачный индекс ухудшает записи, сужает бюджет и искажает последующие решения. В-третьих, само пространство индексных конфигураций комбинаторно велико, уже для умеренного числа таблиц и атрибутов задача выбора близка к NP-трудной постановке [4, 5].

Здесь рациональным становится применение обучения с подкреплением (*reinforcement learning*, RL). В отличие от одношаговых оптимизационных постановок, RL рассматривает систему как среду, в которой агент последовательно принимает решения, наблюдает реакцию среды и стремится максимизировать долгосрочную совокупную выгоду [6, 7]. Создание индекса – не изолированное действие, а шаг, меняющий состояние СУБД: текущее множество индексов, планы оптимизатора, доступный бюджет памяти, стоимость записи и набор будущих релевантных кандидатов. При этом эффект решения может быть отложенным: новый индекс строится некоторое время, затем начинает использоваться не всеми запросами сразу, а его реальная выгода проявляется лишь после накопления эксплуатационной статистики. RL позволяет описать именно такую последовательную, частично наблюдаемую и стохастическую динамику.

Содержательно RL опирается на несколько базовых понятий: агент, среда, состояние, действие, награда и политика [6]. Агентом в рассматриваемой задаче выступает модуль автоматического индексирования. Средой является сама СУБД вместе с данными, запросами, статистиками и эксплуатационными ограничениями. Состояние кодирует наблюдаемую конфигурацию среды в конкретный момент времени; действие соответствует изменению физического дизайна хранения данных, а награда измеряет, насколько полезным оказалось это изменение. Политика задает правило выбора действий по состоянию. В простейшем варианте политика может учиться выбирать один из нескольких индексов-кандидатов, а в более зрелых постановках – одновременно учитывать тип индекса, бюджет, стоимость обслуживания и риск блокировок. Таким образом, обучение с под-

креплением позволяет дать широкую постановку задачи автоматического индексирования.

Прямое применение классических алгоритмов обучения с подкреплением в промышленной СУБД требует существенной адаптации. Агент должен принимать решения безопасно по отношению к текущей нагрузке, учитывать, что каждый новый индекс повышает затраты на изменение данных и сам процесс создания индекса зачастую влияет на работу базы данных. Поэтому современные исследования развиваются сразу по нескольким направлениям – сокращение и обобщение описания рабочей нагрузки, уменьшение числа рассматриваемых действий, построение более содержательного описания состояния среды, учет ограниченности вычислительных ресурсов, учет особенностей текущей нагрузки [5, 8, 9–11]. В данной работе сформулирована широкая постановка задачи автоматического индексирования в терминах обучения с подкреплением, охватывающая как ранние, так и современные исследования. На этой основе будет проведен сравнительный обзор исследований по крупным смысловым блокам: как авторы описывают среду и состояние, формируют действия, считают награду, какие архитектуры агентов используют и как обучают модели. После этого предлагается новый алгоритм безопасного автоматического индексирования СУБД, основанный на продолжающейся задаче MDP с дискретными действиями, признаках, сформированных на основе нагрузки, локализованной по конкретному запросу и раздельной суррогатной и реальной награде.

Обобщенная постановка задачи автоматического индексирования в терминах обучения с подкреплением

В общем виде задача автоматического индексирования – это процесс последовательного управления физическим дизайном хранения данных. Пусть в момент времени *t* среда описывается как

$$\mathbf{E}_t = \mathbf{W}_t, \Sigma_t, \mathbf{I}_t, \Omega_t,$$

где $\mathbf{W}_t$ – актуальная рабочая нагрузка или ее релевантный фрагмент, Σ_t – статистическая информация о выполнении запросов , $\mathbf{I}_t$ – текущая конфигурация индексов, Ω_t – эксплуатационные ограничения: бюджет памяти, допустимость операций определения данных (англ. Data Definition Language, DDL), временные окна обслуживания и т.д. Тогда задача агента состоит в том, чтобы выбирать действия, переводящие систему в более выгодные состояния с учетом долгосрочных последствий.

Состояние должно быть достаточно информативным и универсальным, чтобы определять полезные действия и модель оставалась обучаемой и переносимой между разными схемами баз данных. Наблюдение $\mathbf{o}_t = \psi(\mathbf{E}_t)$ описывается как

$$\mathbf{o}_t = \psi(\mathbf{W}_t, \Sigma_t, \mathbf{I}_t, \Omega_t). \quad (1)$$

Действие a_t может соответствовать построению нового индекса, удалению или модификации существующего индекса, выбору типа индекса или перераспределению бюджета. Поэтому пространство действий целесообразно описывать как множество допустимых физических преобразований $\mathbf{A}_t = \mathbf{A}(\mathbf{o}_t)$. Для эпизодического обучения терминальным состоянием обычно служит исчерпание бюджета индексов или достижение предельного числа шагов. Для неэпизодических сценариев явного терминального состояния нет, а агент периодически переоценивает ситуацию и принимает очередное локальное решение.

$$a_t \in \mathbf{A}(\mathbf{o}_t), \pi(a_t|\mathbf{o}_t) \to \max \mathrm{E}_\pi\left[\sum_{t=0}^{\infty}\gamma^t r_t\right], \quad (2)$$

где γ^t – коэффициент дисконтирования будущей награды r_t.

Награда r_t должна отражать не только увеличение производительности, но и сопутствующие издержки. Для СУБД лучше задавать награду как разность между приростом производительности одних запросов и штрафами за уменьшение производительности других, штрафами за занимаемое пространство и стоимость DDL-операций. В общем виде такую функцию можно записать следующим образом:

$$r(t) = \Delta U_t^{\text{выигр}} - \lambda_{\text{w}}\,\Delta U_t^{\text{пот}} - \\ -\lambda_{\text{s}}\,\Delta S_t - \lambda_{\text{d}}\,\Delta\,\text{DDL}_t, \quad (3)$$

здесь $\Delta U_t^{\text{выигр}}$ обозначает выигрыш в производительности; $\Delta U_t^{\text{пот}}$ – ухудшение производительности; ΔS_t – ограничения бюджета; $\Delta\,\text{DDL}_t$ – стоимость физического изменения конфигурации, а λ_{w}, λ_{s}, λ_{d} – коэффициенты эксплуатационных предпочтений. В конкретных работах эти компоненты могут измеряться по-разному: через реальные времена выполнения, оценку стоимости оптимизатора, или через аппроксимации, полученные на основе моделей, обученных на результатах выполненных запросов. То есть под наградой понимается то, насколько выбранное действие улучшило выполнение рабочей нагрузки.

На рис. 1 показана обобщенная схема взаимодействия агента и среды в задаче автоматического индексирования – «наблюдение → действие → изменение конфигурации → награда → новое наблюдение», которая лежит в основе всех последующих исследований.

Сравнительный анализ исследований по автоматическому индексированию на основе RL

Ниже проведен обзор работ в исследуемой области за последние 10 лет в терминах, определенных в постановке задачи: представление среды и состояния, пространство действий, функции награды, архитектуры агентов и методология обучения.

Представление среды и состояния

Развитие методов особенно хорошо прослеживается по тому, как авторы описывают состояние среды.



Рис. 1. Обобщенная схема автоматического индексирования в терминах обучения с подкреплением

В ранних работах это описание было сравнительно компактным и зависело от конкретной структуры базы данных – набора таблиц, столбцов и допустимых индексов. Так, в rCOREIL состоянием служил текущий набор уже созданных вторичных индексов. Основное внимание в этом подходе уделялось тому, насколько выбранный набор индексов уменьшает стоимость выполнения запросов и насколько дорого обходится его сопровождение при вставке, обновлении и удалении данных. В самой модели это состояние представлялось через признаки, показывающие наличие или отсутствие индексов в текущей конфигурации [12].

В NoDBA описание состояния стало более содержательным: рабочая нагрузка задавалась матрицей селективностей по столбцам, а текущая конфигурация – двоичным вектором уже созданных индексов. То есть агент видел не только то, какие индексы существуют, но и то, насколько сильно разные запросы отбирают данные по различным столбцам. Исторически это был важный шаг, поскольку в одном описании впервые были объединены характеристики запросов и характеристики физической организации данных. Однако и такой подход оставался жестко связанным с заранее известным набором таблиц и столбцов, поэтому не переносился на новые базы данных с другой структурой [13].

SMARTIX логичнее рассматривать не как следующий шаг к более полному описанию состояния, а как отдельную упрощённую линию исследований. В этой работе состояние задаётся двоичным вектором по всем индексируемым столбцам базы данных: для каждого столбца фиксируется, существует на нём индекс или нет [14]. Иными словами, агент видит текущую конфигурацию индексирования на уровне отдельных столбцов, но не получает в самом состоянии подробного описания рабочей нагрузки. По сравнению с NoDBA это более простое представление: там в состояние входили не только уже созданные индексы, но и характеристики запросов, выраженные через селективности по столбцам [13]. Поэтому значение SMARTIX состоит не в усложнении состояния среды, а в непрерывном управлении конфигурацией индексов с действиями создания и удаления, тогда как само описание состояния остаётся предельно компактным.

Следующий этап развития связан с тем, что в состояние начали включать более подробные сведения о том, как именно запросы обращаются к данным. В DRLindex и ряде близких работ учитывалась не только информация о наличии индексов, но и признаки обращений к данным: какие таблицы и столбцы используются в запросах, по каким столбцам выполняются фильтрация и соединение, насколько избирательны условия отбора, а в системах с репликацией – на каких репликах размещаются данные [15, 16]. Такое описание состояния уже ближе к реальной задаче, поскольку агент получает информацию не только о текущем наборе индексов, но и о том, почему тот или иной индекс может быть полезен.

Часть работ решает ту же проблему иначе – не за счет усложнения состояния, а за счет предварительного ограничения множества рассматриваемых индексов. В работе [4] состояние задается относительно заранее подготовленного набора кандидатов: для каждого такого кандидата фиксируется, построен он или нет. Само состояние при этом получается достаточно простым, поскольку агенту не нужно анализировать все теоретически возможные индексы. Однако эта простота достигается тем, что значительная часть работы выносится на предыдущий этап: система заранее формирует разумный набор кандидатов по эвристическим правилам, в том числе для составных индексов. По существу, здесь действует следующий компромисс: чем сильнее пространство возможных индексов сокращено до начала обучения, тем проще можно описать состояние среды, но тем выше зависимость итогового качества от качества предварительного отбора.

MANTIS [17] и DRLISA [18] важны скорее не как работы с принципиально новым описанием состояния, а как работы, расширяющие саму задачу автоматического индексирования: в первом случае – за счет разделения выбора типа и выбора индекса, во втором – за счет перехода к NoSQL, настройке параметров индекса и работе при изменяющейся нагрузке.

В Indexer++ описание состояния становится еще ближе к реальной эксплуатации [19]. Система анализирует не один фиксированный набор запросов, а последовательность недавних окон рабочей нагрузки.

Это позволяет учитывать не только текущее состояние, но и изменение характера запросов во времени. Если новые окна нагрузки похожи на предыдущие, конфигурация индексов может не пересматриваться. Если же распределение запросов заметно меняется, состояние считается существенно изменившимся, и система получает основание для пересмотра текущего набора индексов. Таким образом, здесь состояние описывает уже не просто «снимок» системы в один момент времени, а недавнюю историю ее поведения.

В наиболее поздних работах основной акцент смещается на переносимость и обобщение. Исследователи стремятся построить такое описание состояния, которое не зависело бы слишком жестко от конкретных имен, таблиц, числа столбцов и заранее фиксированной схемы базы данных. В SWIRL для этого используется компактное описание рабочей нагрузки и метаданных [8]. В IA2 состояние строится вокруг связи между текущей нагрузкой и конкретными индексами-кандидатами [10]. Система оценивает не базу данных в целом, а то, насколько данный кандидат может быть полезен именно для текущего набора запросов. В AutoIndexer эта идея развивается дальше: описание рабочей нагрузки предварительно сжимается, дополняется сведениями из оптимизатора запросов и только после этого используется для принятия решения [11].

В целом можно выделить устойчивую тенденцию. Ранние работы описывали состояние в основном через конфигурацию индексов. Затем в состояние начали включать свойства запросов и статистику данных. Далее появились подходы, в которых описание строится уже не для всей базы данных сразу, а относительно окна рабочей нагрузки. Появились подходы, решающие не только задачу построения индекса, но и смежные задачи, такие как выбор типа индекса.

Пространство действий и гранулярность управления

Эволюция работ в этой области хорошо видна по тому, как меняется множество действий агента. Например, в NoDBA агент на каждом шаге мог только создать индекс на одном столбце. Индексы добавлялись последовательно, пока их число не достигало заранее заданного предела, после чего эпизод завершался, все индексы удалялись, и обучение начиналось заново. Главное ограничение такой постановки в том, что агент не мог ни удалить неудачный индекс, ни пропустить ход, ни сохранить и постепенно улучшать конфигурацию. Следующий этап развития связан с расширением набора допустимых действий. В работах [4] и в SWIRL агент по-прежнему выбирает индекс-кандидат для создания, но пространство выбора предварительно сужается за счёт генерации кандидатов и исключения недопустимых действий [4, 8]. Иными словами, здесь прогресс связан с сокращением пространства поиска. Отдельная линия развития проявляется в SMARTIX, где агент уже может не только создавать, но и удалять индекс [14]. В более поздних постановках того же направления появляется и явное действие пропуска шага. Наконец, в DRLISA пространство действий расширяется ещё дальше: система допускает изменение структуры индекса и его параметров, т.е. переходит от простого выбора очередного индекса к управлению текущей индексной конфигурацией [18]. В работе MANTIS рассматриваются уже не только составные индексы, но и несколько типов индексов в пределах заданного бюджета [17].

Награда и способы оценки эффекта индексирования

Функция награды, т.е. правило, по которому агент получает оценку полезности выбранного действия, является одним из ключевых элементов индексирования на основе обучения с подкреплением. В рассмотренных исследованиях можно выделить три основные группы способов задания награды. Первая группа связывает награду с уменьшением суммарной стоимости выполнения рабочей нагрузки. Так, в NoDBA качество конфигурации индексов оценивается по сравнению с исходным вариантом без индексов [13]. В работе [7] награда определяется относительным снижением стоимости рабочей нагрузки после добавления очередного индекса. В MANTIS также используется выигрыш по стоимости, но одновременно учитываются ограничения на доступный объем памяти [17]. Преимущество таких схем состоит в их понятности: агент поощряется за ускорение выполнения запросов. Однако их слабое место заключается в том, что они опираются на модель стоимости, а она не всегда точно отражает поведение реальной системы.

Вторая группа подходов использует награду, основанную не на оценке, а на фактической производительности. Наиболее показательной здесь является работа SMARTIX, где награда вычисляется по метрике QphH@Size, т.е. на основе реального выполнения тестовой нагрузки [14]. Такой подход дает более надежную обратную связь, поскольку агент ориентируется не на прогноз оптимизатора, а на реальные метрики системы. Вместе с тем его практическое применение затруднено высокой ценой такого измерения: для получения награды необходимо запускать нагрузку и наблюдать результат, что слишком дорого для оперативного управления индексами.

В DRLISA награда учитывает два эффекта одновременно – насколько возрастает число обрабатываемых запросов и насколько увеличиваются издержки на вставку, обновление и удаление данных [18]. Это делает постановку более практичной, поскольку индекс должен не только ускорять чтение, но и не создавать чрезмерные нагрузки при вставке и обновлении данных.

Третья группа объединяет подходы, в которых награда учитывает не только выигрыш в производительности, но и затраты ресурсов. В работе Welborn et al. награда определяется через соотношение между снижением стоимости запросов и ростом пространства рассматриваемых решений [20].

В SWIRL выигрыш от индекса сопоставляется с увеличением занимаемого им объема памяти [8]. В BAIT оптимизируется не только качество найденной конфигурации индексов, но и эффективность использования бюджета на её поиск. При этом бюджет измеряется числом вызовов оценочного интерфейса оптимизатора, т.е. числом оценок стоимости для пар «запрос – конфигурация индексов» без реального построения индексов [9].

Почти все существующие работы сталкиваются с одной и той же фундаментальной трудностью: в задачах индексирования награда обычно является отложенной и неоднозначной. Индекс не начинает влиять на систему мгновенно: сначала он должен быть физически построен, затем его эффект может проявиться лишь для части запросов. Кроме того, наблюдаемое изменение времени выполнения запросов не всегда связано именно с новым индексом. Оно может быть вызвано конкуренцией за ресурсы, блокировками, изменением статистик или другими внешними факторами. Во многих исследовательских работах эта проблема упрощается: предполагается, что после каждого действия можно почти сразу получить корректную и однозначно интерпретируемую награду. Для академического эксперимента такое допущение допустимо, однако при переходе к промышленному автоматическому индексированию именно здесь возникает одно из главных расхождений между теоретической моделью и реальной эксплуатацией.

Архитектуры агентов и стратегии обучения

Архитектуры агентов в работах по автоматическому индексированию развивались поэтапно. Сначала использовались методы без глубоких нейронных сетей. Затем появились модели, в которых нейронная сеть стала использоваться для выбора индексов. После этого основной линией развития стало глубокое Q-обучение, а в более поздних работах начали применяться и другие алгоритмы: методы, напрямую обучающие стратегию выбора действий, методы с двумя взаимосвязанными моделями выбора и оценки, а также поиск по дереву [12].

Первый этап представлен работами, такими как rCOREIL. В них агент не использует глубокую нейронную сеть. Вместо этого для каждой конфигурации индексов заранее задается набор признаков, а полезность конфигурации и возможных переходов оценивается как линейная комбинация этих признаков. Обучение в таких работах состоит в том, что по наблюдаемым переходам система постепенно уточняет веса при признаках и тем самым улучшает правило выбора действий [12].

NoDBA относится к ранним работам по глубокому обучению с подкреплением для индексирования, в качестве обучающего алгоритма используется CEM. Для повышения устойчивости обучения авторы используют повторное использование накопленного опыта [13].

Основная группа последующих работ использует глубокое Q-обучение. В этой схеме нейронная сеть получает описание текущего состояния базы данных и оценивает ожидаемую полезность каждого допустимого действия. Затем агент выбирает действие с наибольшей оценкой. В [4] эта схема сочетается с эвристической генерацией кандидатов, что позволяет применять глубокое Q-обучение не только к одноколоночным, но и к составным индексам. В MANTIS задача разбивается на два этапа: одна модель по рабочей нагрузке определяет тип индекса, а сеть глубокого Q-обучения выбирает конкретный индекс в рамках этого типа [17]. В DRLISA также используется схема обучения по Q-функции, но уже в более широкой постановке, где допускается изменение конфигурации индекса и его параметров [18].

По мере роста числа кандидатов одной только замены линейной модели на нейронную сеть оказалось недостаточно. Поэтому вместе с глубоким Q-обучением стали использоваться дополнительные приемы, которые уменьшают число рассматриваемых действий и делают обучение устойчивее. В [4] эту роль играет предварительная генерация кандидатов. В DQN-AMTAS пространство действий организуется иерархически, а из рассмотрения исключаются недопустимые действия [21]. Тем самым развитие идет не только по линии изменения алгоритма обучения, но и по линии сокращения пространства поиска [4, 21].

После этого в литературе появляются работы, использующие другие алгоритмы. В SWIRL применяется метод, в котором обучается сама стратегия выбора действий, а не только оценка их полезности [8].

В BAIT обучение с подкреплением используется не для прямого выбора индекса, а для управления поиском: агент распределяет ограниченный бюджет вызовов механизма гипотетической оценки индексов и решает, какие конфигурации следует проверить следующими [9]. Таким образом, развитие архитектур агентов идет от ранних методов без глубоких нейронных сетей к моделям с нейронной сетью, затем к глубокому Q-обучению и далее к другим алгоритмам, лучше приспособленным к большим пространствам действий и сложным ограничениям задачи [3, 4, 7–21].

Экспериментальная методология и ограничения практического внедрения

Почти все рассмотренные работы проверяют свои идеи на стандартных тестовых наборах запросов. Чаще всего для этого используются TPC-H и TPC-DS; в задачах, связанных с оперативной обработкой и нереляционными системами, применяют YCSB и наборы на основе IMDB/JOB или специально сгенерированные нагрузки [8–11, 14, 17–19, 22]. Такой подход имеет очевидное достоинство: результаты можно воспроизвести и сопоставить между собой. Однако одновременно он задает и существенное ограничение. Во многих исследованиях обучение и проверка проводятся на сравнительно небольшом числе запросов, на статических или частично искусственных нагрузках и в условиях, где неудачное действие агента не приводит к серьезным последствиям для системы. Поэтому метод, который показывает хорошие результаты в лабораторном эксперименте, не гарантирует надежность при длительной работе на

реальной производственной нагрузке, где одновременно действуют жесткие требования к времени отклика, пропускной способности и стабильности сервиса. По этой причине в современных работах все заметнее смещается акцент от полного многократного перестроения конфигурации индексов к более реалистичным сценариям. Исследователи используют предварительное сжатие и анализ рабочей нагрузки, безопасный отбор кандидатов, ограниченный бюджет гипотетических проверок, проверку способности метода работать на новых нагрузках и гибридные схемы, в которых реальные изменения в рабочей базе данных выполняются редко и только после дополнительной верификации [3, 5, 8, 9, 11].

Краткие выводы по анализу существующих исследований

Проведенный обзор показывает, что сильной стороной RL-подходов является способность моделировать индексирование как последовательное принятие решений. Это особенно важно там, где между индексами существуют сложные взаимодействия, а локально выгодное действие может ухудшить дальнейшую конфигурацию. Работы последних лет показывают, что за счет более содержательного описания состояния среды, исключения заведомо недопустимых действий, сжатого описания рабочей нагрузки и поиска с учетом ограниченности вычислительных ресурсов методы обучения с подкреплением способны выходить за рамки простых жадных схем и находить конкурентоспособные конфигурации индексов [8–11].

В то же время почти все исследовательские линии имеют общие слабые места. Во-первых, недостаточно явно учитывается эксплуатационная безопасность: работа с индексом рассматривается как незатратное действие, хотя в реальной СУБД – это дорогостоящая физическая операция. Во-вторых, отложенная награда обычно упрощается до мгновенной оценки после шага. В-третьих, многие представления состояния остаются либо слишком бедными, либо слишком привязанными к конкретной схеме и набору кандидатов [3, 8, 14, 18].

Следовательно, перспективный путь состоит в том, чтобы сделать агента наблюдательнее и безопаснее: опираться на состояния, ограниченные оптимизируемым запросом, раздельно учитывать суррогатную и реальную награду, ограничивать DDL-операции и применять обучение с подкреплением не как автономный механизм, а как компонент принятия решений в составе более широкой системы, основанной на сборе и обработке эксплуатационной телеметрии.

Запрос-ориентированный метод автоматического индексирования

Метод предназначен для автоматического управления индексной конфигурацией базы данных на основе наблюдаемой рабочей нагрузки. Решение последовательно принимается в рамках конкретного оптимизируемого запроса. В каждом шаге выбирается запрос, дающий наибольший вклад в текущую стоимость нагрузки, вокруг него формируется локальный контекст связанных запросов, после чего происходит построение множества индексов кандидатов и для каждого индексного кандидата выбирается действие CREATE, DROP или SKIP. Такая постановка уменьшает размер пространства решений и при этом учитывает влияние индекса не только на оптимизируемый запрос, но и на другие запросы, которые обращаются к тем же объектам и используют те же столбцы.

Рабочая нагрузка в момент времени t рассматривается в окне наблюдения $\mathbf{W}_t$. Для запроса q используется оценка стоимости оптимизатора $C_{\text{опт}}(q,t)$. Чтобы выбирать запросы с учетом количества вызовов, вводится нагрузочный вес запроса

$$H(q,t) = \tilde{C}_{\text{опт}}(q,t) \cdot \log(1+ f(q,t)), \qquad (4)$$

где $f(q,t)$ – число выполнений запроса q в окне $\mathbf{W}_t$, а $\tilde{C}_{\text{опт}}(q,t)$ – нормированная оценочная стоимость оптимизатора. Нормирование числовых признаков выполняется внутри окна наблюдения по логарифмической шкале

$$\tilde{x} = \frac{\log(1+x)}{\max\limits_{y \in \mathbf{X}_t} \log(1+y) + \varepsilon}. \qquad (5)$$

Фактические метрики выполнения – логические чтения, процессорное время, длительность, ожидания и признаки сканирований – включаются в состояние агента и используются при расчете фактической награды после совершения действий.

Выбор запроса, локального контекста и индексных кандидатов

Вначале выбирается запрос с максимальным весом. Такой запрос рассматривается как инициатор цикла оптимизации посредством построения индекса. После чего оценка индексного действия формируется на основе локального подмножества рабочей нагрузки, связанного с теми же объектами:

$$q_t^* = \arg\max_{q \in \mathbf{W}_t} H(q,t). \qquad (6)$$

Для запроса q_t^* строится локальный контекст $\mathbf{L}_t(q_t^*)$, содержащий запросы из текущего окна наблюдения, которые обращаются к тем же базовым объектам. Контекст нужен для того, чтобы индексное решение оценивалось по группе связных запросов, работающих с одинаковыми объектами:

$$\mathbf{L}_t(q_t^*) = \left\{ q \in_t \mathbf{W} \middle| \mathbf{O}(q) \cap \mathbf{O}(q_t^*) \neq \varnothing \right\}, \qquad (7)$$

где $\mathbf{O}(q)$ – множество базовых объектов, к которым обращается запрос q. На основе инициирующего запроса, локального контекста и текущей индексной конфигурации $\mathbf{I}_t$ формируется множество кандидатов

$$\mathbf{C}_t = G(q_t^*, \mathbf{L}_t(q_t^*), \mathbf{I}_t). \qquad (8)$$

Здесь G – генератор новых индексных структур. Он формирует кандидаты по столбцам, участвующим в аргументах поиска, предикатах фильтрации, соеди-

нениях, сортировке и группировке. В множество $\mathbf{C}_t$ также включаются существующие индексы $\mathbf{I}_t$.

Каждый кандидат c описывает индексную структуру и имеет признак присутствия $p(c,t)\in\{0,1\}$: $p(c,t)=1$ означает, что такая структура уже существует в конфигурации $\mathbf{I}_t$, а $p(c,t)=0$ означает, что она является потенциальным новым индексом. Для каждого кандидата используется единое пространство действий

$$\mathbf{A}=\{\text{CREATE},\text{DROP},\text{SKIP}\}. \quad (9)$$

Действие CREATE означает переход к конфигурации, в которой кандидат присутствует; действие DROP – переход к конфигурации, в которой кандидат отсутствует; действие SKIP сохраняет текущую конфигурацию. Формально целевое значение признака присутствия после действия задается как

$$p^a(c,t+1)=\begin{cases}1, & a=\text{CREATE},\\ 0, & a=\text{DROP},\\ p(c,t), & a=\text{SKIP}.\end{cases} \quad (10)$$

Такая запись позволяет использовать одно и то же множество действий для новых и уже существующих индексных структур.

Множество затронутых запросов и кодирование состояния

Для оценки конкретного кандидата используется не весь локальный контекст, а только те запросы, на которые данный кандидат потенциально способен повлиять. Для кандидата c вводится множество смежных запросов:

$$\mathbf{L}(c,t)=\{q\in\mathbf{W}_t\,|\,\mathbf{O}(q)\cap\mathbf{O}(c)\neq\varnothing\wedge\mathbf{R}(q,c)\neq\varnothing\}, \quad (11)$$

где $\mathbf{O}(c)$ – объект или объекты, для которых определен кандидат c, а $\mathbf{R}(q,c)$ – множество ролей столбцов кандидата в запросе q. В $\mathbf{R}(q,c)$ включаются признаки того, что столбец используется как аргумент поиска, остаточный предикат, столбец соединения, сортировки, группировки или покрытия результата. Если запрос обращается к тому же объекту, но не использует столбцы кандидата ни в одной из этих ролей, он не включается в $\mathbf{L}(c,t)$.

Состояние среды для кандидата c задается вектором признаков

$$\mathbf{s}(c,t)=[\mathbf{x}_q,\mathbf{x}_\text{L},\mathbf{x}_\text{T},\mathbf{x}_\text{I},\mathbf{x}_c,\mathbf{x}_\text{E}]. \quad (12)$$

Блок $\mathbf{x}_q$ описывает инициирующий запрос q_t^*: оценочную стоимость оптимизатора, логические чтения, процессорное время, полную длительность, частоту выполнений, наличие сортировок, выгрузок во временные структуры и сканирований. Ожидания кодируются по нескольким устойчивым классам: блокировки, ввод-вывод, ожидание процессора и сетевые ожидания. Все временные величины выражаются в миллисекундах и нормируются внутри окна наблюдения.

Блок $\mathbf{x}_\text{L}$ содержит агрегированные характеристики множества $\mathbf{L}(c,t)$: суммарную частоту выполнений, суммарную оценочную стоимость, средние и максимальные значения логических чтений, процессорного времени и длительности, а также доли запросов, для которых кандидат может использоваться как индекс поиска, как покрывающий индекс или как индекс, сохраняющий требуемый порядок для сортировки либо группировки.

Блок $\mathbf{x}_\text{T}$ описывает объект данных: размер таблицы в байтах, число строк, интенсивность чтения, а также интенсивность операций INSERT, UPDATE и DELETE. В методе различаются общая интенсивность обращений к объекту и DML-интенсивность: первая учитывает все запросы к объекту, вторая – только операции модификации данных.

Блок $\mathbf{x}_\text{I}$ описывает текущую индексную конфигурацию на объекте: число существующих индексов, их суммарный размер, наличие индекса с тем же ведущим ключевым столбцом, максимальную длину совпадающего ключевого префикса, долю столбцов кандидата, уже покрытых существующими индексами, и признак существования почти дублирующей структуры.

Блок $\mathbf{x}_c$ описывает сам кандидат: число ключевых столбцов, число включенных столбцов, предполагаемый размер, оценочную стоимость построения, ожидаемый прирост стоимости операций записи, признак покрытия инициирующего запроса и признак сохранения порядка данных. Для существующего индекса эти признаки вычисляются по его текущему определению; для нового кандидата – по его предполагаемому определению.

Блок $\mathbf{x}_\text{E}$ содержит эксплуатационные условия, влияющие на допустимость изменения: доступное место, оценочную нагрузку на журнал транзакций, текущий уровень параллелизма, наличие конкурирующих операций изменения структуры и признак допустимого эксплуатационного окна. Эти признаки входят в состояние агента, но окончательное разрешение на изменение все равно выдается отдельным контуром безопасного применения.

Для запросов из $\mathbf{L}(c,t)$ дополнительно определяется степень пригодности кандидата:

$$\psi(q,c)=\begin{cases}0, & \text{нет изменения доступа},\\ 1, & \text{поиск без покрытия},\\ 2, & \text{покрытие запроса},\\ 3, & \text{покрытие и порядок}.\end{cases} \quad (13)$$

Уровень 0 означает, что кандидат не меняет способ доступа; уровень 1 – что кандидат позволяет выполнить поиск по ключу, но требует последующего обращения к данным таблицы; уровень 2 – что кандидат покрывает запрос; уровень 3 – что кандидат одновременно покрывает запрос и сохраняет требуемый порядок для сортировки или группировки. На основе этой шкалы задается вес запроса при оценке кандидата

$$w(q,c,t)=\frac{f(q,t)C_{\text{опт}}(q,t)(1+\psi(q,c))}{\sum_{u\in\mathbf{L}(c,t)}f(u,t)C_{\text{опт}}(u,t)(1+\psi(u,c))+\varepsilon}. \quad (14)$$

Чтобы выявлять дублирование и близкие замены существующих индексов, вычисляется мера структурной близости кандидата к текущей конфигурации

$$D(c,\mathbf{I}_t)=\max_{i\in\mathbf{I}_t}\left(\alpha_{\text{p}}P(c,i)+\alpha_{\text{вкл}}\,\text{Inc}(c,i)+\alpha_{\text{o}}\,\text{Ord}(c,i)\right), \quad (15)$$

где $P(c,i)$ – доля совпадающего ключевого префикса; $\text{Inc}(c,i)$ – доля включенных столбцов кандидата, уже присутствующих в индексе i, а $\text{Ord}(c,i)$ – совместимость порядка ключевых столбцов с требованиями сортировки и группировки. Значение $D(c,\mathbf{I}_t)$ используется как признак состояния и как возможный штраф при создании почти дублирующего индекса.

Суррогатная и фактическая награда

Индексное действие сначала оценивается в гипотетической конфигурации, а затем проверяется по фактической статистике. Пусть $\mathbf{I}_t^{c,a}$ – конфигурация, полученная из $\mathbf{I}_t$ после применения действия a к кандидату c. Для каждого запроса $q\in\mathbf{L}(c,t)$ оптимизатор оценивает стоимость выполнения в конфигурациях $\mathbf{I}_t$ и $\mathbf{I}_t^{c,a}$. Относительный ожидаемый выигрыш определяется как

$$\Delta\hat{C}(q,c,a,t)=\frac{\hat{C}(q,\mathbf{I}_t)-\hat{C}(q,\mathbf{I}_t^{c,a})}{\max(\hat{C}(q,\mathbf{I}_t),\varepsilon)}. \quad (16)$$

Суррогатная награда действия задается выражением

$$\hat{r}(c,a,t)=\sum_{q\in\mathbf{L}(c,t)}w(q,c,t)\Delta\hat{C}(q,c,a,t)-\lambda_{\text{m}}\Delta M(c,a,t)-$$
$$-\lambda_{\text{s}}\Delta S(c,a,t)-\lambda_{\text{b}}B(c,a,t). \quad (17)$$

Здесь $\Delta M(c,a,t)$ – ожидаемое изменение стоимости операций INSERT, UPDATE и DELETE; $\Delta S(c,a,t)$ – изменение занимаемого пространства; $B(c,a,t)$ – стоимость и операционный риск физического изменения структуры. Величины ΔM и ΔS имеют знак: при создании индекса они обычно положительны, а при удалении индекса могут быть отрицательны, если действие снижает стоимость записи и освобождает место. Для минимально значимого выигрыша $G_{\min}=0{,}10$, т.е. 10% взвешенного относительного снижения оценочной стоимости затронутых запросов по сравнению с текущей конфигурацией, допустимое относительное увеличение стоимости операций записи M принимается равным 10%, а бюджет дополнительного пространства S равным 20% от суммарного начального размера данных таблиц. Показатель стоимости и риска B нормируется к интервалу ([0;1]) по фиксированной шкале. Коэффициенты штрафов соответственно рассчитываются по общему правилу

$$\lambda_j=\frac{G_{\min}}{L_j},\quad j\in m,s,b.$$

После применения изменения система переходит в режим наблюдения. Фактическая награда вычисляется не сразу, а после накопления не менее N выполнений затронутых запросов. Момент расчета определяется как

$$\tau_N=\min\{\tau>t:\sum_{q\in\mathbf{L}(c,t)}n_q(t,\tau)\ge N\}, \quad (18)$$

где $n_q(t,\tau)$ – число выполнений запроса q на интервале от t до τ. Если часть запросов из $\mathbf{L}(c,t)$ не выполнялась в период наблюдения, фактическая награда рассчитывается по выполненным запросам с весами, пересчитанными по фактическим частотам.

Для выполненного запроса q используются относительные изменения длительности T, процессорного времени P и логических чтений R

$$\Delta T_q=\frac{T_q^{\text{до}}-T_q^{\text{после}}}{\max(T_q^{\text{до}},\varepsilon)},$$
$$\Delta P_q=\frac{P_q^{\text{до}}-P_q^{\text{после}}}{\max(P_q^{\text{до}},\varepsilon)},$$
$$\Delta R_q=\frac{R_q^{\text{до}}-R_q^{\text{после}}}{\max(R_q^{\text{до}},\varepsilon)}. \quad (19)$$

Фактическая награда имеет вид

$$r(c,a,\tau_N)=$$
$$=\sum_{q\in\mathbf{L}_N(c,t)}\bar{w}(q,c,\tau_N)\kappa_q\left(\beta_{\text{T}}\Delta T_q+\beta_P\Delta P_q+\beta_R\Delta R_q\right)-$$
$$-\mu_{\text{m}}\Delta M^{\text{факт}}(c,a)-\mu_{\text{s}}\Delta S^{\text{факт}}(c,a)-\mu_{\text{b}}B^{\text{факт}}(c,a). \quad (20)$$

Здесь $\mathbf{L}_N(c,t)$ – подмножество затронутых запросов, реально выполненных за период наблюдения, $\bar{w}(q,c,\tau_N)$ – веса, пересчитанные по фактическим частотам, а коэффициент $\kappa_q\in[0,1]$ снижает вклад запроса, если изменение его длительности в основном связано не с индексом, а с внешними ожиданиями: блокировками, насыщением дисковой подсистемы, ожиданием процессора или сетевыми задержками. После появления фактической награды она заменяет суррогатную награду в обучающей выборке. В фактической награде используются те же коэффициенты, что и для суррогатной награды: $\mu_m=\lambda_m$, $\mu_s=\lambda_s$, $\mu_b=\lambda_b$.

Обучающий агент

Задача выбора индексного действия рассматривается как повторяемая задача принятия решений в

изменяющемся состоянии рабочей нагрузки и текущей индексной конфигурации. На вход агенту подается состояние $\mathbf{s}(c,t)$, описывающее кандидата, затронутые запросы, объект данных, текущие индексы и эксплуатационные условия. На выходе агент выбирает одно из трех действий: CREATE, DROP или SKIP.

В качестве обучающей модели используется двойная дуэльная глубокая Q-сеть (англ. Double Dueling Deep Q-Network, Double Dueling DQN) . Дуэльная архитектура применяется потому, что для индексного кандидата полезно раздельно оценивать общую перспективность состояния и относительное преимущество конкретного действия. Оценка Q-функции задается как

$$Q(\mathbf{s},a;\theta)=V(\mathbf{s};\theta_{\mathrm{V}})+ \\ +\left(A(\mathbf{s},a;\theta_{\mathrm{A}})-\frac{1}{|\mathbf{A}|}\sum_{a'\in\mathbf{A}}A(\mathbf{s},a';\theta_{\mathrm{A}})\right). \qquad (21)$$

Здесь $V(\mathbf{s};\theta_{\mathrm{V}})$ – оценка ценности состояния, а $A(\mathbf{s},a;\theta_{\mathrm{A}})$ – оценка преимущества действия a относительно других действий в том же состоянии. Такая архитектура полезна, когда часть кандидатов заведомо малоперспективна независимо от выбранного действия, например, из-за малого размера таблицы, высокой DML-интенсивности или почти полного дублирования существующей структуры.

Двойная схема используется для уменьшения завышения оценок Q при обучении на шумных данных. Гипотетические оценки оптимизатора могут отличаться от фактического эффекта, а рабочая нагрузка может изменяться между моментом выбора действия и моментом расчета фактической награды. Поэтому выбор следующего действия и оценка его целевого значения разделяются:

$$y_t=r_t+\gamma Q\left(\mathbf{s}_{t+1},a^*,\theta_t^-\right), \\ a^*=\arg\max_{a'\in\mathbf{A}}Q\left(\mathbf{s}_{t+1},a';\theta_t\right), \qquad (22)$$

где θ_t – параметры основной сети; θ_t^- – параметры целевой сети; γ – коэффициент дисконтирования. Обучение выполняется по накопленным переходам вида

$$(\mathbf{s}_t,a_t,\hat{r}_t,\mathbf{s}_{t+1})\to(\mathbf{s}_t,a_t,r_t,\mathbf{s}_{t+1}), \qquad (23)$$

т.е. сначала в буфер помещается переход с суррогатной наградой, а после завершения периода наблюдения эта награда заменяется фактической. Агент не выполняет DDL-операции напрямую. Его результат является рекомендацией, которая передается в контур безопасного применения.

Безопасное применение индексных изменений

Рекомендация агента применяется только после проверки эксплуатационных ограничений. Контур безопасного применения проверяет доступное место, допустимую нагрузку на журнал транзакций, наличие конкурирующих DDL-операций, текущее состояние блокировок, допустимый уровень параллелизма и наличие эксплуатационного окна. Если хотя бы одно ограничение нарушено, действие не применяется, а система переходит к наблюдению без изменения индексной конфигурации.

Физическое изменение конфигурации выполняется не чаще одного раза за цикл. После CREATE или DROP система не принимает новое индексное решение до завершения периода наблюдения и накопления достаточного числа выполнений затронутых запросов. Это ограничение нужно для того, чтобы фактическая награда могла быть связана с конкретным действием, а не с несколькими одновременными изменениями.

Построение индекса выполняется в онлайн-режиме, если такая возможность поддерживается используемой СУБД. При отсутствии безопасного режима выполнения контур применения отклоняет действие или переносит его до подходящего эксплуатационного окна. В результате предложенный метод не является автономным механизмом бесконтрольного изменения схемы. Обучающий агент решает задачу ранжирования индексных действий в локальном контексте нагрузки, а фактическое изменение конфигурации выполняется только через безопасный контур. За счет этого метод сочетает адаптацию к текущей рабочей нагрузке, учет реального эффекта индекса и эксплуатационные ограничения промышленной СУБД.

Результаты

Для оценки производительности стратегии индексирования использовался тестовый набор TPC-H. В качестве системы управления базами данных был выбран Microsoft SQL Server 2022. Цель эксперимента состояла в сравнительной оценке трёх стратегий индексирования в одинаковых условиях: базовой индексации, отсутствия индексов и стратегии, сформированной методом обучения с подкреплением.

Рассматривались три сценария нагрузки. Первый сценарий соответствовал последовательному выполнению 22 аналитических запросов TPC-H в одном потоке.

Второй сценарий моделировал конкурентную нагрузку: два потока параллельно выполняли аналитические запросы, а третий поток выполнял стандартные операции обновления данных RF1 и RF2.

Третий сценарий представлял собой дополнительную синтетическую нагрузку из 100 SQL-запросов к тем же данным. Она использовалась для проверки поведения стратегий индексирования за пределами стандартного набора запросов TPC-H. На рис. 2 представлено сравнение трёх стратегий индексирования при последовательном выполнении 22 запросов TPC-H по суммарному времени выполнения и метрике Power@Size.

Метрика Power@Size используется для оценки последовательного однопоточного выполнения запросов. Времена выполнения 22 запросов агрегируются с использованием геометрического среднего, после чего результат нормируется.

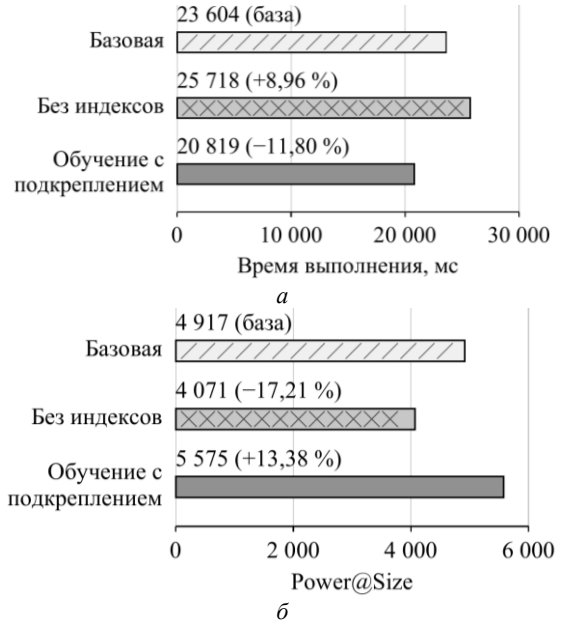


Рис. 2. Результаты последовательного выполнения 22 аналитических запросов TPC-H: *а* – суммарное время выполнения; *б* – метрика Power@Size

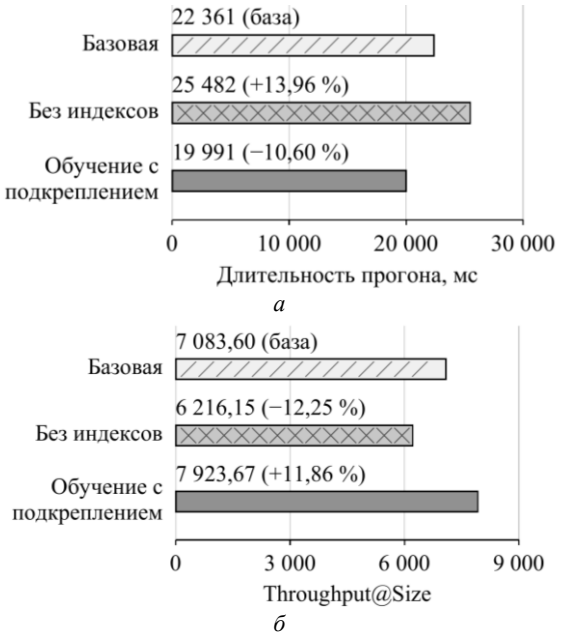


Рис. 3. Результаты выполнения конкурентной нагрузки с двумя потоками аналитических запросов и потоком операций обновления RF1 и RF2: *а* – длительность прогона; *б* – метрика Throughput@Size

Этот показатель используется как сравнительная оценка сценариев индексирования.

В первом сценарии при базовой индексации набор из 22 запросов был выполнен за 23604 мс. Среднее время выполнения одного запроса составило 1072,91 мс, а значение интегральной метрики Power@Size – 4917. В сценарии без индексов суммарное время увеличилось до 25718 мс, среднее время запроса – до 1169,00 мс, а значение Power@Size снизилось до 4071.

Таким образом, отказ от индексов в режиме только чтения привёл к ухудшению на 8,96% по общему времени выполнения и на 17,21% по Power@Size. При использовании стратегии обучения с подкреплением общее время выполнения составило 20819 мс, среднее время запроса – 946,32 мс, а значение Power@Size – 5575. Это соответствует улучшению на 11,80% по времени выполнения и на 13,38% по Power@Size относительно базовой индексации.

Метрика Throughput@Size характеризует способность системы обрабатывать конкурентную нагрузку. Она рассчитывается по результатам многопоточного испытания как количество выполненных запросов в час с учётом числа параллельных потоков, общего времени выполнения испытания и масштаба набора данных.

Во втором сценарии при базовой индексации общее время прогона составило 22361 мс. На рис. 3 представлено сравнение трёх стратегий при конкурентном выполнении аналитических запросов и операций обновления по длительности выполнения и метрике Throughput@Size.

Поскольку в испытании использовались два параллельных потока чтения, среднее время одного аналитического выполнения рассчитывалось по 44 выполненным запросам и составило 508,20 мс. Итоговая метрика Throughput@Size составила 7083,60.

В сценарии без индексов общее время увеличилось до 25482 мс, среднее время запроса – до 579,14 мс, а значение Throughput@Size снизилось до 6216,15.

Это соответствует ухудшению на 13,96% по общему времени выполнения и на 12,25% по метрике пропускной способности относительно базовой индексации.

Для стратегии обучения с подкреплением общее время выполнения составило 19991 мс, среднее время запроса – 454,34 мс, а значение Throughput@Size – 7923,67. Это соответствует улучшению на 10,60% по времени прогона и на 11,86% по метрике пропускной способности относительно базовой индексации.

В качестве третьего сценария была добавлена дополнительная синтетическая нагрузка из 100 SQL-запросов к тем же данным TPC-H. Нагрузка формировалась шаблонным генератором: были заданы 10 групп запросов, и для каждой группы создано по 10 параметризованных экземпляров, что дало итоговый набор из 100 SQL-запросов. В шаблоны входили фильтры по датам и справочникам, соединения таблиц, агрегирование с GROUP BY, сортировки с выбором первых N строк, подзапросы с EXISTS и NOT EXISTS, а также запросы, в которых соединения порождают большое число связанных строк.

Параметры выбирались из доменов, присутствующих в данных. На рис. 4 представлено сравнение трёх стратегий на синтетической нагрузке из 100 SQL-запросов по суммарному времени выполнения и метрике Power@Size.

Для синтетической нагрузки из 100 запросов при базовой индексации были получены следующие показатели производительности: общее время выполнения – 156 с, среднее время запроса – 1,56 с, условная интегральная метрика Power@Size – 2950.

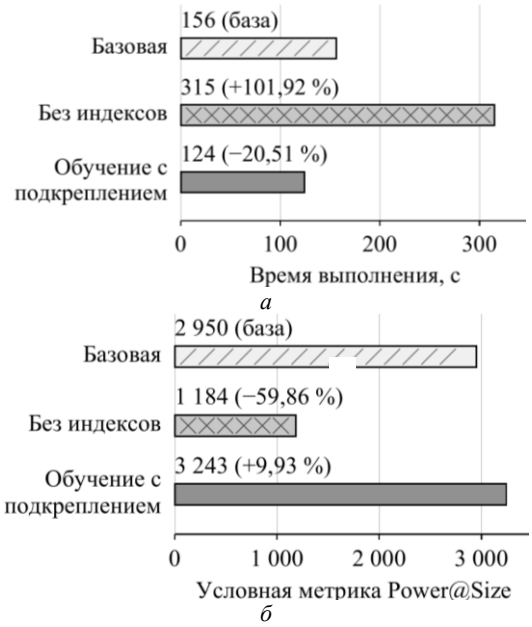


Рис. 4. Результаты выполнения синтетической нагрузки из 100 SQL-запросов: *а* – суммарное время выполнения; *б* – условная интегральная метрика Power@Size

Сценарий без индексов оказался наиболее медленным: общее время выполнения составило 315 с, среднее время запроса – 3,15 с, а значение условной интегральной метрики снизилось до 1184. Таким образом, выполнение без индексов оказалось в 2,02 раза медленнее по времени выполнения и хуже на 59,86% по интегральной метрике относительно базового набора индексов. Для стратегии обучения с подкреплением общее время выполнения составило 124 с, среднее время запроса – 1,24 с, а значение условной интегральной метрики – 3243. Это соответствует улучшению на 20,51% по времени выполнения и на 9,93% по производительности относительно базового набора индексов.

Заключение

В работе рассмотрена задача автоматического индексирования баз данных как задача последовательного управления физической структурой хранения данных с использованием обучения с подкреплением. Показано, что выбор индексной конфигурации не сводится к одноразовой настройке для фиксированной рабочей нагрузки: каждое индексное действие изменяет последующее состояние системы, влияет на планы выполнения запросов, занимаемое пространство и множество дальнейших допустимых решений.

Проведенный анализ существующих исследований показал, что методы обучения с подкреплением учитывают последовательный характер задачи индексирования и выходят за пределы жадных стратегий. Вместе с тем выявлены ограничения большинства рассмотренных подходов: недостаточный учет эксплуатационной безопасности, упрощенное представление отложенной награды, зависимость состояния от конкретной схемы базы данных и слабая связь между гипотетической оценкой индекса и фактическим поведением рабочей системы.

Предложен запрос-ориентированный метод автоматического индексирования. Цикл оптимизации начинается с выбора запроса, вносящего наибольший вклад в текущую стоимость рабочей нагрузки. Для каждого индексного кандидата агент выбирает одно из дискретных действий, а оценка действия строится не только по исходному запросу, но и по затронутым запросам, для которых данный индекс может изменить план выполнения. Особенностью метода является разделение суррогатной и фактической награды: первая вычисляется на основе гипотетической оценки оптимизатора, вторая – по статистике выполненных запросов.

Экспериментальная проверка показала, что стратегия, сформированная методом обучения с подкреплением, превзошла базовую индексацию TPC-H во всех рассмотренных сценариях. В последовательном выполнении 22 аналитических запросов получено улучшение на 11,80% по общему времени выполнения и на 13,38% по метрике Power@Size. В конкурентном сценарии улучшение составило 10,60% по времени прогона и 11,86% по Throughput@Size.

На дополнительной синтетической нагрузке стратегия снизила общее время выполнения на 20,51% и повысила условную интегральную метрику производительности на 9,93% относительно базовой индексации.

Полученные результаты подтверждают, что обучение с подкреплением может использоваться как основа для динамического управления индексной конфигурацией базы данных.

Дальнейшее развитие метода связано с расширением экспериментальной проверки на реальные производственные нагрузки, увеличением длительности наблюдения за изменениями запросов, сравнением с промышленными средствами автоматической настройки индексов и переносом подхода на другие типы СУБД.

Литература

1. Agrawal S. AutoAdmin: Self-Tuning Database Systems Technology / S. Agrawal, N. Bruno, S. Chaudhuri, V. Narasayya // IEEE Data Engineering Bulletin. – 2006. – Vol. 29, No. 3. – P. 7–15.

2. Chaudhuri S. Self-Tuning Database Systems: A Decade of Progress / S. Chaudhuri, V. Narasayya // VLDB '07: Proceedings of the 33rd international conference on Very large data bases. – Vienna Austria: ACM, 2007. – P. 3–14.

3. Wu Y. Automatic Index Tuning: A Survey / Y. Wu, X. Zhou, Y. Zhang, G. Li // IEEE Transactions on Knowledge and Data Engineering. – 2024. – Vol. 36, No. 12. – P. 7657–7676.

4. Lan H. An Index Advisor Using Deep Reinforcement Learning / H. Lan, Z. Bao, Y. Peng // Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM). – New York: ACM, 2020. – P. 2105–2108.

5. Shi J. Learned Index Benefits: Machine Learning Based Index Performance Estimation / J. Shi, G. Cong, X.-L. Li // Proceedings of the VLDB Endowment. – 2022. – Vol. 15, No. 13. – P. 3950–3962.

6. Sutton R.S. Reinforcement Learning: An Introduction / R.S. Sutton, A.G. Barto. – 2nd ed. – Cambridge, MA: MIT Press, 2018. – 552 p.

7. Human-level Control through Deep Reinforcement Learning / V. Mnih, K. Kavukcuoglu, D. Silver et al. // Nature. – 2015. – Vol. 518. – P. 529–533.

8. Kossmann J. SWIRL: Selection of Workload-aware Indexes Using Reinforcement Learning / J. Kossmann, A. Kastius, R. Schlosser // Proceedings of the 25th International Conference on Extending Database Technology (EDBT). – Edinburgh, UK, 2022. – P. 155–168.

9. Budget-aware Index Tuning with Reinforcement Learning / W. Wu, C. Wang, T. Siddiqui et al. // Proceedings of the 2022 International Conference on Management of Data (SIGMOD). – USA: ACM, 2022. – P. 1528–1541.

10. Wang T. IA2: Leveraging Instance-Aware Index Advisor with Reinforcement Learning for Diverse Workloads / T. Wang, E. Yoneki // Proceedings of the 4th Workshop on Machine Learning and Systems (EuroMLSys). – New York: ACM, 2024. – P. 10–17.

11. Wang T. AutoIndexer: A Reinforcement Learning-Enhanced Index Advisor Towards Scaling Workloads / T. Wang, E. Yoneki // arXiv e-prints. – 2025. – Art. no. arXiv:2507. 23084.

12. Basu D. Regularized Cost-Model Oblivious Database Tuning with Reinforcement Learning / D. Basu, Q. Lin, W. Chen et al. // Transactions on Large-Scale Data- and Knowledge-Centered Systems: Lecture Notes in Computer Science. – Cham: Springer, 2016. – Vol. 28. – P. 96–132.

13. Sharma A. The Case for Automatic Database Administration using Deep Reinforcement Learning / A. Sharma, F.M. Schuhknecht, J. Dittrich // arXiv e-prints. – 2018. – Art. no. arXiv:1801.05643.

14. SMARTIX: A Database Indexing Agent Based on Reinforcement Learning / G.P. Licks, J.C. Couto, P. Miehe et al. // Applied Intelligence. – 2020. – Vol. 50, No. 8. – P. 2575–2588.

15. Sadri Z. Online Index Selection Using Deep Reinforcement Learning for a Cluster Database / Z. Sadri, L. Gruenwald, E. Leal // 2020 IEEE 36th International Conference on Data Engineering Workshops (ICDEW). – Dallas, TX: IEEE, 2020. – P. 158–161.

16. Sadri Z. DRLindex: Deep Reinforcement Learning Index Advisor for a Cluster Database / Z. Sadri, L. Gruenwald, E. Leal // Proceedings of the 24th International Database Engineering & Applications Symposium (IDEAS). – Seoul, Republic of Korea: ACM, 2020. – Article 11. – P. 1–8.

17. Sharma V. MANTIS: Multiple Type and Attribute Index Selection using Deep Reinforcement Learning / V. Sharma, C. Dyreson, N. Flann // Proceedings of the 25th International Database Engineering & Applications Symposium (IDEAS). – New York, USA, 2021. – P. 56–64.

18. Yan Y. Index Selection for NoSQL Database with Deep Reinforcement Learning / Y. Yan, S. Yao, H. Wang, M. Gao // Information Sciences. – 2021. – Vol. 561. – P. 20–30.

19. Sharma V. Indexer++: Workload-Aware Online Index Tuning with Transformers and Reinforcement Learning / V. Sharma, C. Dyreson // Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing (SAC). – New York: ACM, 2022. – P. 372–380.

20. Welborn J. Learning Index Selection with Structured Action Spaces / J. Welborn, M. Schaarschmidt, E. Yoneki // arXiv e-prints. – 2019. – Art. no. arXiv:1909.07440.

21. Wu Y. Index Advisor via DQN with Invalid Action Mask in Tree-Structured Action Space / Y. Wu, Y. Zhang, N. Li // Advanced Data Mining and Applications. ADMA 2022. Lecture Notes in Computer Science. – Cham: Springer, 2022. – Vol. 13781. – P. 434–445.

22. Zhou W. Breaking It Down: An In-depth Study of Index Advisors / W. Zhou, C. Lin, X. Zhou, G. Li // Proceedings of the VLDB Endowment. – 2024. – Vol. 17, No. 10. – P. 2405–2418. DOI: 10.14778/3675034.3675035.

Коломиец Виктор Владимирович
Аспирант каф. информационных технологий
на транспорте Сибирского государственного
университета путей сообщения
Дуси Ковальчук ул., 191, г. Новосибирск, Россия, 630049
Тел.: +7-953-777-72-71
Эл. почта: viccont65@gmail.com

Поступила в редакцию: 17.06.2026.
Принята к публикации: 11.09.2026.

Kolomiets V.V.

Automatic database indexing using reinforcement learning

Relevance. Indexes have a significant impact on the performance of database management systems, and their automatic selection under changing workloads remains a challenging multicriteria problem. ***Purpose of the study.*** The purpose of the study is to develop a query-oriented method for automatic index management based on reinforcement learning for operation in a high-load environment. ***Methods.*** The study employs the following tools and approaches: formulation of the problem as a Markov decision process, comparative analysis of modern indexing approaches, the query optimizer's costing mechanism, DBMS operational telemetry, and validation of results on TPC-H workloads. ***Scientific novelty.*** An automatic indexing method based on reinforcement learning is proposed, in which the indexing action is selected in the local context of the query being optimized, while the optimizer's surrogate reward is separated from the actual reward calculated from the statistics of executed queries. ***Results.*** In sequential, concurrent, and synthetic scenarios, the reinforcement learning strategy outperformed baseline indexing in terms of query execution time and aggregate performance metrics. ***Practical significance.*** The proposed method can be used as a component of a safe automatic index-tuning loop in industrial DBMSs.

Keywords: reinforcement learning, automatic indexing, query-oriented index selection, dynamic index tuning, physical database design, query execution statistics, query cost estimation.
DOI: 10.21293/1818-0442-2026-29-2-116-128

References

1. Agrawal S., Bruno N., Chaudhuri S., Narasayya V. AutoAdmin: Self-Tuning Database Systems Technology. IEEE Data Engineering Bulletin, 2006, vol. 29, no. 3, pp. 7–15.

2. Chaudhuri S., Narasayya V. Self-Tuning Database Systems: A Decade of Progress. VLDB '07: Proceedings of the 33rd international conference on Very large data bases. Vienna Austria, ACM, 2007, pp. 3–14.

3. Wu Y., Zhou X., Zhang Y., Li G. Automatic Index Tuning: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, 2024, vol. 36, no. 12, pp. 7657–7676.

4. Lan H., Bao Z., Peng Y. An Index Advisor Using Deep Reinforcement Learning. Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM), New York, ACM, 2020, pp. 2105–2108.

5. Shi J., Cong G., Li X.-L. Learned Index Benefits: Machine Learning Based Index Performance Estimation. *Proceedings of the VLDB Endowment*, 2022, vol. 15, no. 13, pp. 3950–3962.

6. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. 2nd ed. Cambridge, MA, MIT Press, 2018, 552 p.

7. Mnih V., Kavukcuoglu K., Silver D. et al. Human-level Control through Deep Reinforcement Learning. *Nature*, 2015, vol. 518, pp. 529–533.

8. Kossmann J., Kastius A., Schlosser R. SWIRL: Selection of Workload-aware Indexes Using Reinforcement Learning. Proceedings of the 25th International Conference on Extending Database Technology (EDBT), Edinburgh, UK, 2022, pp. 155–168.

9. Wu W., Wang C., Siddiqui T. et al. Budget-aware Index Tuning with Reinforcement Learning. Proceedings of the 2022 International Conference on Management of Data (SIGMOD), USA, ACM, 2022, pp. 1528–1541.

10. Wang T., Yoneki E. IA2: Leveraging Instance-Aware Index Advisor with Reinforcement Learning for Diverse Workloads. Proceedings of the 4th Workshop on Machine Learning and Systems (EuroMLSys), New York, ACM, 2024, pp. 10–17.

11. Wang T., Yoneki E. AutoIndexer: A Reinforcement Learning-Enhanced Index Advisor Towards Scaling Workloads. *arXiv e-prints*, 2025, art. no. arXiv:2507.23084.

12. Basu D., Lin Q., Chen W. et al. Regularized Cost-Model Oblivious Database Tuning with Reinforcement Learning. Transactions on Large-Scale Data– and Knowledge-Centered Systems, *Lecture Notes in Computer Science*, Cham, Springer, 2016, vol. 28, pp. 96–132.

13. Sharma A., Schuhknecht F.M., Dittrich J. The Case for Automatic Database Administration using Deep Reinforcement Learning. *arXiv e-prints*, 2018, art. no. arXiv:1801.05643.

14. Licks G.P., Couto J.C., Miehe P. et al. SMARTIX: A Database Indexing Agent Based on Reinforcement Learning. *Applied Intelligence*, 2020, vol. 50, no. 8, pp. 2575–2588.

15. Sadri Z., Gruenwald L., Leal E. Online Index Selection Using Deep Reinforcement Learning for a Cluster Database. 2020 IEEE 36th International Conference on Data Engineering Workshops (ICDEW), Dallas, TX, IEEE, 2020, pp. 158–161.

16. Sadri Z., Gruenwald L., Leal E. DRLindex: Deep Reinforcement Learning Index Advisor for a Cluster Database. Proceedings of the 24th International Database Engineering & Applications Symposium (IDEAS), Seoul, Republic of Korea, ACM, 2020, article 11, pp. 1–8.

17. Sharma V., Dyreson C., Flann N. MANTIS: Multiple Type and Attribute Index Selection using Deep Reinforcement Learning. Proceedings of the 25th International Database Engineering & Applications Symposium (IDEAS), New York, USA, 2021, pp. 56–64.

18. Yan Y., Yao S., Wang H., Gao M. Index Selection for NoSQL Database with Deep Reinforcement Learning. *Information Sciences*, 2021, vol. 561, pp. 20–30.

19. Sharma V., Dyreson C. Indexer++: Workload-Aware Online Index Tuning with Transformers and Reinforcement Learning. Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing (SAC), New York, ACM, 2022, pp. 372–380.

20. Welborn J., Schaarschmidt M., Yoneki E. Learning Index Selection with Structured Action Spaces. *arXiv e-prints*, 2019, art. no. arXiv:1909.07440.

21. Wu Y., Zhang Y., Li N. Index Advisor via DQN with Invalid Action Mask in Tree-Structured Action Space. Advanced Data Mining and Applications. ADMA 2022. *Lecture Notes in Computer Science*, Cham, Springer, 2022, vol. 13781, pp. 434–445.

22. Zhou W., Lin C., Zhou X., Li G. Breaking It Down: An In-depth Study of Index Advisors. *Proceedings of the VLDB Endowment*, 2024, vol. 17, no. 10, pp. 2405–2418. DOI:10.14778/3675034.3675035.

Viktor V. Kolomiets
Postgraduate student, Department of Information Technologies in Transport, Siberian Transport University
191, Dusi Kovalchuk St., Novosibirsk, Russia, 630049
Phone: +7-953-777-72-71
Email: viccont65@gmail.com

Received: 17.06.2026.
Accepted: 11.09.2026.

УДК 004.942
Шифр специальности ВАК: 2.3.3

М.Е. Захаров, М.Н. Чернышов, В.М. Дмитриев, К.Е. Захаров, Л.М. Ульев

Методика нелинейного предиктивного управления технологическими комплексами нефтегазоподготовки

Актуальность. Нелинейность, инерционность и перекрёстные связи процессов нефтегазоподготовки в сочетании с жёсткими технологическими ограничениями снижают эффективность классических систем автоматического регулирования и требуют перехода к предиктивным алгоритмам управления на основе цифровых двойников. ***Цель исследования.*** Разработка методики нелинейного предиктивного управления технологическими комплексами нефтегазоподготовки на базе цифрового двойника в отечественной среде имитационного моделирования. ***Методы.*** Параметрическая идентификация по ретроспективным данным SCADA-систем градиентной коррекцией; линеаризация в рабочей точке с использованием матриц Якоби и последующая дискретизация; сведение задачи оптимального управления к квадратичному программированию на скользящем горизонте; анализ устойчивости, управляемости и наблюдаемости замкнутой системы; верификация модели по паспортным данным и численным экспериментам в MatLab. ***Научная новизна.*** Впервые предложена методика интеграции цифрового двойника, построенного методом компонентных цепей, в контур предиктивного управления, учитывающая специфику идентификации по данным промышленной SCADA, явные технологические ограничения и проверку устойчивости по Ляпунову; разработан алгоритм перехода от нелинейной дифференциально-алгебраической модели к квадратичной задаче оптимизации, решаемой в реальном времени. ***Результаты.*** Подтверждена адекватность цифрового двойника теплообменного аппарата: отклонение расчётных статических характеристик от паспортных не превышает 13,5%, а переходные характеристики согласованы с верификационным расчётом в MatLab. Корректность процедур линеаризации и дискретизации обоснована проверкой критериев управляемости и наблюдаемости Калмана и решением уравнения Риккати. Сформулирована постановка задачи предиктивного управления с ограничениями на скользящем горизонте. ***Практическая значимость.*** Методика предназначена для внедрения в АСУ ТП промысловых объектов нефтегазоподготовки: её применение повысит энергоэффективность и надёжность регулирования, снизит колебательность технологических параметров и может быть использована при проектировании усовершенствованных систем управления теплообменного оборудования.

Ключевые слова: предиктивное управление, имитационное моделирование, цифровой двойник, среда моделирования MAPC, метод компонентных цепей, нефтегазоподготовка, нелинейные системы, оптимизация, теплообменные аппараты, АСУ ТП.
DOI: 10.21293/1818-0442-2026-29-2-129-134

Введение

Современные технологические объекты нефтегазоподготовки представляют собой сложные многосвязные динамические системы, функционирующие в условиях неопределённости входных воздействий, изменчивости состава сырья и наличия жёстких технологических ограничений. Процессы тепло- и массообмена характеризуются значительной инерционностью, нелинейностью и наличием перекрёстных связей между переменными.

Классические алгоритмы автоматического регулирования (ПИД-регуляторы), характерные для классической АСУ ТП, ориентированы на компенсацию текущей ошибки и не учитывают будущую динамику системы. В условиях многофакторного влияния и запаздывания процессов это приводит к возникновению колебательных режимов и неэффективному расходу энергоресурсов. Колебательные процессы при регулировании на объектах подготовки газа могут вызывать значительные экономические потери из-за повышенного потребления электрической энергии на аппаратах воздушного охлаждения.

Решение данной проблемы возможно посредством внедрения предиктивного управления (англ. Model Predictive Control, MPC [1–3]), основанного на использовании динамической математической модели объекта – цифрового двойника [4, 5], что позволяет уменьшить амплитуду колебательных процессов при регулировании.

Целью настоящей статьи является разработка методики синтеза нелинейного предиктивного управления для объектов нефтегазоподготовки, реализуемой на базе среды MAPC. В основе методики лежат принципы компонентного моделирования, открытой архитектуры для параметрической идентификации по данным SCADA-систем и динамического расчета переходных процессов.

В настоящей работе впервые предложена методика интеграции цифрового двойника, построенного на базе метода компонентных цепей (МКЦ) [6, 7], в контур предиктивного управления с учётом:

- специфики параметрической идентификации по данным промышленной SCADA;
- явного учёта технологических ограничений нефтегазоподготовки;
- процедуры линеаризации [8, 9] нелинейной модели в рабочей точке с проверкой устойчивости по Ляпунову [10, 11].

Математическая модель и методика

В основу математического обеспечения среды моделирования MAPC [12] положен МКЦ, позволяющий автоматизировать формирование системы диф-

ференциально-алгебраических уравнений (ДАУ) на основе графической топологии технологической схемы.

В результате формируется система ДАУ общего вида

$$F\left(x,\dot{x},z,u,p,t\right)=0, \quad (1)$$

где x – дифференциальные переменные состояния; z – алгебраические переменные; u – управляющие воздействия; p – параметры модели.

При условии невырожденности Якобиана

$$\det\left(\frac{\partial F}{\partial \dot{x}}\right)\neq 0 \quad (2)$$

система приводится к нормальной форме Коши. Приведение выполнено численно-символическим методом, реализованным в моделируемой среде: используется метод обратных разностей для жестких систем, алгебраические переменные $z(t)$ исключаются методом подстановки на каждом шаге интегрирования. В результате получаем

$$\dot{x}(t)=f\left(x(t),u(t),p\right). \quad (3)$$

Таким образом, цифровой двойник в МАРС представляет собой нелинейную систему первого порядка. Нелинейность в правой части обусловлена теплофизическими зависимостями, однако в рабочей точке выполняется линеаризация, поэтому данная нелинейность является несущественной с точки зрения устойчивости по Ляпунову.

Для полного описания системы в пространстве состояний также задается уравнение наблюдения

$$y(t)=\mathbf{C}_x(t)+\mathbf{D}_u(t), \quad (4)$$

где $y(t)$ – измеряемые выходы (например, температура на выходе теплообменника); $\mathbf{C}$ – матрица наблюдения (в простейшем случае единичная, если все состояния измеряемы); $\mathbf{D}$ – матрица прямого прохода (при отсутствии мгновенной связи $\mathbf{D}=0$).

В данной работе $\mathbf{C}$ и $\mathbf{D}$ задаются исходя из доступных датчиков по данным SCADA.

Исходя из описанного выше, алгоритм работы предиктивного управления в среде МАРС организовывается следующим образом:

1. Определение вектора состояний – определение динамических переменных в технологическом процессе.

2. Определение управляющих воздействий

$$u(t)=\left[Q_{\text{heat}}F_{\text{feed1}}F_{\text{feed2}}\ldots\right]^T. \quad (5)$$

3. Определение технологических ограничений (диапазон работы оборудования, заданные нормы ведения технологического режима согласно регламенту)

$$x_{\min}\le x(t)\le x_{\max},$$
$$u_{\min}\le u(t)\le u_{\max}. \quad (6)$$

4. Построение динамической модели в МАРС.

5. Параметрическая идентификация

$$J_{\text{id}}(p)=\sum_{k=1}^{N}\left\|y_k-y_k(p)\right\|^2, \quad (7)$$

где y_k – реальные измерения; $y_k(p)$ – значения, полученные в модели.

Выполнение градиентной коррекции:

$$p_{i+1}=p_i-\alpha\nabla J_{\text{id}}(p), \quad (8)$$

где критерий сходимости

$$\nabla J_{\text{id}}(p)<\varepsilon. \quad (9)$$

После идентификации цифровой двойник становится репрезентативным.

Этап линеаризации. Линеаризация модели выполняется в рабочей точке с использованием матриц Якоби

$$\mathbf{A}=\frac{\partial f}{\partial x},\ \mathbf{B}=\frac{\partial f}{\partial u},\ \Delta\dot{x}=\mathbf{A}\Delta x+\mathbf{B}\Delta u. \quad (10)$$

Этап дискретизации с шагом Δt дает

$$x_{k+1}=\Phi_{x_k}+\Gamma_{u_k}, \quad (11)$$

где

$$\Phi=e^{\mathbf{A}\Delta t},\quad \Gamma=\int_0^{\Delta t}e^{\mathbf{A}\Delta t}\mathbf{B}d\tau.$$

В табл. 1 приведены числовые значения коэффициентов полученных уравнений для теплообменного аппарата в рабочей точке (расход греющего теплоносителя 12,5 м^3/ч, температура входа 85 °С) при порядке модели $n=2$ и шаге дискретизации $\Delta t=10$ с.

Для теплообменного аппарата с характерным временем переходного процесса $\tau\approx 120$–180 с и шагом дискретизации $\Delta t=10$ с приняты следующие параметры горизонта:

- горизонт прогноза $N_p=18$ тактов (180 с);
- горизонт управления $N_u=6$ тактов (60 с).

Задача предиктивного управления формулируется как поиск вектора управления u, минимизирующего квадратичный функционал J:

$$J=\sum_{i=1}^{N}\left\|y_{k+i}+r_{k+i}\right\|_Q^2+\sum_{i=0}^{M-1}\left\|u_{k+i}\right\|_R^2, \quad (12)$$

где $y_{k+i}-r_{k+i}$ – ошибка отслеживания задания (отклонения); Δu – изменение управления; Q,R – весовые коэффициенты матрицы.

При проведении оптимизации в границах

$$x_{\min}\le x(t)\le x_{\max},$$
$$u_{\min}\le u(t)\le u_{\max}.$$

Таблица 1

Пример значений коэффициентов уравнения состояния линеаризованной дискретной модели теплообменного аппарата в рабочей точке

Матрица	Значение (пример)
A	$\begin{bmatrix}-0{,}0412 & 0{,}0087\\ 0{,}0031 & -0{,}0295\end{bmatrix}$
B	$\begin{bmatrix}0{,}1520\\ 0{,}0043\end{bmatrix}$
C	$\begin{bmatrix}1 & 0\end{bmatrix}$
D	0

Далее формируется стандартная квадратичная задача

$$\min_{\mathbf{U}} \frac{1}{2}\mathbf{U}^T\mathbf{H}\mathbf{U} + \mathbf{f}^T\mathbf{U}, \qquad (13)$$

где **U** – вектор искомых переменных; **H** – матрица Гессиана (положительно определенная), описывающая кривизну целевой функции; **f** – линейный вектор, учитывающий текущую ошибку и состояние системы.

Устойчивость замкнутой системы проверяется решением дискретного уравнения Риккати

$$P = \mathbf{A}^T P\mathbf{A} - \mathbf{A}^T P\mathbf{B}(R + \mathbf{B}^T P\mathbf{B})^{-1}\mathbf{B}^T P\mathbf{A} + Q, \qquad (14)$$

если $P \geq 0$, система устойчива.

Критерий управляемости Калмана применим к линейной стационарной системе, полученной на этапе линеаризации нелинейной модели в фиксированной рабочей точке и последующей дискретизации. Матрицы **A** и **B** в данном случае постоянны (не зависят от времени), что обеспечивает корректность применения рангового критерия:

$$\text{rank}[\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \ldots, \mathbf{A}^{n-1}\mathbf{B}] = n, \qquad (15)$$

где n – размерность вектора состояния системы, которая определяется количеством независимых переменных, необходимых для полного описания её динамики.

В случае управления двумя параметрами порядок системы принимается равным $n = 2$. Наблюдаемость проверяется по аналогичной матрице, составленной из **C** и **A**.

Алгоритм синтеза предиктивного управления:

1. Определение вектора состояний и управляющих воздействий.
2. Задание технологических ограничений.
3. Построение динамической модели в МАРС и параметрическая идентификация по данным SCADA (минимизация функции ошибки градиентным методом).
4. Линеаризация в рабочей точке с использованием матриц Якоби и дискретизация с шагом Δt.
5. Формулировка задачи оптимизации: поиск вектора управления u, минимизирующего квадратичный функционал J с весовыми коэффициентами и ограничениями.
6. Проверка устойчивости (уравнение Риккати), управляемости и наблюдаемости (критерий Калмана).

Последовательность операций на каждом такте управления наглядно иллюстрирует обобщённая структурная схема контура краткосрочного прогнозирования и предиктивного управления теплообменным аппаратом (рис. 1).

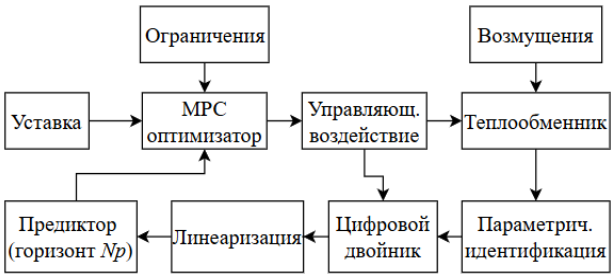


Рис. 1. Обобщённая структурная схема контура краткосрочного прогнозирования и предиктивного управления теплообменным аппаратом

Результаты и обсуждение

Для разработанного цифрового двойника теплообменного аппарата проведено сравнение расчётных статических характеристик с паспортными данными (табл. 2). Максимальная относительная погрешность составила 13,5%, что позволяет считать модель адекватной для синтеза алгоритмов управления.

Для подтверждения расчётных данных в динамике выполнен численный эксперимент в среде MatLab: получены переходные характеристики температур горячего и холодного потоков на выходе теплообменного аппарата при ступенчатом возмущении по расходу теплоносителя (рис. 2).

Разработанный цифровой двойник обладает потенциалом для реализации предиктивного управления. Как показано в работе [15], применение МРС для нелинейных объектов со значительным запаздыванием позволяет повысить качество регулирования более чем на 20% по сравнению с ПИД-регулятором.

Однако в работе [15] рассматривался общий подход к синтезу регулятора, тогда как в настоящей статье предложена полная методика – от построения цифрового двойника в отечественной среде имитационного моделирования МАРС до формирования квадратичной задачи оптимизации с учётом специфики нефтегазоподготовки (изменчивость сырья, жёсткие ограничения по давлению/температуре).

Таблица 2

Сравнение расчётных статических характеристик с паспортными данными

Параметр	Паспортные данные	Расчетные данные	Отклонение, %
Тепловая нагрузка	1 250	1 315	5,2
Коэффициент теплопередачи	480	545	13,5
Сопротивление в трубках	45	39	13,3
Сопротивление в межтрубном пространстве	60	52	6,6

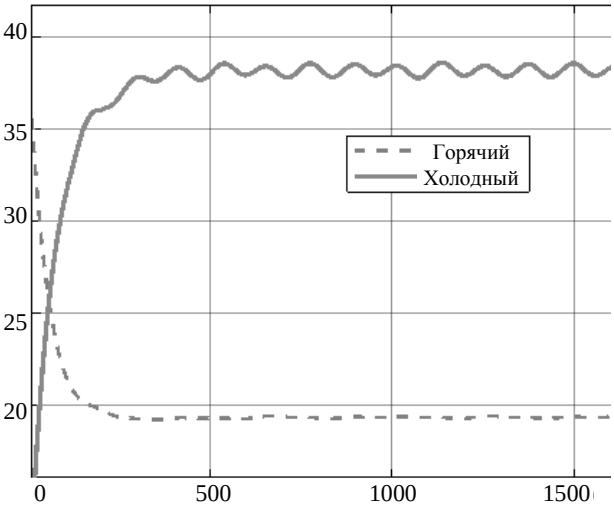


Рис. 2. Зависимость температур горячего и холодного потоков на выходе теплообменного аппарата при ступенчатом возмущении по расходу теплоносителя

Заключение

Показано, что метод компонентных цепей, реализованный в среде MAPC, обеспечивает формирование адекватных динамических моделей (цифровых двойников) технологических аппаратов с учётом нелинейности, инерционности и перекрёстных связей.

Предложена методика параметрической идентификации моделей на основе ретроспективных данных SCADA, позволяющая адаптировать цифровой двойник к реальным условиям эксплуатации.

Описана процедура перехода от нелинейной модели ДАУ к линейному дискретному представлению в пространстве состояний, пригодному для решения задачи оптимального управления.

Сформулирована постановка задачи предиктивного управления на скользящем горизонте с квадратичным критерием качества и учётом технологических ограничений.

Перспективы дальнейших исследований:

• отказ от линеаризации в пользу нелинейного решателя NMPC [16–18] в реальном времени для расширения рабочей области;

• интеграция разработанного алгоритма в промышленный контур АСУ ТП с тестированием на реальном объекте нефтегазоподготовки;

• разработка адаптивного механизма коррекции весовых коэффициентов функционала J в зависимости от режима эксплуатации.

Практическая реализация предложенного подхода открывает возможности для повышения энергоэффективности и надёжности функционирования промысловых объектов за счёт перехода от классического регулирования к усовершенствованному управлению на основе прогнозирующих моделей.

Литература

1. Model predictive control for gas turbine shaft speed based on model compensation using extended state observer / F. Zhang, Ya. Xue, B. Wei, B. Zhang // Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering. – 2023. – Vol. 237, No. 3. – P. 513–523. DOI: 10.1177/09596518221127511.

2. Fan Sh. Model predictive current control method for PMSM drives based on an improved prediction model / Sh. Fan, Ch. Tong // Journal of Power Electronics. – 2020. – Vol. 20. – P. 1456–1466. DOI: 10.1007/s43236-020-00125-4.

3. Optimization of an Integrated Production Model Using Long Short-Term Memory and Model Predictive Control under Constraints / M. Fadaei, M. Ja. Ameri, Y. Rafiei et al. // Digital Technologies Research and Applications. – 2025. – Vol. 4, No. 3. – P. 97–112. DOI: 10.54963/dtra.v4i3.1604.

4. The connotation of digital twin, and the construction and application method of shop-floor digital twin / C. Zhuang, T. Miao, J. Liu, H. Xiong // Robotics and Computer-Integrated Manufacturing. – 2021. – Vol. 68. – P. 102075. DOI: 10.1016/j.rcim.2020.102075.

5. Hypermedia-based digital twin ecosystems: A proposal for digital twin interoperability / A. Giulianelli, S. Burattini, A. Ciortea, A. Ricci // Future Generation Computer Systems. – 2026. – Vol. 177. – P. 108219. DOI: 10.1016/j.future.2025.108219.

6. Дмитриев В.М. Методика построения многоуровневых компонентных цепей для моделирования химико-технологических систем / В.М. Дмитриев, Т.В. Ганджа // Доклады ТУСУР. – 2017. – Т. 20, № 3. – С. 98–102. DOI: 10.21293/1818-0442-2017-20-3-82-86.

7. Дмитриев В.М. Среда многоуровневого компьютерного моделирования химико-технологических систем / В.М. Дмитриев, Т.В. Ганджа. – Томск: Изд-во ТГУ, 2017. – 332 с. – ISBN 978-5-7511-2514-1.

8. Мануйлов Ю.С. Линеаризация модели движения многосвязного динамического объекта методом псевдокоординатной линеаризации / Ю.С. Мануйлов, А.В. Толоконников // Интеллектуальные технологии на транспорте. – 2025. – № 3 (43). – С. 58–67. DOI: 10.20295/2413-2527-2025-343-58-67.

9. Гаюбов Т.Н. Некоторые проблемы теории линеаризации приемной информации микропроцессорного приемника рельцовых цепей кодовой автоблокировки / Т.Н. Гаюбов, А.Н. Садиков, С.К. Убайдуллаев // Universum: технические науки. – 2022. – № 4-12 (97). – С. 59–62.

10. Курбаншоев С.З. Решение систем разностных уравненнй при помощи функций Ляпунова / С.З. Курбаншоев, Н.С. Якубов // Вестник Бохтарского гос. ун-та им. Носира Хусрава. – Сер. естественных наук. – 2026. – № 2-1 (144). – С. 5–7.

11. Aloev R.D. Numerical study of Lyapunov stability of an upwind difference scheme for a quasilinear hyperbolic system / R.D. Aloev, A.S. Berdishev, D.E. Nematova // Problems of Computational and Applied Mathematics. – 2025. – No. 3(67). – P. 83–96. DOI: 10.71310/pcam.3_67.2025.07.

12. Кочергин М.И. Методика и алгоритмы визуального моделирования непрерывных и дискретно-непрерывных физико-технических задач методом компонентных цепей: специальность 05.13.18: дис. … канд. техн. наук. – Томск: ТУСУР, 2019. – 242 с.

13. Захаров М.Е. Применение компьютерного моделирования для усовершенствованного управления перекачивания углеводородного сырья с целью исключения гидроударов в трубопроводной системе / М.Е. Захаров, Т.В. Ганджа // Электронные средства и системы управления: матер. докладов XXI Междунар. науч.-практ. конф.: в 3 ч. – Томск: ТУСУР; В-Спектр (ИП В.М. Бочкарева), 2025. – Ч. 2. – С. 171–173.

14. Захаров М.Е. Применение метода прогнозирующего управления модели для оптимизации автоматизированных систем управления технологическими процессами нефтегазовых объектов // Технологии обустройства нефтяных, газовых и газоконденсатных месторождений: сб. тезисов VIII науч.-техн. конф., Томск, 24–25 сентября 2025 г. – Томск: ООО «РН-Проектирование Добыча», 2025. – С. 233–236.

15. Колодин А.А. Разработка и исследование регулятора на основе прогнозирущей модели / А.А. Колодин, В.В. Елшин // Вестник Самарского гос. техн. ун-та. – Сер.: Технические науки. – 2021. – Т. 29, № 1(69). – С. 36–45. DOI: 10.14498//tech.2021.1.3.

16. Hierarchical NMPC for energy-efficient torque vectoring in four in-wheel motor electric vehicles / S. Park, Ju. Kim, D.G. Nguyen et al. // Control Engineering Practice. – 2025. – Vol. 164. – P. 106477. DOI: 10.1016/j.conengprac.2025.106477.

17. NMPC design for a self-aligning compliant gait rehabilitation robot / Y. Jin, T. Goyal, P. K. Jamwal et al. // Robotics and Autonomous Systems. – 2025. – Vol. 194. – P. 105128. DOI: 10.1016/j.robot.2025.105128.

18. Xu J. N-GKLS: An NMPC problem generator and a test platform for NMPC solvers / J. Xu, L. Xu // Expert Systems with Applications. – 2023. – Vol. 214. – P. 119029. DOI: 10.1016/j.eswa.2022.119029.

Захаров Михаил Евгеньевич
Аспирант каф. компьютерных систем в управлении и проектировании (КСУП)
Томского государственного университета систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050;
инж. проектирования АСУ 1-го уровня отд. АСУ ТП
ООО «РН-Проектирование Добыча»
Говорова ул., 19б, г. Томск, Россия, 634057
Тел.: +7-913-869-99-40
Эл. почта: epsn97@mail.ru

Чернышов Михаил Николаевич
Вед. инж. отд. концептуального проектирования
ООО «РН-Проектирование Добыча»
Мира пр-т, 70А, г. Томск, Россия, 634027;
аспирант Национального исследовательского Томского политехнического университета (НИ ТПУ)
Ленина пр-т, 30, г. Томск, Россия, 634050
Тел.: +7 (382-2) 70-17-77
Эл. почта: mixailcher35@gmail.com

Дмитриев Вячеслав Михайлович
Д-р техн. наук, проф. КСУП ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0002-1089-6993
Тел.: +7-913-867-01-56
Эл. почта: dmitriewvm@gmail.com

Захаров Кирилл Евгеньевич
Специалист инженерной школы ядерных технологий (ИЯТШ) НИ ТПУ
Ленина пр-т, 30, г. Томск, Россия, 634050
Тел.: +7 (382-2) 70-17-77
Эл. почта: kzaharov18@gmail.com

Ульев Леонид Михайлович
Д-р техн. наук, проф. НИ ТПУ
Ленина пр-т, 30, г. Томск, Россия, 634050
ORCID: 0000-0002-2732-6444
Тел.: +7 (382-2) 70-17-77
Эл. почта: ulyevlm@mail.ru

Поступила в редакцию: 03.03.2026.
Принята к публикации: 19.08.2026.

Zakharov M.E., Chernyshov M.N., Dmitriev V.M., Zakharov K.E., Uliev L.M.

Methodology for nonlinear predictive control of oil and gas treatment complexes

Relevance. The nonlinearity, inertia, and cross-couplings inherent in oil and gas treatment processes, combined with strict technological constraints, reduce the efficiency of conventional automatic control systems and necessitate the transition to predictive control algorithms based on digital twins. ***Purpose of the study.*** Development of a methodology for nonlinear predictive control of oil and gas treatment complexes based on a digital twin implemented in the domestic MARS simulation environment. ***Methods.*** Parametric identification is performed using retrospective SCADA data via gradient correction; linearization at the operating point is carried out using Jacobian matrices followed by discretization. The optimal control problem is reduced to quadratic programming on a receding horizon. Stability, controllability, and observability of the closed-loop system are analyzed. Model verification is conducted against nameplate data and through numerical experiments in MATLAB. ***Scientific novelty.*** For the first time, a methodology is proposed for integrating a digital twin built by the component chain method into the predictive control loop, which accounts for the specifics of identification based on industrial SCADA data, explicit technological constraints, and Lyapunov stability verification. An algorithm is developed for converting a nonlinear differential-algebraic model into a quadratic optimization problem solvable in real time. ***Results.*** The adequacy of the heat exchanger digital twin has been confirmed: the deviation of the calculated steady-state characteristics from the nameplate values does not exceed 13.5%, and the transient responses agree with the verification computation performed in MATLAB. The correctness of the linearization and discretization procedures is substantiated by verifying the Kalman controllability and observability criteria and by solving the Riccati equation. The problem statement for constrained receding-horizon predictive control has been formulated. ***Practical significance.*** The methodology is intended for implementation in the industrial control systems (ICS) of oil and gas field treatment facilities. Its application is expected to enhance energy efficiency and control reliability, reduce the oscillations of process parameters, and can be used in the design of advanced control systems for heat exchange equipment.

Keywords: predictive control, simulation modeling, digital twin, MARS simulation environment, component chain method, oil and gas treatment, nonlinear systems, optimization, heat exchangers, industrial control systems (ICS).
DOI: 10.21293/1818-0442-2026-29-2-129-134

References

1. Zhang F., Xue Y., Wei B., Zhang B. Model predictive control for gas turbine shaft speed based on model compensation using extended state observer. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering,* 2023, vol. 237, no. 3, pp. 513–523. DOI: 10.1177/09596518221127511.

2. Fan S., Tong C. Model predictive current control method for PMSM drives based on an improved prediction model. *Journal of Power Electronics*, 2020, vol. 20, pp. 1456–1466. DOI: 10.1007/s43236-020-00125-4.

3. Fadaei M., Ameri M.J., Rafiei Y., Al-Alimi A., Hatami Golmakani H., Vafadar Eidgahi M.M. Optimization of an Integrated Production Model Using Long Short□Term Memory and Model Predictive Control under Constraints. *Digital Technologies Research and Applications*, 2025, vol. 4, no. 3, pp. 97–112. DOI: 10.54963/dtra.v4i3.1604.

4. Zhuang C., Miao T., Liu J., Xiong H. The connotation of digital twin, and the construction and application method of shop-floor digital twin. *Robotics and Computer-Integrated Manufacturing*, 2021, vol. 68, Art. 102075. DOI: 10.1016/j.rcim.2020.102075.

5. Giulianelli A., Burattini S., Ciortea A., Ricci A. Hypermedia-based digital twin ecosystems: A proposal for digital twin interoperability. *Future Generation Computer Systems*, 2026, vol. 177, art. 108219. DOI: 10.1016/j.future.2025.108219.

6. Dmitriev V.M., Gandzha T.V. [Method to construct multilevel component chains for modeling chemical-technological systems]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2017, vol. 20, no. 3, pp. 98–102. DOI: 10.21293/1818-0442-2017-20-3-82-86 (in Russ.).

7. Dmitriev V.M., Gandzha T.V. *Sreda mnogourovnevogo komp'yuternogo modelirovaniya khimiko-tekhnologicheskikh system* [Multilevel computer modeling environment for chemical-technological systems]. Tomsk: Tomsk University Publishing House, 332 p. (in Russ.).

8. Manuylov Yu.S., Tolokonnikov A.V. [Linearization of the motion model of a multi-connected dynamic object by pseudo-coordinate linearization]. *Intellectual Technologies on Transport*, 2025, no. 3(43), pp. 58–67. DOI: 10.20295/2413-2527-2025-343-58-67 (in Russ.).

9. Gayubov T., Sadikov A., Ubaydullayev S. Some problems of the theory of linearization of receiving information of a microprocessor receiver of code auto-locking rail circuits. *Universum: Technical Sciences*, 2022, no. 4-12 (97), pp. 59–62.

10. Kurbanshoev S.Z., Yakubov N.S. [Solving systems of difference equations using Lyapunov functions]. *Bulletin of Bokhtar State University named after Nosiri Khusrav. Series of natural sciences*, 2026, no. 2-1 (144), pp. 5–7 (in Russ.).

11. Aloev R.D., Berdishev A.S., Nematova D.E. Numerical study of Lyapunov stability of an upwind difference scheme for a quasilinear hyperbolic system. Problems of Computational and Applied Mathematics, 2025, no. 3(67), pp. 83–96. DOI: 10.71310/pcam.3_67.2025.07.

12. Kochergin M.I. *Metodika i algoritmy vizual'nogo modelirovaniya nepreryvnykh i diskretno-nepreryvnykh fiziko-tekhnicheskikh zadach metodom komponentnykh tsepey* [Methodology and algorithms for visual modeling of continuous and discrete-continuous physical and technical problems by the method of component circuits. Cand. Diss. Abstract]. Tomsk, 2019, 242 p (in Russ.).

13. Zakharov M.E., Gandzha T.V. *Primenenie komp'yuternogo modelirovaniya dlya usovershenstvovannogo upravleniya perekachivaniya uglevodorodnogo syr'ya s tsel'yu isklyucheniya gidroudarov v truboprovodnoy sisteme* [Application of computer modeling for improved control of hydrocarbon raw materials pumping to eliminate water hammers in the pipeline system]. Elektronnye sredstva i sistemy upravleniya. [*Electronic control means and systems*], Proc. of the XXI International Scientific and Practical Conference, Tomsk, TUSUR, V-Spectr (Individual Entrepreneur V.M. Bochkareva), 2025, 2nd part, pp. 171–173 (in Russ.).

14. Zakharov M.E. *Primenenie metoda prognoziruyushchego upravleniya modeli dlya optimizatsii avtomatizirovannykh sistem upravleniya tekhnologicheskimi protsessami neftegazovykh ob'ektov* [Application of the model predictive control method for optimizing automated control systems for technological processes of oil and gas facilities]. Tekhnologii obustroystva neftyanykh, gazovykh i gazokondensatnykh mestorozhdeniy [*Technologies for the development of oil, gas and gas condensate fields*] *Proc. of the VIII scientific and technical* conf., Tomsk, LLC «RN-Upstream Design», September 24–25, 2025, pp. 235–236 (in Russ.).

15. Kolodin A.A., Elshin V.V. [Research and development of the controller based on the model predictive control*]. Vestnik of Samara State Technical University (Technical Sciences Series)*, 2021, vol. 29, no. 1 (69), pp. 36–45. DOI: 10.14498/tech.2021.1.3 (in Russ.).

16. Park S., Kim Ju., Nguyen D.G. et al. Hierarchical NMPC for energy-efficient torque vectoring in four in-wheel motor electric vehicles. *Control Engineering Practice*, 2025, vol. 164, pp. 106477. DOI: 10.1016/j.conengprac.2025.106477.

17. Jin Y., Goyal T., Jamwal P. K. et al. NMPC design for a self-aligning compliant gait rehabilitation robot. *Robotics and Autonomous Systems*, 2025, vol. 194, pp. 105128. DOI: 10.1016/j.robot.2025.105128.

18. Xu J., Xu L. N-GKLS: An NMPC problem generator and a test platform for NMPC solvers. *Expert Systems with Applications*, 2023, vol. 214, pp. 119029. DOI: 10.1016/j.eswa.2022.119029.

Mikhail E. Zakharov
Postgraduate Student, Department of Computer Control and Design Systems (CCDS), Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050;
Control System Design Engineer (Level 1) Industrial Automation Department (I&C), LLC «RN-Upstream Design»
19b, Govorova St., Tomsk, Russia, 634057
Phone: +7-913-869-99-40
Email: epsn97@mail.ru

Mikhail N. Chernyshov
Lead Engineer, Conceptual Design Department,
LLC «RN-Upstream Design»
70A, Mira Ave., Tomsk, Russia, 634027;
Postgraduate Student,
National Research Tomsk Polytechnic University (TPU)
30, Lenin Ave., Tomsk, Russia, 634050
Phone: +7 (382-2) 70-17-77
Email: mixailcher35@gmail.com

Vyacheslav M. Dmitriev
Doctor of Science in Engineering, Professor,
Department of CCDS, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0002-1089-6993
Phone: +7-913-867-01-56
Email: dmitriewvm@gmail.com

Kirill E. Zakharov
Specialist, School of Nuclear Science & Engineering, TPU
30, Lenin Ave., Tomsk, 634050, Russia
Phone: +7 (382-2) 70-17-77
Email: kzaharov18@gmail.com

Leonid M. Uliev
Doctor of Science in Engineering, Professor, TPU
30, Lenin Ave., Tomsk, 634050, Russia.
ORCID: 0000-0002-2732-6444
Phone: +7 (382-2) 70-17-77
Email: ulyevlm@mail.ru

Received: 03.03.2026.
Accepted: 19.08.2026.

УДК 691.5
Шифр специальности ВАК: 2.3.3

М.М. Свинцов, Л.П. Мышляев, Г.В. Макаров, В.В. Грачев

Математические модели объектов управления с координатно-параметрическими связями

Актуальность. Учет координатно-параметрических связей (КПС) в объектах управления является важным условием для повышения эффективности автоматизированных технологических комплексов горнодобывающей, металлургической и химической промышленности. ***Цель исследования.*** Разработка новых структур моделей объектов с учетом взаимовлияния координатных и параметрических связей, формирование классификации новых структур математических моделей объектов управления с КПС. ***Методы.*** Использованы методы системного анализа, теории автоматического управления, математического моделирования динамических систем с запаздыванием и параметрическими зависимостями. ***Научная новизна.*** Создание математических моделей и схем объектов управления с КПС, учитывающих взаимное влияние координатных и параметрических переменных (в том числе для объектов с рециклом), формирование классификации объектов управления с координатно-параметрическими связями. ***Результаты.*** Предложены структуры математических моделей для четырех классов объектов с КПС, а также приведены примеры таких объектов в технологическом комплексе углеобогащения. ***Практическая значимость*** заключается в расширении инструментария анализа и синтеза эффективных автоматизированных технологических комплексов, содержащих объекты управления с КПС, за счет оперативного совместного определения координатных воздействий и параметров объекта.

Ключевые слова: система управления, автоматизированный технологический комплекс, координатно-параметрические связи (КПС), рецикл.
DOI: 10.21293/1818-0442-2026-29-2-135-139

Введение

Первым из условий системных исследований является необходимость наличия в системе двух или более системных связей [1]. Определение понятия «связь» в контексте систем управления постоянно развивается, вместе с чем растёт и его практическая значимость [2]. Например, конкретизация принципа однонаправленности систем при формировании полной модели объекта изучения основывается на том, что полный объект исследования не должен охватываться прямыми и обратными связями, которые должны быть заключены внутри этого объекта [3].

Теоретическим и практическим развитием системных связей послужили принципы и конкретные положения теории новых типов обратных связей [4, 5]. Координатные и операторные (параметрические и структурные) типы связей в [1] предложено реализовать в управляющей части системы управления.

Полезно развить такой подход непосредственно и на объект управления, учитывая наличия разных типов связей между элементами и воздействиями внутри самого объекта управления. Это диктуется в том числе и более «тонким» математическим описанием многих объектов управления технических и социальных комплексов, что необходимо для повышения эффективности их управления [6].

При этом новые модели объектов управления требуют введения в алгоритмы управления функций совместного оперативного определения параметров объектов и координатных воздействий.

Новые математические модели объектов и их классификация

Для многих технологических объектов горнодобывающей, металлургической, химической и других отраслей промышленности имеется ряд объектов, которые можно представить моделями с координатными и параметрическими связями (КПС), когда координатные воздействия влияют на параметры объектов, что влияет через выработанное управление и на координатные воздействия [7]. Сформирована следующая классификация таких объектов и представлены их математические модели.

1. Объекты с влиянием выходных воздействий на его параметры.

В общем виде эти объекты можно представить следующей математической моделью:

$$y(t)=\varphi_y(t)\times u(t-\tau_u)+\varphi_w(t)\times w(t-\tau_w), \quad (1)$$

$$\varphi_y(t)=\psi_y\times y(t)\ ;\varphi_w(t)=\psi_w\times y(t), \quad (2)$$

где $y(t),u(t),w(t)$ – выходное, управляющее и внешнее воздействия; $\varphi_y(t),\varphi_w(t)$ – операторы преобразования $u(t)$ и $w(t)$ в параметры объекта; ψ_y,ψ_w – операторы преобразования $y(t)$ в $\varphi_y(t)$ и $\varphi_w(t)$; τ_u,τ_w – запаздывание в канале преобразования управляющих и внешних воздействий; t – время.

Новизна предложенной модели заключается в введении дополнительной зависимости (2).

Простым примером рассматриваемого объекта может служить демпфирующая емкость (бункер) для сыпучих или жидких материалов (рис. 1).

Математическая модель этого агрегата имеет вид

$$y(t)=\int_{t-\Theta}^{t}k(t)\left[u(t-\Theta)+w(t-\Theta)\right]dt, \quad (3)$$

$$k(t)=\begin{cases}k_0 \text{ при } y(t)>y^*,\\ k_0+a\left[y^*-y(t)\right] \text{ при } y(t)\le y^*,\end{cases} \quad (4)$$

где y – уровень материала в ёмкости; y^* – величина уровня, при котором цилиндрическая часть ёмкости переходит в конусообразную; k_0, a – коэффициенты, зависящие от геометрии емкости.

При более сложной геометрии емкости вместо величины y можно брать площадь сечения бункера. Тогда коэффициент будет иметь следующий вид:

$$k(t)=\beta\times S_{y(t)}, \quad (5)$$

где $S_{y(t)}$ – площадь поперечного сечения емкости при $y(t)$; β – коэффициент пересчета $S_{y(t)}$ в $k(t)$.

2. Объекты с влиянием управляющих воздействий на его параметры

$$y(t)=y_\tau(t-\tau(t)), \quad (6)$$

$$\tau(t)=C\times u(t), \quad (7)$$

где τ – время запаздывания; C – пересчётный коэффициент.

Примером такого объекта может служить конвейер, когда количество материала на выходе с конвейера зависит от его скорости (рис. 2).

Математическая модель конвейера имеет вид

$$Q_2(t)=Q_1(t-\tau(t)), \quad (8)$$

$$\tau(t)=C\times u(t), \quad (9)$$

где $Q_1(t)$ – расход на выходе питателя (производительность питателя); $Q_2(t)$ – расход на выходе конвейера (производительность конвейера); τ – время нахождения материала на конвейере.

Новизна данной модели заключается в введении зависимости (9).

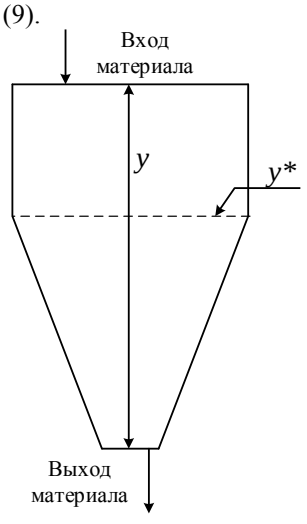


Рис. 1. Схема демпфирующей ёмкости

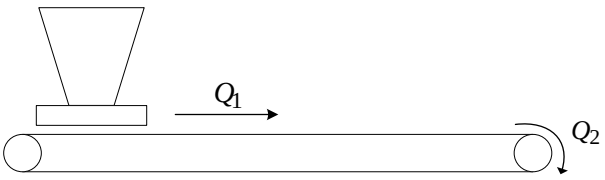


Рис. 2. Схема конвейера

Более сложный вариант объекта – проточные нагревательные емкости и технологические агрегаты с конечным временем обработки сырья, например, установка металлизации окисленных окатышей [8, 9], представленная на рис. 3.

В зависимости от скорости выгрузки восстановленных окатышей изменяется время нахождения окатышей в агрегате и, соответственно, коэффициент передачи, характеризующий степень восстановления окатышей, а также время отсечки. Математическая модель установки имеет вид [10]

$$y(t)=\int_{t-\theta_u}^{t} k_u(t)\times U(t-\theta_u(t))dt, \quad (10)$$

$$\theta_u(t)=C_\theta\times U(t), \quad (11)$$

$$k_u(t)=C_u\times\theta_u(t), \quad (12)$$

где $\theta_u(t)$ – время отсечки, время нахождения материала в установке; C_θ и C_u – пересчетные коэффициенты.

3. Объекты с внутренней положительной обратной связью (с рециклом), для которых рецикл влияет на параметры элементов прямой цепи [11]. Математическая модель с запаздыванием в пространстве состояний имеет вид [12]

$$\dot{\mathbf{X}}(t)=\mathbf{A}(t)\mathbf{X}(t-\tau_X)+\mathbf{B}(t)\mathbf{U}(t-\tau_U)+\mathbf{C}(t)\mathbf{W}(t-\tau_W), \quad (13)$$

$$\mathbf{Y}(t)=\mathbf{D}(t)\times\mathbf{X}(t-\tau_Y), \quad (14)$$

где $\mathbf{X, U, W, Y}$ – векторы состояний, управления, внешних и выходных воздействий; τ_X, τ_U, τ_W – соответствующее время запаздывания; $\mathbf{A}(t), \mathbf{B}(t), \mathbf{C}(t), \mathbf{D}(t)$ – матрицы соответствующих размерностей.

Примером таких объектов могут служить дробильные агрегаты с возвратом на вход крупной фракции материала, которые меняют характеристики (параметры) непосредственно механизмов дробления [13]. Структурная схема такого агрегата показана на рис. 4.

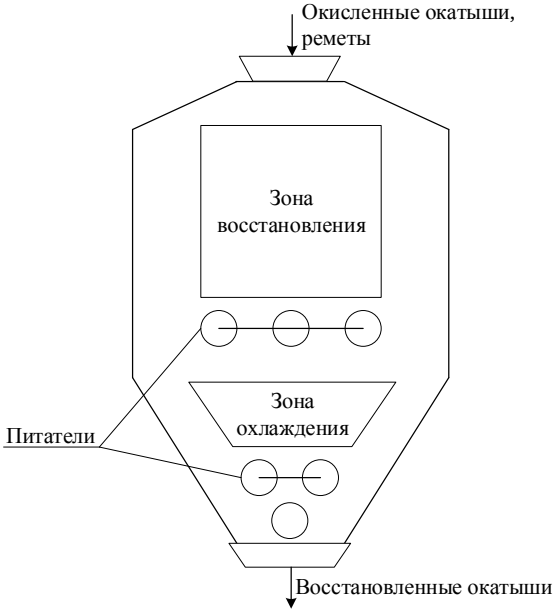


Рис. 3. Схема установки металлизации окисленных окатышей



Рис. 4. Схема дробильного агрегата

4. Комбинированные комплексы, объединяющие предыдущие три группы объектов и предназначенные для описания сложных технологических производств [14].

Реализация описанных выше моделей проводилась в рамках разработки алгоритмов управления технологическими объектами на программно-аппаратных комплексах действующего производства.

Пример технологического комплекса, включающего объекты с КПС

Описанные объекты встречаются во многих автоматизированных технологических комплексах горнодобывающей, металлургической и химической промышленности.

На рис. 5 представлен фрагмент видеокадра углеобогатительной фабрики, на котором контурами отмечены объекты с КПС. Контуром в виде сплошной линии отмечены демпфирующие ёмкости насосов 237, 252 и 253. Они обеспечивают работу тяжелосредных гридоциклонов и характеризуются влиянием выходных воздействий на параметры объекта – так, при уменьшении уровня в ёмкости будет увеличиваться влияние производительности связанного насоса на скорость изменения уровня.

Контуром в виде пунктирной линии отмечены конвейеры 50, 52, 54, 55, 201, 202, 336 и 340. В частности, конвейерами 54 и 55 подаются отдельные компоненты шихты в требуемых пропорциях, конвейер 201 подаёт готовую шихту в технологический процесс. Особенностью данного объекта является влияние управляющих воздействий на его параметры – в зависимости от скорости конвейера 201 будет уменьшаться запаздывание объекта, что необходимо учитывать при изменении соотношения компонентов шихты.

Контуром в виде двойной линии отмечена дробилка 102, являющаяся объектом с внутренней положительной обратной связью (с рециклом), где крупная фракция, полученная в результате дробления, повторно подаётся на вход агрегата.

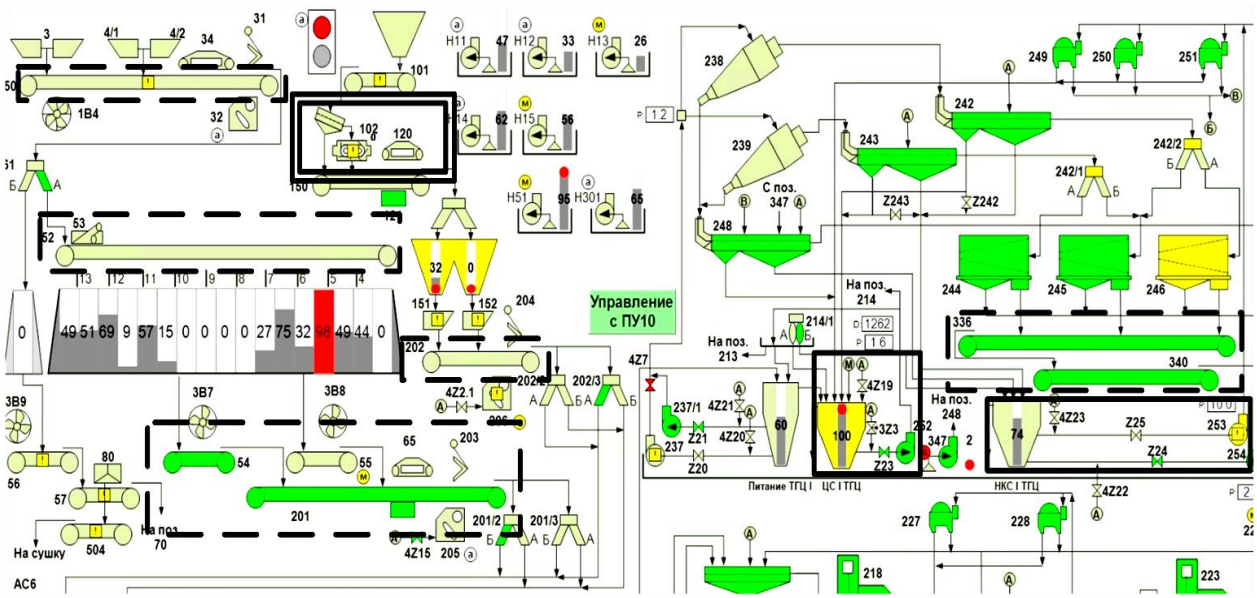


Рис. 5. Фрагмент главного видеокадра углеобогатительной фабрики

Заключение

В работе разработаны новые структуры моделей объектов управления с КПС, охватывающие основные типы взаимного влияния координатных и параметрических переменных, сформирована классификация таких объектов, представлены их математические модели. Наличие КПС существенно изменяет свойства объектов управления, что должно учитываться при синтезе систем управления путем совместного определения координатных воздействий и параметров объекта, а для объектов с рециклом еще и за счет цепей модельной компенсации рециклов [15].

Приведены примеры предложенных структур моделей объектов в технологическом комплексе углеобогатительной фабрики.

Полученные результаты расширяют возможности анализа и синтеза систем управления и могут быть использованы для дальнейших исследований в области моделирования и разработки эффективных автоматизированных технологических комплексов.

Литература

1. Блауберг И.В. Становление и сущность системного подхода / И.В. Блауберг, Э.Г. Юдин. – М.: Наука, 1973. – 270 с.

2. Optner S.L. Systems analysis for business and industrial problem solving. – N.J.: Prentice-Hall, 1965. – 116 p.

3. Laske O. Advanced Systems-Level Problem Solving, Volume 3: Manual of Dialectical Thought Forms. – Cham: Springer, 2023. – 234 p.

4. Емельянов С.В. Основные принципы системного анализа / С.В. Емельянов, Э.Л. Наппельбаум // Матер.

II Всесоюз. науч.-техн. конф.: Проблемы научной организации управления социалистической промышленностью / под ред. Д.М. Гвишиани, С.Е. Каменицера. – М.: Экономика, 1974. – С. 92–99.

5. Емельянов С.В. Новые типы обратной связи: Управление при неопределённости / С.В. Емельянов, С.К. Коровин. – М.: Наука, Изд-во «Физ.-мат. лит.», 1997. – 352 с.

6. Cohen M. Adaptive and Learning-Based Control of Safety-Critical Systems / M. Cohen, C. Belta. – Cham: Springer, 2023. – 194 p.

7. Васильков Ю.В. Математическое моделирование объектов и систем автоматического управления / Ю.В. Васильков, Н.Н. Василькова. – Москва; Вологда: Инфра-Инженерия, 2025. – 428 с.

8. Семёнов А.Д. Математические модели систем управления / А.Д. Семёнов, А.В. Волков, О.В. Ермилина. – Москва; Вологда: Инфра-Инженерия, 2022. – 200 с.

9. Иванова А.А. Математическое моделирование тепловых процессов непрерывной разливки металлов: монография / А.А. Иванова, А.Б. Бирюков. – М.: Инфра-Инженерия, 2022. – 284 с.

10. Зароченцев В.М. Математическое и компьютерное моделирование в металлургии цинка: монография. – Барнаул: Новый формат, 2025. – 396 с.

11. Блочно-модульное моделирование и исследование бимолекулярного гетерогенного взаимодействия при обжиге сульфида металла / В.М. Зароченцев, А.Л. Рутковский, Т.В. Кондратенко, А.К. Макоева // Вестник Магнитогорского гос. техн. ун-та им. Г.И. Носова. – 2024. – Т. 22, № 1. – С. 20–28.

12. Математическая модель металлизации титаномагнетитовых руд Качканарского месторождения в шахтной печи Мидрекс / А.Н. Дмитриев, М.О. Золотых, Г.Ю. Витькина, Л.А. Маршук, М.С. Ялунин // Черная металлургия. Бюллетень научно-технической и экономической информации. – 2021. – Т. 77, № 10. – С. 1039–1045.

13. Особенности теории и практики идентификации объектов в системах управления / Л.П. Мышляев, Е.И. Львова, К.А. Ивушкин, Е.Н. Тараборина // Фундаментальные исследования. – 2014. – № 11–1. – С. 519–524.

14. Фуртат И.Б. Методы анализа и синтеза линейных и нелинейных систем управления при наличии возмущений и запаздывания: монография / И.Б. Фуртат, П.А. Гущин. – Москва; Ижевск: Ин-т компьютерных исследований, 2021. – 194 с.

15. Aguirre L.A. An Introduction to Nonlinear System Identification // Lectures on Nonlinear Dynamics / ed. by J.R. Castilho Piqueira, C.E. Nigro Mazzilli, C.P. Pesce, G.R. Franzini. – Cham: Springer, 2023. – P. 133–154.

Свинцов Максим Максимович
Аспирант каф. автоматизации и информационных систем
Сибирского государственного индустриального
университета (СибГИУ)
Кирова ул., 42, г. Новокузнецк, Россия, 654007
Нач. отд. систем технологического управления
Научно-исследовательского центра
систем управления» (ООО «НИЦ СУ»)
Строителей ул., 55, корп. А, оф. 25,
г. Новокузнецк, Россия, 654005
Тел.: +7-923-630-84-20
Эл. почта: svintzovmax@yandex.ru

Мышляев Леонид Павлович
Д-р техн. наук, проф., гл. н.с. ООО «НИЦ СУ»
Строителей ул., 55, корп. А, оф. 25,
г. Новокузнецк, Россия, 654005
ORCID: 0000-0002-0833-4772
Тел.: +7-905-911-51-16
Эл. почта: myshl_948@mail.ru

Макаров Георгий Валентинович
Канд. техн. наук, доцент каф. менеджмента качества
и инноваций Сибирского государственного
индустриального университета (СибГИУ)
Кирова ул., 42, г. Новокузнецк, Россия, 654007;
директор ООО «НИЦ СУ»
Строителей ул., 55, корп. А, оф. 25,
г. Новокузнецк, Россия, 654005
ORCID: 0000-0002-9933-4769
Тел.: +7-913-138-42-19
Эл. почта: gmakarov@nicsu.ru

Грачев Виталий Викторович
Канд. техн. наук, доцент каф. автоматизации
и информационных систем СибГИУ
Кирова ул., 42, г. Новокузнецк, Россия, 654007;
зам. дир. ООО «НИЦ СУ»
ул. Строителей, 55, корп. А, оф. 25,
г. Новокузнецк, Россия, 654005
ORCID: 0000-0002-0615-3820
Тел.: +7-909-509-67-76
Эл. почта: vitaly.grachev@nicsu.ru

Поступила в редакцию: 21.05.2026.
Принята к публикации: 20.08.2026.

Svintsov M.M., Myshlyaev L.P., Makarov G.V., Grachev V.V.
Mathematical models of control objects with coordinate-parametric couplings

Relevance*.** Accounting for coordinate-parameter couplings (CPCs) in control objects is an essential prerequisite for improving the efficiency of automated technological complexes in the mining, metallurgical, and chemical industries. ***Purpose of the study*.** To develop new structural models of objects taking into account the mutual influence of coordinate and parameter couplings, and to form a classification of new structures of mathematical models of control objects with CPCs. ***Methods*.** The methods of system analysis, automatic control theory, and mathematical modeling of dynamic systems with delay and parametric dependencies were used. ***Scientific novelty*.** The no-velty lies in the creation of mathematical models and structural schematics for control objects with CPCs that incorporate the mutual influence of coordinate and parametric variables (including for objects with recycling), as well as in the development of a classification framework for control objects exhibit-ting coordinate-parameter couplings. ***Results*.** Mathematical model structures for four classes of objects with CPCs are proposed, along with examples of such objects in a coal preparation process complex. ***Practical significance lies in expanding the toolset for analyzing and synthesizing effective automated technological complexes containing control objects with CPCs through the rapid joint determination of coordinate effects and object parameters.
Keywords: control system, automated technological complex, coordinate-parameter couplings (CPCs), recycling.
DOI: 10.21293/1818-0442-2026-29-2-135-139

References

1. Blauberg I.V., Yudin Eh.G. *Stanovlenie i sushchnost sistemnogo podkhoda* [Formation and essence of the systems approach]. – M.: Nauka, 1973, 270 p. (in Russ.).

2. Optner S.L. Systems analysis for business and industrial problem solving, N.J., Prentice-Hall, 1965, 116 p.

3. Laske O. Advanced Systems-Level Problem Solving, Volume 3: Manual of Dialectical Thought Forms. Cham, Springer, 2023, 234 p.

4. Emelyanov S.V., Nappelbaum Eh.L. *Osnovnye printsipy sistemnogo analiza* [Basic principles of systems analysis]. Pro. of the Second All-Union Scientific and Technical Conference «Problems of Scientific Organization of Management of Socialist Industry»], ed. by D.M. Gvishiani, S.E. Kamenitser, Moscow: Ehkonomika, 1974, pp. 92–99. (in Russ.).

5. Emelyanov S.V., Korovin S.K. *Novye tipy obratnoi svyazi: Upravlenie pri neopredelennosti* [New types of feedback: Management under uncertainty. Moscow: Nauka, Publishing company «Phys.-Math. Literature», 1997, 348 p. (in Russ.).

6. Cohen M., Belta C. Adaptive and Learning-Based Control of Safety-Critical Systems. Cham, Springer, 2023, 194 p.

7. Vasilkov Yu.V., Vasilkova N.N., *Matematicheskoe modelirovanie obĕktov i sistem avtomaticheskogo upravleniya* [Mathematical modeling of objects and automatic control systems]. M.; Vologda: Infra-Engineering, 2025, 428 p. (in Russ.).

8. Semenov A.D. Volkov A.V., Ermilina O.V. *Matematicheskie modeli sistem upravleniya* [Mathematical models of control systems]. Moscow; Vologda: Infra-Engineering, 2022, 200 p. (in Russ.).

9. Ivanova A.A., Biryukov A.B. *Matematicheskoe modelirovanie teplovy`x processov neprery`vnoj razlivki metallov: monografiya* [Mathematical modeling of thermal processes of continuous casting of metals: monograph]. M.: Infra-Engineering Publ., 2022, 284 p. (in Russ.).

10. Zarochentsev V.M. *Matematicheskoe i komp`yuternoe mode-lirovanie v metallurgii cinka: monografiya* [Mathematical and computer modeling in zinc metallurgy: monograph]. Barnaul: New Format Publ., 2025, 396 p. (in Russ.).

11. Zarochentsev V.M., Rutkovskii A.L., Kondratenko T.V., Makoeva A.K. [Block-modular simulation and study on a bimolecular heterogeneous interaction during roasting of metal sulfide]. *Vestnik Magnitogorskogo gosudarstvennogo tekhnicheskogo universiteta im. G.I. Nosova [Vestnik of Nosov Magnitogorsk State Technical University]*, 2024, vol. 22, no. 1, pp. 20–28 (in Russ.).

12. Dmitriev A.N., Zolotykh M.O., Vitkina G.Yu., Marshuk L.A. [A mathematical model of metallization of titanium-magnetite ores of kachkanar deposit in midrex shaft furnace]. *Ferrous metallurgy. bulletin of scientific, technical and economic information*, 2021, vol. 10, pp. 1039–1045 (in Russ.).

13. Myshlyaev L.P., Lvova E.I., Ivushkin K.A., Taraborina E.N. [Features of the theory and practice of identification of objects in control systems]. *Fundamental research*, 2014, № 11–1. pp. 519–524 (in Russ.).

14. Furtat I.B., Gushchin P.A. *Metody` analiza i sinteza linejny`x i nelinejny`x sistem upravleniya pri nalichii vozmushhenij i zapazdy`vaniya: monografiya* [Methods of analysis and synthesis of linear and nonlinear control systems in the presence of disturbances and delays: monograph]. Moscow; Izhevsk: Institute of Computer Research, 2021, 193 p. (in Russ.).

15. Aguirre L.A. An Introduction to Nonlinear System Identification. In the book: Lectures on Nonlinear Dynamics, ed. by J.R. Castilho Piqueira, C.E. Nigro Mazzilli, C.P. Pesce, G.R. Franzini. Cham: Springer, 2023, pp. 133–154.

Maksim M. Svintsov
Postgraduate student, Department of Automation and Information Systems, Siberian State Industrial University (SibSIU)
42, Kirova St., Novokuznetsk, Russia, 654007;
Head, Technological Control Systems Department, Limited Liability Company «Research Center for Control Systems»
55/A, Stroiteley Str., Off. 25, Novokuznetsk, Russia, 654005
Phone: +7-923-630-84-20
Email: svintzovmax@yandex.ru

Leonid P. Myshlyaev
Doctor of Technical Sciences, Professor
Chief Research Fellow at the Limited Liability Company «Research Center for Control Systems»
55/A, Stroiteley Str., Off. 25, Novokuznetsk, Russia, 654005
ORCID: 0000-0002-0833-4772
Phone: +7-905-911-51-16
Email: myshl_948@mail.ru

Georgy V. Makarov
Candidate of Technical Sciences, Associate Professor, Department of Quality Management and Innovation, SibSIU
42 Kirova St., Novokuznetsk, Russia, 654007;
Director of the Limited Liability Company «Research Center for Control Systems»
55/A, Stroiteley Str., Off. 25, Novokuznetsk, Russia, 654005
ORCID: 0000-0002-9933-4769
Phone: +7-913-138-42-19
Email: gmakarov@nicsu.ru

Vitaly V. Grachev
Candidate of Technical Sciences, Associate Professor, Department of Automation and Information Systems, SibSIU
42 Kirova St., Novokuznetsk, Russia, 654007;
Deputy Director of the Limited Liability Company «Research Center for Control Systems»
55/A, Stroiteley Str., Off. 25, Novokuznetsk, Russia, 654005
ORCID: 0000-0002-0615-3820
Phone: +7-909-509-67-76
Email: vitaly.grachev@nicsu.ru

Received: 21.05.2026.
Accepted: 20.08.2026.

УДК 004.94:681.7
Шифр специальности ВАК: 2.3.6

К.И. Чуков, В.А. Фаерман

Программный комплекс для моделирования оптического тракта системы квантовой коммуникации

Актуальность. Представлен программный комплекс для моделирования оптического тракта систем квантовой коммуникации, основанный на дискретно-событийной обработке фотонных посылок на единой временной шкале. Достоверное моделирование квантового канала необходимо для проектирования и исследования систем квантового распределения ключей. ***Цель исследования.*** Разработка программного комплекса для моделирования оптического канала и оценки влияния его параметров на характеристики системы. ***Методы.*** Применено дискретно-событийное моделирование. ***Научная новизна.*** Предложен программный фреймворк, объединяющий физический и протокольный уровни моделирования, а также математическое обеспечение для оценки совместного влияния физических факторов на характеристики системы. ***Результаты.*** Разработан программный комплекс, позволяющий оценивать влияние параметров канала и оборудования на QBER и скорость генерации ключа; корректность модели подтверждена вычислительными экспериментами. ***Практическая значимость.*** Комплекс предназначен для проектирования, прототипирования и верификации систем квантового распределения ключей.
Ключевые слова: квантовый канал, волоконно-оптическая линия связи, моделирование, лазерный импульс, однофотонный детектор, C++.
DOI: 10.21293/1818-0442-2026-29-2-140-147

Введение

Квантовое распределение ключей (КРК) представляет собой технологию, предназначенную для обеспечения защиты коммуникационных каналов в условиях нарастающей угрозы компрометации традиционных криптосистем, основанных на использовании открытого ключа [1].

Развитие данной технологии происходит во всём мире. По данным Kept (входит в перечень организаций «квантовой экосистемы» ОАО «РЖД»), глобальный рынок квантовых коммуникаций растёт на 31% каждый год, его объём оценивается в 600 млн долл. К 2030 г. он достигнет отметки в 3 млрд долл. [2, 3].

Для изучения и совершенствования систем квантовой коммуникации применяется компьютерное моделирование [4]. Оно позволяет исследовать как верхнеуровневое функционирование системы (протоколы выработки общего ключа, коррекция ошибок и др.), так и низкоуровневое (генерация и рассеивание импульсов, срабатывания детекторов и др.). Это открывает широкие возможности для решения ряда актуальных научно-технических задач, однако большинство существующих моделей рассматривают отдельные компоненты или процессы изолированно, без учета функционирования системы в целом.

Целью данной работы является создание программного комплекса, который позволит сформулировать требования при проектировании и прототипировании оптических каналов систем квантовой коммуникации. Программный комплекс позволит как генерировать данные для их непосредственного анализа, так и для дальнейшего использования в других моделях (например, для согласования ключевого материала или усиления секретности). Модернизация комплекса позволит создать платформу, интегрирующую полный технологический стек КРК.

Обзор комплексов для моделирования систем КРК

Исследование характеристик и особенностей функционирования систем КРК требует применения специализированных программных инструментов. В настоящее время существует широкий спектр решений, созданных как на базе универсальных платформ моделирования [5–12] (MatLab Simulink, COMSOL Multiphysics, NS-3, OMNeT++), так и реализованных как самостоятельные программные комплексы на языках C++, Python и др.

Поиск по запросу «QKD simulation software» в поисковой системе Google Scholar за период с 2021 по 2025 г. показывает около 9550 материалов.

Часто встречающимся инструментом является QKDNetSim [5, 6], разработанный на базе симулятора сетевых протоколов NS-3 [7]. Комплекс ориентирован на изучение поведения сетей, использующих КРК, и позволяет моделировать генерацию и распространение ключей между узлами. Однако данный инструмент не включает моделирование физических эффектов в оптическом тракте, что ограничивает его применение.

OpenQKDSecurity [8], представляет собой модель для оценки нижней границы скорости выработки безопасного ключа. Комплекс использует инструменты MatLab Simulink и позволяет исследовать эффективность различных схем КРК. Основным ограничением является зависимость от проприетарного программного обеспечения MatLab Simulink, которое официально не доступно к покупке.

Среди проприетарных решений, которые возможно приобрести на территории Российской Федерации, можно выделить ViPNet QKDSim [9], поддерживающий протоколы BB84, B92 и ГОКС [10]. Программа позволяет учитывать параметры подключае-

мой аппаратной периферии, такие как мощность лазерного излучения, чувствительность детекторов и уровень шумов, а также моделировать различные типы атак. Однако данный продукт в основном предназначен больше для обучения специалистов, чем для проведения исследований.

На базе платформы OMNeT++ построен qkdX [11]. Комплекс реализует несколько протоколов КРК и поддерживает моделирование физических эффектов и явлений. Тем не менее исходный код программы не находится в открытом доступе, что делает невозможным расширение модели или адаптацию под исследовательские задачи.

Среди полностью свободных решений выделяется комплекс SeQUeNCe [12], разработанный на языке Python. Он предоставляет широкий набор возможностей для моделирования квантовых коммуникаций, включая протоколы КРК и физические процессы в канале передачи. Ограничением является сравнительно невысокая производительность, характерная для интерпретируемых языков программирования, что осложняет моделирование больших сетей или длительных процессов передачи.

Ниже представлена сравнительная таблица программных решений для моделирования систем КРК.

В ходе обзора было определено, что существующие программные платформы характеризуются значительным разнообразием функциональных возможностей, однако на данный момент не существует универсального и производительного программного решения, которое бы объединяло низкоуровневое и верхнеуровневое функционирование системы.

В связи с этим было решено разработать программный комплекс:

– обеспечивающий одновременный учет физических процессов в оптическом тракте и параметров протокольного уровня;

– реализующий дискретно-событийную обработку фотонных посылок на единой временной шкале, что позволяет учитывать временные эффекты взаимодействия компонентов;

– обладающий открытой архитектурой, допускающей модификацию и расширение модели;

– обеспечивающий более высокую производительность по сравнению с решениями на интерпретируемых языках, что позволяет моделировать интенсивные потоки импульсов.

Программные комплексы для моделирования системы КРК

Название решения	Работа в реальном времени	Функциональные возможности	Ограничения	Среда моделирования	Тип лицензии	Источник
QKDNetSim	Нет	Генерация, управление и распространение ключей среди узлов сети	Отсутствует моделирование физических эффектов и явлений, отсутствует реализация протоколов КРК	NS-3 (разработан на C++)	Свободная	[5, 6]
OpenQKDSecurity	Нет	Расчёт нижней границы скорости выработки ключа	Отсутствует моделирование физических эффектов и явлений, отсутствует реализация протоколов КРК, зависит от библиотек MatLab, нет поддержки многопоточности	MatLab	Модель распространяется под свободной лицензией, однако MatLab распространяется под проприетарной лицензией	[8]
ViPNet QKDSim	Программный продукт позволяет подключать аппаратную платформу и анализировать поступающую информацию в процессе работы	Демонстрация режимов однофотонного и квазиоднофотонного режимов. Поддерживает протоколы BB84, B92 и ГОКС. Демонстрирует влияние параметров оборудования, таких как чувствительность и шумы детектора, а также влияние атак	Является закрытым ПО	Открытой информации нет	Проприетарное ПО	[9]
qkdX	Нет	Поддержка протоколов различных протоколов КРК. Моделирование физических эффектов и явлений	Является закрытым ПО	OMNeT++	Исходный код не был предоставлен в общий доступ	[11]
SeQUeNCe	Нет	Поддержка протоколов различных протоколов КРК. Моделирование физических эффектов и явлений	Производительность ограничена языком программирования Python	Разработан на языке программирования Python	Свободная	[12]

Описание архитектуры программного комплекса

Для реализации программного комплекса было решено логически разделить его на две составляющие: 1) управляющая утилита; 2) сервисы (набор модулей для моделирования с общим API-интерфейсом). Схема взаимодействия на рис. 1.

Управляющая утилита может представлять собой как программу собственной разработки на каком-либо языке программирования (например, Python), так и инструмент командной строки (например, curl). Основное назначение данной утилиты – организовать взаимодействие с сервисами.

Сервис моделирования включает в себя HTTP-сервер для обработки API-запросов и специализированную библиотеку для моделирования. Ключевым условием для обеспечения унифицированного взаимодействия является соответствие REST API. Выбор языка программирования для реализации сервиса определяется исключительно из задач моделирования.

Требования к сервису моделирования

Для реализации сервиса для моделирования оптического тракта системы КРК были составлены следующие требования:

1. *Язык программирования.* Сервис должен обеспечивать высокую производительность за счёт эффективного управления памятью и многопоточностью, так как количество обрабатываемых импульсов может насчитывать миллионы единиц. В связи с этим решено было использовать C++, который является популярным выбором в высоконагруженных системах.

2. *Структура программного обеспечения.* Структура сервиса должна обеспечивать возможность добавления нового функционала и модернизацию старого. Для соответствия этому требованию разрабатываемый комплекс должен соответствовать базовым принципам объектно-ориентированного программирования (ООП) и SOLID в частности.

3. *Контейнеризация.* Разрабатываемый сервис должен работать в контейнере, чтобы обеспечить переносимость и масштабируемость. Это позволит разворачивать программный комплекс на любой платформе без необходимости устанавливать все зависимости на собственной системе или использовать тяжеловесную виртуальную машину.

4. *Сторонние библиотеки.* Использование сторонних библиотек может негативно сказаться на читаемости программного кода. Поэтому решено было ограничиться только теми библиотеками, которые необходимы для работы с JSON-файлами и HTTP-сервером.

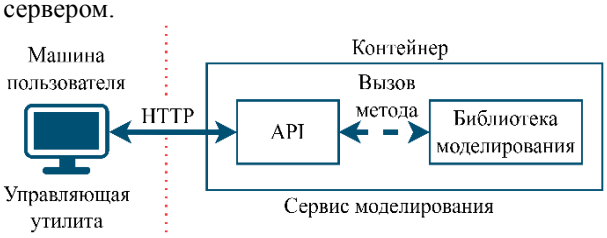


Рис. 1. Схема взаимодействия управляющий утилиты и сервиса моделирования

Описание архитектуры сервиса моделирования

Структура сервиса моделирования представляет собой API-сервер и библиотеку для моделирования (рис. 2), на рис. 3 представлена диаграмма классов. HTTP-сервер представляет собой классы APIRouter и APIController. Библиотека состоит из интерфейсов: ILaser, ISequenceGenerator, IModulator, IQuantumChannel, IBeamSplitter, IPhotodetector, а также классов: AttLaser, mtPNRG, PolarizationModulator, FOCL, PBS, SPAD.

Назначение классов HTTP-сервера:

− APIRouter отвечает за маршрутизацию входящих HTTP-запросов к соответствующим обработчикам API-методов.

− APIController реализует логику обработки запросов. Принимает данные для конфигурации, инициирует запуск выбранного протокола, позволяет получить результаты моделирования и др.

Для реализации HTTP-сервера была выбрана «header-only» библиотека httplib [13], которая позволяет использовать нужный функционал без необходимости инсталляции различных зависимостей. Данная библиотека открыто распространяется в сети и имеет активное сообщество разработчиков.

Назначение классов сервиса моделирования:

− AttLaser – класс для моделирования лазера с аттенюатором. Учитывается распределение Пуассона для числа фотонов в импульсе, скорость генерации импульсов [14].

− mtPNRG – класс, который реализует интерфейс ISequenceGenerator на основе алгоритма Mersenne Twister. Класс обеспечивает псевдослучайное распределение битов и базисов с большой длиной периода.

− PolarizationModulator – класс, который позволяет моделировать поляризационный модулятор. Реализует преобразование к матрицам Джонса для горизонтальной, вертикальной, диагональной и антидиагональной поляризаций.

− FOCL – класс для моделирования передачи фотонов через волоконно-оптическую линию связи (ВОЛС). Учитывается влияние затухания сигнала и хроматической дисперсии при различной протяженности канала.

− PBS – класс, который представляет собой реализацию интерфейса IBeamSplitter и моделирует поведение поляризационного светоделителя.

– SPAD – класс, реализующий однофотонный лавинный фотодетектор (ОЛФД). Позволяет задавать основные характеристики ОЛФД, такие как квантовая эффективность, темновой счёт, мёртвое время, вероятность послеимпульса [15].

Описание REST-интерфейса

В предложенной реализации для взаимодействия с библиотекой моделирования предусмотрено 5 основных запросов:

1. *Загрузка конфигурации.* С помощью управляющей утилиты формируется POST-запрос для передачи входных параметров моделирования, например, протокол КРК, мощность лазера, длина квантового канала и др.

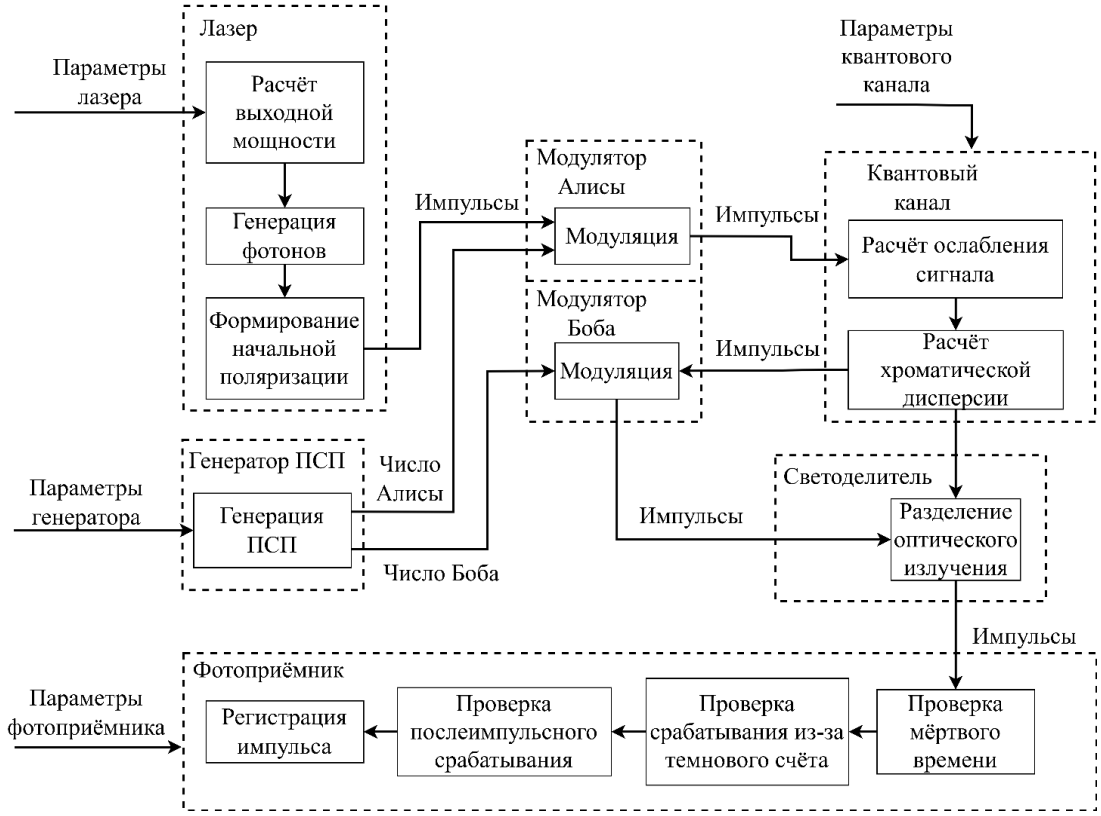


Рис. 2. Архитектура библиотеки моделирования

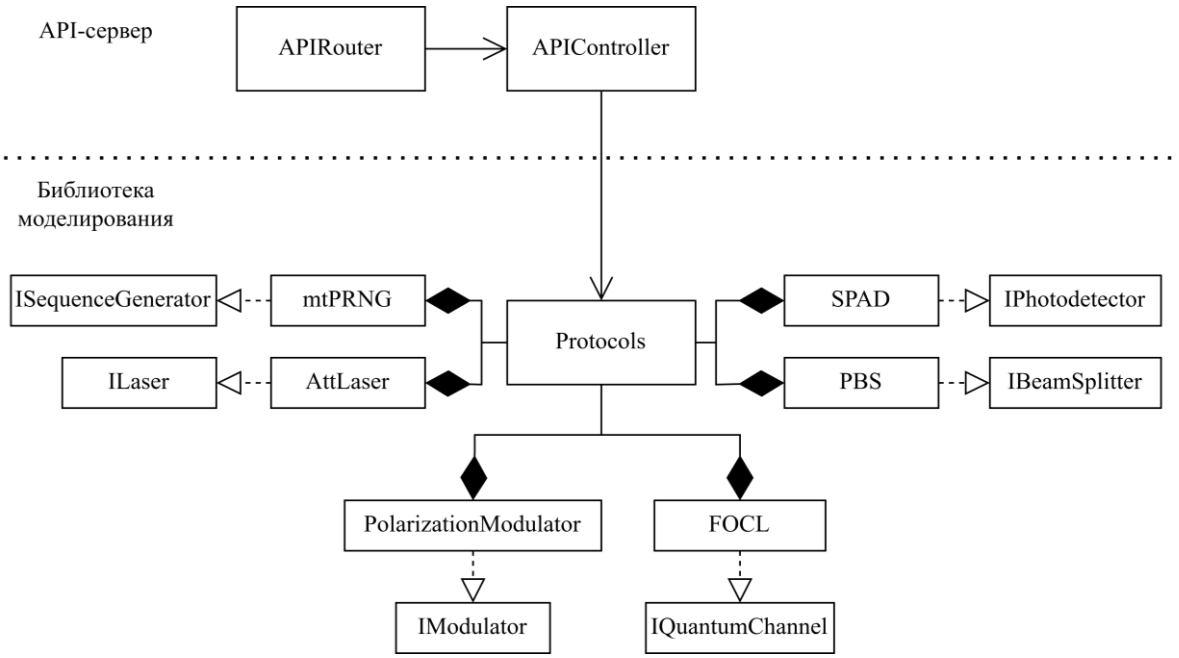


Рис. 3. Диаграмма классов сервиса моделирования

2. *Запуск моделирования.* Для запуска моделирования отправляется простой JSON-объект следующего вида: {"start": true}. В ответ формируется уникальный идентификатор (UUID), соответствующий запущенному сценарию.

3. *Результат моделирования.* По ранее полученному UUID пользователь может получить результаты запущенного сценария. Результаты предоставляются в формате CSV и содержат как входные параметры, так и выходные: выбранные базисы Алисы и Боба, результат измерения Боба, а также просеянный ключ.

4. *Остановка моделирования.* При работе с сервисом может потребоваться прервать его текущую задачу. Для выполнения данной функции необходимо осуществить отправку соответствующего POST-запроса.

5. *Статус моделирования.* Посредством GET-запроса пользователь может получить информацию,

доступен ли сервис для моделирования в текущий момент.

По завершении моделирования формируется csv-файл с итоговыми результатами. В csv-файле присутствуют входные параметры и следующие заголовки:

• total_pulses – количество отправленных импульсов;

• detected_H – количество зарегистрированных импульсов на детекторе, регистрирующем горизонтальные базисы;

• detected_H – количество зарегистрированных импульсов на детекторе, регистрирующем вертикальные базисы;

• raw_key_length – длина сырого ключевого материала;

• QBER – отношение корректно принятых фотонных посылок к отправленным;

• key_rate – скорость выработки ключа;

• sifted_key – просеянный ключевой материал.

Проведение экспериментов

В рамках работы эксперименты проводились с целью поэтапной верификации разработанной физико-математической модели [16] и программной реализации системы квантового распределения ключей.

Экспериментальная методика строилась на последовательном усложнении условий моделирования: от идеализированных режимов, позволяющих изолированно проверить корректность алгоритмов и протокольной логики, к более реалистичным сценариям, учитывающим затухание в волоконно-оптическом канале, хроматическую дисперсию, конечную квантовую эффективность фотодетекторов, а также ограничения, связанные с временным разрешением и мёртвым временем регистрации детекторов.

Разработанный программный комплекс позволяет проводить параметрические исследования, основанные на систематическом варьировании характеристик основных элементов оптического тракта:

– источника излучения (длина волны, длительность импульса, частота генерации, среднее число фотонов, задаваемое через мощность и аттенюацию), что позволяет анализировать влияние временного перекрытия импульсов и статистики фотонного распределения на вероятность регистрации и уровень ошибок;

– волоконно-оптического канала (длина линии, коэффициент затухания, хроматическая дисперсия), что даёт возможность оценить деградацию сигнала при распространении и её влияние на квантовую битовую ошибку и скорость генерации ключа. При этом предусмотрена возможность изолированного анализа отдельных эффектов за счёт их программного отключения;

– фотодетектора (квантовая эффективность, мёртвое время, вероятность послеимпульсных срабатываний, временное окно регистрации), что позволяет исследовать влияние ограничений однофотонной регистрации на точность измерений и пропускную способность системы.

Помимо однофакторного анализа, комплекс поддерживает проведение многофакторных экспериментов, в рамках которых варьируются несколько параметров одновременно. Это позволяет выявлять нелинейные эффекты и области параметров, в которых наблюдается существенное ухудшение характеристик системы, например, при совместном влиянии длины канала и мёртвого времени детектора.

В результате моделирования формируются выходные данные, на основе которых строятся функциональные зависимости, отражающие поведение системы при изменении параметров, в частности:

$$\mathrm{QBER} = f(L), R = f(\nu, \tau), \mathrm{QBER} = f(\tau_d, \eta),$$

где L – длина канала, ν – частота генерации импульсов, τ – длительность импульса, τ_d – мёртвое время детектора, η – квантовая эффективность.

На рис. 4 представлен график для эксперимента со следующими параметрами: длина волны лазерного излучения 1550 нм; длина канала 50 км; коэффициент хроматической дисперсии 18 пс/(нм·км); временное окно детектора 2 нс; мёртвое время детектора 4 нс; квантовая эффективность детектора 50%. В качестве основных метрик оценки использовались скорость генерации ключа и отношение зарегистрированных импульсов к отправленным (C_{reg}), что позволяет комплексно охарактеризовать как информационную, так и физическую эффективность системы КРК.

Экспериментальные исследования влияния хроматической дисперсии демонстрируют, что возникающие искажения имеют параметрическую природу и проявляются в характеристиках системы только при определенных сочетаниях длительности импульса, частоты генерации и временного разрешения детектора [16].

Экспоненциальный рост QBER и снижение скорости генерации ключа при увеличении длины канала и при учёте затухания согласуются с теоретическими моделями распространения квантовых состояний в ВОЛС и подтверждают адекватность учёта потерь [17].

Анализ влияния квантовой эффективности фотодетектора и параметров импульсного режима демонстрирует, что модель корректно воспроизводит реальные ограничения однофотонной регистрации, включая влияние временного окна и мёртвого времени детектора.

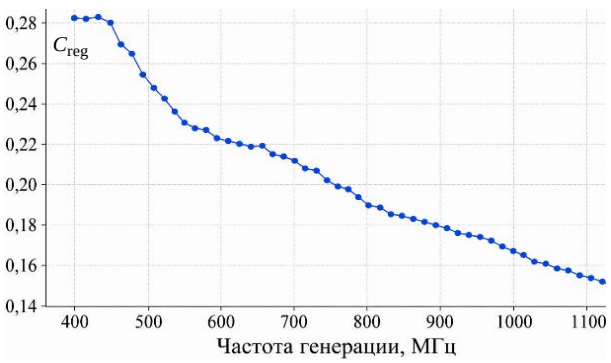


Рис. 4. Зависимость отношения зарегистрированных импульсов от частоты генерации и длительности импульсов

Таким образом, предложенный программный комплекс обеспечивает проведение воспроизводимых вычислительных экспериментов и позволяет исследовать параметрическое пространство модели оптического тракта с учётом временных и вероятностных эффектов, возникающих при передаче и регистрации фотонных посылок, а полученные зависимости подтверждают корректность реализованной модели и алгоритмов.

Заключение

В работе описан разработанный программный комплекс для моделирования оптического тракта системы квантовой коммуникации. Программный комплекс ориентирован на комплексный учёт параметров оптических элементов тракта, на эксплуатационные характеристики системы в целом.

Проведённый обзор существующих решений показал отсутствие универсального и производительного инструмента, объединяющего низкоуровневое моделирование квантового канала и удобные средства интеграции. Предложена архитектура, основанная на сервисной модели с REST API и реализации ключевых компонентов на языке C++, архитектура комплекса ориентирована на эффективную обработку интенсивных потоков событий за счёт реализации на языке C++, модульность и расширяемость, что позволяет адаптировать комплекс под различные исследовательские сценарии.

Предложен дискретно-событийный подход к моделированию оптического тракта систем квантовой коммуникации, основанный на последовательной обработке фотонных посылок на единой временной шкале, что позволяет учитывать взаимное влияние временных характеристик элементов системы (длительности импульсов, мёртвого времени детекторов, временных окон регистрации). Разработана модель представления оптического тракта в виде композиции взаимодействующих компонентов, в которой каждый элемент описывается как преобразование потока событий, что обеспечивает модульное расширение модели без нарушения её целостности. Обеспечена возможность проведения вычислительных экспериментов с высокой интенсивностью потока событий за счёт эффективной реализации алгоритмов обработки посылок на языке C++.

Результаты вычислительных экспериментов подтверждают корректность реализованной физико-математической модели и адекватность учёта основных факторов, влияющих на характеристики систем КРК, включая затухание, дисперсию и параметры однофотонных детекторов. Полученные зависимости квантовой битовой ошибки и скорости генерации ключа согласуются с теоретическими оценками, что позволяет рассматривать разработанный программный комплекс как перспективный инструмент для анализа, верификации и дальнейшего развития систем квантового распределения ключей.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках базовой части государственного задания ТУСУРа на 2026–2028 гг. (проект FEWM-2026-0009).

Литература

1. Реализация средства криптографической защиты информации, использующего квантовое распределение ключей / А.Г. Втюрина, В.Л. Елисеев, А.Е. Жиляев, А.С. Николаева, В.Н. Сергеев, А.В. Уривский // Доклады ТУСУР. – 2018. – Т. 21, № 2. – С. 15–21.

2. 50-кубитный компьютер, шифратор в ЦОД и генераторы для 5G. Как развивалась квантовая индустрия в 2024 году // СберПро [Электронный ресурс]. – URL: https://sber.pro/publication/50-kubitnii-kompyuter-shifrator-v-tsod-i-generatori-dlya-5g-kak-razvivalas-kvantovaya-industriya-v-2024-godu/?utm_source=vc_sberpro (дата обращения: 12.11.2025).

3. Центр экспертизы в области квантовых технологий // Kept. – 2024 [Электронный ресурс]. – URL: https://assets.kept.ru/upload/pdf/2023/09/ru-kept-quantum-services.pdf (дата обращения: 12.11.2025).

4. Система массового обслуживания в сетях квантового распределения ключей / Д.С. Гекк, В.Д. Пинчук, Д.В. Ожигов, А.О. Терехин // Электронные средства и системы управления: материалы докладов XX Междунар. науч.-практ. конф. (20–22 ноября 2024 г.): в 2 ч. – Ч. 2. – Томск: В-Спектр (ИП В.М. Бочкарева); ТУСУР, 2024. – С. 78–80.

5. Dervisevic E. Large-scale quantum key distribution network simulator / E. Dervisevic, M. Voznak, M. Mehic // Journal of Optical Communications and Networking. – 2024. – Vol. 16, No. 4. – P. 449–462.

6. Soler D. QKDNetSim+: Improvement of the quantum network simulator for NS-3 / D. Soler, I. Cillero et al. // SoftwareX. – 2024. – Vol. 26. – P. 101685.

7. Riley G.F. The ns-3 network simulator. Chapter in Modeling and tools for network simulation / G.F. Riley, T.R. Henderson. – Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. – 537 p.

8. Winick A. Reliable numerical key rates for quantum key distribution / A. Winick, N. Lütkenhaus, P.J. Coles // Quantum. – 2018. – Vol. 2. – P. 77.

9. ViPNet QKDSim. Программный комплекс симуляции квантового распределения ключей [Электронный ресурс]. – URL: https://infotecs.ru/products/vipnet-qkdsim/ (дата обращения: 12.11.2025).

10. Жиляев А.Е. Сети квантового распределения ключей в кибербезопасности / А.Е. Жиляев, А.Г. Сабанов, А.А. Шелупанов, А.А. Конев, Д.С. Брагин. – М.: Горячая линия – Телеком, 2023. – 152 с.

11. Engle R.D.L. Modeling quantum optical components, pulses and fiber channels using om-net++ / R.D.L. Engle, D.D. Hodson, et al. / arXiv e-prints. arXiv preprint. – 2015. – Art. no. arXiv:1509.03091. DOI: 10.48550/arXiv.1509.03091.

12. Wu X. SeQUeNCe: a customizable dis-crete-event simulator of quantum networks / X. Wu, A. Kolar, et al. // Quantum Science and Technology. – 2021. – Vol. 6, No. 4. – P. 045027.

13. cpp-httplib // github.com. – URL: https://github.com/yhirose/cpp-httplib (дата обращения: 24.10.2025).

14. Махорин Д.А. Статистический контроль распределения числа фотонов в информационных сообщениях в системе квантового распределения ключа с временным кодированием / Д.А. Махорин, А.С. Задорин, С.Ю. Решетников // Электронные средства и системы управления: матер. докладов междунар. науч.-практ. конф. – Томск: ТУСУР, 2015. – № 1-1. – С. 270–273.

15. Курочкин Ю.В. Детекторы одиночных фотонов на основе лавинных фотодиодов / Ю.В. Курочкин, В.Л. Курочкин // Известия вузов. Физика. – 2011. – Т. 54, №. 2-2. – С. 202–205.

16. Динамическая модель оптического канала системы квантового распределения ключей в среде MatLab Simulink / К.И. Чуков, В.А. Фаерман, Д.В. Курков, Е.В. Антропов // Известия Том. политехн. ун-та. Промышленная кибернетика. – 2025. – Т. 3, № 2. – С. 24–35.

17. Ландсберг Г.С. Оптика: учеб. пособие. – 6-е изд., стер. – М.: Физматлит, 2003. – 848 с.

18. A practical transmitter device for passive state BB84 quantum key distribution / Yu. Kurochkin, M. Papadovasilakis, A. Trushechkin, R. Piera, J.A. Grieve // arXiv e-prints. – 2024. – Art. no. arxiv.org/html/2405.08481v1.

Чуков Кирилл Игоревич
Аспирант каф. безопасности информационных систем (БИС) Томского государственного университета систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0009-0005-5029-3504
Тел.: +7 (382-2) 70-15-29
Эл. почта: kl-chukov@yandex.ru

Фаерман Владимир Андреевич
Ст. преп. каф. комплексной информационной безопасности электронно-вычислительных систем ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0002-9643-0245
Тел.: +7 (382-2) 70-15-29
Эл. почта: fva@fb.tusur.ru

Поступила в редакцию: 28.01.2026.
Принята к публикации: 17.07.2026.

Chukov K.I., Fayerman V.A.

Software package for modeling the optical path of a quantum communication system

The paper presents a software package for modeling the optical path of quantum communication systems based on discrete event processing of photonic packets on a unified time scale. ***Relevance.*** Reliable modeling of a quantum channel is essential for the design and investigation of quantum key distribution systems. ***Purpose of the study.*** Development of a software package for modeling an optical channel and evaluating the effect of its parameters on system characteristics. ***Methods.*** Discrete event modeling was employed. ***Scientific novelty.*** A software framework is proposed that integrates the physical and protocol levels of modeling, together with mathematical support for assessing the combined influence of physical factors on system performance. ***Results.*** A software package has been developed to assess the influence of channel and equipment parameters on the quantum bit error rate (QBER) and key generation rate; the model's validity has been confirmed through computational experiments. ***Practical significance.*** The software package is intended for the design, prototyping and verification of quantum key distribution systems.

Keywords: quantum channel, fiber-optic communication line, simulation, laser pulse, single-photon detector, C++.

Funding: The work was supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of the basic part of the TUSUR state assignment for 2026-2028 (project FEWM-2026-0009).
DOI: 10.21293/1818-0442-2026-29-2-140-147

References

1. Vtyurina A.G., Eliseev V.L., Zhilyaev A.E., Nikolaeva A.S., Sergeev V.N., Urivskiy A.V. [On the principal decisions of the practical implementation of the cryptographic devices with quantum key distribution]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2018, vol. 21, no. 2, pp. 15–21.

2. *50-kubitnyy komp'yuter, shifrator v TSOD i generatory dlya 5G. Kak razvivalas' kvantovaya industriya v 2024 godu* [50-qubit computer, data center encoder and generators for 5G. How the quantum industry developed in 2024]. Available at: https://sber.pro/publication/50-kubitnii-kompyuter-shifrator-v-tsod-i-generatori-dlya-5g-kak-razvivalas-kvanto-vaya-industriya-v-2024-godu/?utm_source=vc_sberpro (accessed: 11 December 2025) (in Russ.)

3. *Tsentr ekspertizy v oblasti kvantovykh tekhnologiy* [Center of expertise in the field of quantum technologies]. Kept, 2024. Available at: https://assets.kept.ru/upload/pdf/2023/09/ru-kept-quantum-services.pdf (accessed: 12 November 2025).

4. Gekk D.S., Pinchuk V.D., Ozhigov D.V., Terekhin A.O. *Sistema massovogo obsluzhivaniya v setyakh kvantovogo raspredeleniya klyuchey* [Queuing system in quantum key distribution networks]. Elektronnyye sredstva i sistemy upravleniya [*Electronic means and control systems*]. Proc. of the International Scientific and Practical Conference, Tomsk, V-Spektr (Bochkareva V.M.), TUSUR, 2024, 2 part, pp. 78–80 (in Russ.).

5. Dervisevic E., Voznak M., Mehic M. Large-scale quantum key distribution network simulator. *Journal of Optical Communications and Networking*, 2024, vol. 16, no. 4, pp. 449–462.

6. Soler D. et al. QKDNetSim+: Improvement of the quantum network simulator for NS-3. *SoftwareX*, 2024, vol. 26, pp. 101685.

7. Riley G. F., Henderson T. R. The ns-3 network simulator. In book: Modeling and tools for network simulation. Berlin, Heidelberg, Springer Berlin Heidelberg, 2010, pp. 15–34.

8. Winick A., Lutkenhaus N., Coles P. J. Reliable numerical key rates for quantum key distribution. *Quantum*, 2018, vol. 2, pp. 77.

9. ViPNet QKDSim. *Programmnyy kompleks simulyatsii kvantovogo ras-predeleniya klyuchey*. Available: https://infotecs.ru/products/vipnet-qkdsim (accessed: 11 December 2025) (in Russ.).

10. Zhilyaev A.E., Sabanov A.G., Shelupanov A.A., Konev A.A., Bragin D.S. Quantum key distribution networks in cybersecurity. Moscow: Goryachaya liniya – Telekom Publ., 2023, 152 p. (in Russ.).

11. Engle R. D. L. et al. Modeling quantum optical components, pulses and fiber channels using om-net++. arXiv preprint, 2015, Art. no. arXiv:1509.03091,2015. DOI: 10.48550/arXiv.1509.03091.

12. Wu X. et al. SeQUeNCe: a customizable dis-crete-event simulator of quantum networks. Quantum Science and Technology, 2021, vol. 6, no. 4, pp. 045027.

13. cpp-httplib. Available: https://github.com/yhirose/cpp-httplib (accessed: 24 October 2025).

14. Makhorin D.A., Zadorin A.S., Reshetnikov S.Y. *Statisticheskiy kontrol' raspredeleniya chisla fotonov v infor-*

matsionnykh soobshcheniyakh v sisteme kvantovogo raspredeleniya klyucha s vremennym kodirovaniyem [Statistical control of the distribution of the number of photons in information messages in the system of quantum key distribution with time coding] Elektronnyye sredstva i sistemy upravleniya [*Electronic means and control systems*]. Proc. of the International Scientific and Practical Conference, Tomsk: TUSUR, 2015, no. 1-1, pp. 270–273 (in Russ.).

15. Kurochkin Yu.V., Kurochkin V.L. *Detektory odinochnykh fotonov na osnove lavinnykh fotodiodov* [Single photon detectors based on avalanche photodiodes]. *Izvestiya vuzov. Fizika,* 2011, vol. 54, no. 2-2, pp. 202–205 (in Russ.).

16. Chukov K.I., Fayerman V.A., Kurkov D.V., Antropov E.V. [Dynamic model of an optical channel for a quantum key distribution system in MatLab Simulink]. *Bulletin of the Tomsk Polytechnic University. Industrial cybernetics*, 2025, vol. 3, no. 2, pp. 24–35 (in Russ.).

17. Landsberg G.S. *Optika* [Optics]. A textbook, 6th ed., reprinted, Moscow, Fizmatlit, 2003, 848 p (in Russ.).

18. Kurochkin Yu., Papadovasilakis M., Trushechkin A., Piera R., Grieve J.A. A practical transmitter device for passive state BB84 quantum key distribution. arXiv e-prints, 2024, Art. no. arxiv.org/html/2405.08481v1.

Kirill I. Chukov
Postgraduate student, Department of Television and Control, Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0009-0005-5029-3504
Phone: +7 (382-2) 70-15-29
Email: kl-chukov@yandex.ru

Vladimir A. Fayerman
Senior Lecturer, Department of Radio Engineering Systems, TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0002-9643-0245
Phone: +7 (382-2) 70-15-29
Email: fva@fb.tusur.ru

Received: 28.01.2026.
Accepted: 17.07.2026.

УДК 004.056.5
Шифр специальности ВАК: 2.3.6

С.Е. Штука, А.Е. Боршевников

Модификация нейросетевого алгоритма высоконадёжной биометрической аутентификации по радужной оболочке глаза

Актуальность. Рост числа систем, обрабатывающих конфиденциальные данные, требует механизмов аутентификации, устойчивых к компрометации хранимых эталонов. ***Цель исследования.*** Модификация нейросетевого алгоритма высоконадёжной биометрической аутентификации по радужной оболочке глаза, обеспечивающая восстановление криптографического ключа с точным побитовым совпадением. ***Методы.*** Свёрточная нейронная сеть для выделения области радужной оболочки, нормализация кадрированием, совместное обучение свёрточной модели для извлечения признаков и многослойного персептрона с использованием комбинированной функции потерь, включающей метрические слагаемые и два слагаемых, действующих непосредственно в битовом выходе модели; выбор рабочей точки по критерию допустимости; перебор в окрестности Хэмминга при проверке. ***Научная новизна.*** В отличие от известных схем, точное восстановление ключа достигается без вспомогательных данных и внешней коррекции ошибок, что позволяет хранить только хэш ключа. ***Результаты.*** Для 100 зарегистрированных пользователей у экземпляров, допущенных к эксплуатации, доля истинного принятия составила 0,875 при нулевой вероятности ложного допуска на 8000 предъявлений «Чужих»; перебор в окрестности радиуса два бита снижает вероятность ошибки первого рода до 0,107, не изменяя вероятность ошибки второго рода. Отказ по обеим попыткам при точном совпадении ключа наблюдался у 1 пользователя из 100. Показано, что запас, обеспечиваемый на группе подбора порога, переносится на тестовую выборку лишь частично. ***Практическая значимость.*** Алгоритм применим в системах, где недопустимо хранение биометрических эталонов и криптографических ключей в открытом виде.

Ключевые слова: биометрическая аутентификация, радужная оболочка, сегментация изображений, нейронные сети, криптографический ключ, Triplet Loss, Center Loss, биометрическая криптография.
DOI: 10.21293/1818-0442-2026-29-2-148-159

Введение

Развитие цифровых сервисов и распределённых информационных систем требует повышения надёжности механизмов аутентификации пользователей. Данная задача особенно актуальна для систем, обрабатывающих конфиденциальные данные и критически важную информацию.

Используемые на практике парольные схемы аутентификации обладают рядом ограничений, связанных с возможностью компрометации учётных данных вследствие утечек баз данных, фишинга и человеческого фактора. Альтернативой служат биометрические технологии, использующие уникальные физиологические характеристики человека. Такой подход не требует от человека запоминания дополнительной информации и наличия дополнительных устройств [1]. Среди различных биометрических характеристик особый интерес представляет радужная оболочка глаза, отличающаяся высокой сложностью текстурного рисунка, значительной энтропией признаков и стабильностью во времени [1, 2]. Индивидуальность и неизменность радужной оболочки на протяжении жизни делают её перспективной основой для высоконадёжных систем аутентификации.

Целью настоящей работы является модификация нейросетевого алгоритма высоконадёжной биометрической аутентификации по радужной оболочке глаза, обеспечивающая точное побитовое восстановление криптографического ключа и, как следствие, возможность хранения только хэша ключа. Для достижения поставленной цели решались следующие задачи: сравнение архитектур сегментации в единых условиях и выбор архитектуры для конвейера; выбор способа нормализации выделенной области; разработка функции потерь, обеспечивающей одновременно восстановление ключа «Своего» и удаление кодов «Чужих» от базы ключей; формирование процедуры выбора рабочей точки, гарантирующей нулевую оценку вероятности ошибки второго рода. Научная новизна работы состоит в переходе от порогового решения по расстоянию Хэмминга к решению по точному совпадению ключа без использования вспомогательных данных, а также во введении двух дополнительных членов функции потерь, действующих непосредственно в битовом выходе модели.

Основные понятия и критерии оценки

Согласно ГОСТ Р 52633.0–2006 [3], под высоконадёжной биометрической аутентификацией понимается биометрическая аутентификация с приемлемой вероятностью ошибок первого рода и гарантированно малой вероятностью ошибок второго рода, сопоставимой по своему значению с вероятностью случайного подбора кода неизвестного криптографического ключа при малом числе попыток подбора.

Высоконадёжная биометрическая аутентификация является частным случаем биометрической аутентификации и наследует её классическую архитектуру, включающую три последовательных ключевых этапа: предобработку данных, извлечение биометрических признаков и восстановление криптографического ключа. Поскольку в данной работе в качестве биометрической характеристики используется радужная оболочка глаза, исходные данные представляют собой изображения. Для изображений этап пре-

добработки, как правило, сводится к выделению целевой области (сегментации) и приведению её к стандартному виду (нормализации), что позволяет устранить вариативность, вызванную условиями съёмки, и подготовить данные к извлечению признаков.

Сегментация представляет собой классификацию пикселей исходного изображения на два класса: «радужная оболочка» и «фон» (склера, зрачок, веки, ресницы, блики). Результатом является бинарная маска, указывающая точное расположение радужки. В рамках нормализации выделенная область радужной оболочки приводится к стандартному виду для стабильной работы последующих этапов.

Качество решения каждой из задач определяет итоговые значения метрик. Для объективного сравнительного анализа различных подходов в научном сообществе используется устоявшийся набор стандартизированных показателей.

Точность сегментации оценивается с помощью F1-меры [4] и коэффициента Жаккара (Intersection over Union – IoU) [4], которые вычисляются как

$$\mathrm{F1} = \frac{2 \times \mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FP} + \mathrm{FN}}; \quad (1)$$

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}, \quad (2)$$

где TP – количество элементов положительного класса, классифицированных верно, FP – количество элементов отрицательного класса, ошибочно отнесённых к положительному, FN – количество элементов положительного класса, ошибочно отнесённых к отрицательному. Также используются пиксельная точность (Accuracy), точность (Precision) и полнота (Recall)

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}; \quad (3)$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}; \quad (4)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \quad (5)$$

где TN – количество элементов отрицательного класса, классифицированных верно. При оценке сегментации положительным классом считаются пиксели радужной оболочки, отрицательным – пиксели фона.

Качество извлечения признаков оценить напрямую невозможно, только через анализ качества аутентификации, которая характеризуется ошибками первого и второго рода. Метрики FAR (False Acceptance Rate, вероятность ошибки второго рода), FRR (False Rejection Rate, вероятность ошибки первого рода), GAR (Genuine Acceptance Rate, доля истинного принятия) вычисляются следующим образом [5]:

$$\mathrm{FAR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}; \quad (6)$$

$$\mathrm{FRR} = \frac{\mathrm{FN}}{\mathrm{FN} + \mathrm{TP}}; \quad (7)$$

$$\mathrm{GAR} = 1 - \mathrm{FRR}. \quad (8)$$

В контексте аутентификации положительным классом считается множество легитимных пользователей, отрицательным – множество незарегистрированных пользователей.

Обзор методов сегментации радужной оболочки

Сегментация является критически важным этапом, поскольку ошибки выделения области радужной оболочки непосредственно влияют на качество последующего извлечения признаков. Исторически первыми применялись граничные методы, направленные на выделение контуров объектов по резким изменениям интенсивности. Известны три основных подхода этой группы [6]: интегродифференциальный оператор Даугмана, максимизирующий производную от средней интенсивности вдоль кругового контура [7], преобразование Хафа для поиска окружностей и метод, основанный на анализе распределения граничных точек [8, 9]. Согласно приведённому сравнению, оператор Даугмана точнее остальных, но существенно медленнее: преобразование Хафа работает в 2,5 раза быстрее при меньшей точности, а анализ распределения граничных точек – почти в 6 раз быстрее при точности на уровне преобразования Хафа. Общим ограничением этой группы остаётся чувствительность к бликам и перекрытию радужной оболочки веками и ресницами.

Существенно более высокие показатели качества сегментации дают методы на основе свёрточных нейронных сетей (CNN). Архитектура U-Net, изначально предложенная для сегментации биомедицинских изображений [10], была применена к сегментации радужной оболочки [11] и стала основой для специализированных решений: замена обычных свёрток расширенными увеличивает рецептивное поле без потери разрешения [12], плотные соединения между слоями улучшают поток градиентов [13], остаточные блоки в сочетании с модулем ASPP обеспечивают многомасштабный анализ [14], а сокращение числа параметров до 148 тыс. позволяет работать при ограниченных вычислительных ресурсах почти без потери качества [15].

Отдельное направление образуют архитектуры на основе трансформеров и гибридные решения, в которых свёрточные слои выделяют локальные признаки, а трансформерные блоки учитывают глобальные зависимости [16–19]. Достигаемые ими значения F1-меры ниже, чем у лучших свёрточных решений, к тому же число параметров и требования к объёму размеченных данных существенно выше.

Сводное сопоставление классов методов приведено в табл. 1. Прямое сравнение опубликованных значений затруднено тем, что работы выполнены на разных наборах данных и с разной разметкой, поэтому в таблице приведены достигнутые в первоисточниках максимумы, а сопоставление проводится по классам, а не по отдельным моделям. Из него следует, что для рассматриваемой задачи предпочтительны U-Net-подобные свёрточные архитектуры: они дают наивысшие значения F1-меры при числе параметров, на порядок меньшем, чем у трансформеров, и не требуют столь большого объёма размеченных данных. Именно эта группа исследована далее.

Таблица 1

Сопоставление классов методов сегментации и извлечения признаков

Класс методов	Этап	Преимущества	Ограничения	Результативность
Граничные	Сегментация	Не требуют обучения, низкая вычислительная сложность	Чувствительность к бликам и перекрытиям	Accuracy до 0,90 [6]
Свёрточные (U-Net и модификации)	Сегментация	Наивысшая точность, автоматическое выделение признаков	Требуют размеченных данных	F1 до 0,9914 [14]
Трансформерные и гибридные	Сегментация	Учёт глобальных зависимостей	Большое число параметров, высокие требования к объёму данных	Accuracy до 0,9958 [16], F1 до 0,9548 [17–19]
Статистические	Признаки	Простота реализации	Неустойчивость к изменению освещения	Зависит от классификатора
Спектральные	Признаки	Компактный бинарный код	Жёстко заданное преобразование, нет адаптации к данным	Зависит от классификатора
Нейросетевые	Признаки	Адаптация к набору данных, возможность дообучения	Требуют обучения и размеченных данных	Зависит от классификатора

Методы нормализации и извлечения признаков

После сегментации выполняется преобразование выделенной области, компенсирующее изменение размера зрачка и перекрытие радужной оболочки веками. Классическим подходом является нормализация по методу Даугмана – отображение области из декартовой системы координат в псевдополярную [7, 20]. Альтернатива состоит в кадрировании изображения по границам маски с приведением к фиксированному размеру. Выбор в пользу кадрирования сделан по результатам [21]: при прочих равных условиях нормализация по методу Даугмана дала вероятность ошибки первого рода 0,337 против 0,0714 при кадрировании. Причина в том, что кадрирование сохраняет исходные текстурные соотношения, тогда как переход в псевдополярные координаты требует интерполяции и вносит дополнительное сглаживание.

Извлечение признаков определяет дискриминативную способность системы; исторически сложились статистический, спектральный и нейросетевой подходы. Статистические методы работают с распределением интенсивностей пикселей: анализом гистограмм и моментами изображения [22]; они просты в реализации, но плохо устойчивы к изменению освещения. Спектральные методы анализируют изображение в частотной области: фильтры Габора, положенные в основу алгоритма IrisCode [7], и дискретное косинусное преобразование [23]. Общим свойством обеих групп является жёстко заданное преобразование, не адаптирующееся к конкретному набору данных.

Нейросетевые методы такой адаптацией обладают. Предобученные свёрточные сети в роли моделей для извлечения признаков показали высокую эффективность [24], гибридные подходы дополняют их предварительной фильтрацией круговыми фильтрами Габора [25], отдельное направление образуют сиамские сети, обучающиеся на парах изображений [26, 27].

Для настоящей работы выбраны предобученные свёрточные сети. Сиамские сети, несмотря на успешность в задачах идентификации, для высоконадёжной аутентификации не подходят принципиально: их обучение сближает представления похожих изображений, а не привязывает представление к заданному эталонному выходу, поэтому формируемый код оказывается производным от биометрических данных, что противоречит требованию ГОСТ Р 52633.0–2006 [3] о недопустимости формирования ключа непосредственно из биометрии. Также не рассматривались сквозные автоэнкодеры и трансформерные преобразователи: они не допускают послойного контролируемого обучения, предусмотренного ГОСТ Р 52633.5–2011 [28].

Методы аутентификации по радужной оболочке глаза

Завершающий этап аутентификации может быть реализован классификацией признаков методами машинного обучения (случайный лес, *k*-ближайших соседей, метод опорных векторов и др.) [20] либо восстановлением криптографического ключа с последующей проверкой: по хэшу, что требует точного восстановления каждого бита, или путём сравнения по расстоянию Хэмминга [21]. Первый подход прост, но не позволяет получить криптографический ключ и не является высоконадёжным. Второй включает как нечёткие методы и коды коррекции ошибок, так и методы, которые регламентируются линейкой стандартов ГОСТ Р 52633: ключ должен иметь энтропию не менее 128 бит, не храниться в открытом виде и формироваться с использованием сертифицированных генераторов случайных чисел [3].

Стандарты допускают применение нейросетевых преобразователей «биометрия–код», в частности, многослойных персептронов [28].

Нечёткие методы и коды коррекции ошибок позволяют компенсировать вариабельность биометрических данных [29]. Однако такие методы требуют хранения вспомогательной информации (helper data), что снижает общую безопасность системы.

Схемы, восстанавливающие ключ точно, для радужной оболочки реализованы в ряде работ; их сопоставление с предлагаемым решением приведено в табл. 2. В [30] ключ длиной до 252 бит связывается с отменяемым шаблоном радужки по схеме нечёткого обязательства, что при нулевой оценке вероятности ложного допуска даёт вероятность ошибки первого

рода 0,07; вспомогательные данные при этом хранятся на смарт-карте. В [31] по изображению радужки и секретному ключу формируется открытый ключ, допускающий незащищённое хранение, при объёме восстанавливаемого секрета 65 бит и нулевой вероятности восстановления секрета по изображению другого человека. В [32] нечёткое обязательство дополнено кодами с малой плотностью проверок на чётность и отбором наиболее устойчивых битов шаблона. Наиболее близкой по постановке является работа [33], в которой нечёткий экстрактор обеспечивает 105 бит энтропии ключа при доле истинного принятия 0,92.

Общим для перечисленных схем является то, что точность восстановления ключа достигается за счёт вспомогательных данных, сохраняемых или публикуемых при регистрации, и внешней коррекции ошибок. Таким образом, задача побитового восстановления ключа без вспомогательных данных, при котором проверка сводится к сравнению хэшей, для радужной оболочки остаётся нерешённой; именно эта задача рассматривается в настоящей работе.

В исследовании [21] была продемонстрирована эффективность системы, совмещающей свёрточную сеть для извлечения признаков и MLP для восстановления ключа в рамках одной модели. При нормализации кадрированием и пороге сравнения по расстоянию Хэмминга, равном 32 из 256 бит, система достигла вероятности ошибки первого рода 0,0714 при нулевой оценке вероятности ложного допуска. Однако такой подход имеет два существенных ограничения: во-первых, внутриклассовая вариативность признаков не позволяет добиться полного побитового восстановления ключей; во-вторых, аутентификация по порогу не позволяет использовать криптографическое хэширование и требует хранения ключей в открытом виде, что снижает безопасность.

В настоящей работе этот алгоритм взят за основу и модифицируется.

Предлагаемый алгоритм высоконадёжной аутентификации

В рамках данного исследования предложен и экспериментально оценён нейросетевой алгоритм, при проектировании которого исследована эффективность различных подходов к сегментации, нормализации, а также методов снижения внутриклассовой вариативности на этапах извлечения признаков и восстановления ключа. Тестирование проводилось на наборе данных CASIA-Iris-Thousand, который содержит 20000 изображений радужной оболочки от 1000 субъектов. Левые и правые глаза рассматривались как независимые биометрические объекты, что позволило сформировать выборку из 2000 классов по 10 образцов в каждом.

Общая структура алгоритма приведена на рис. 1.

Таблица 2

Сравнение решений по восстановлению криптографического ключа по радужной оболочке глаза*

Работа	Длина ключа, бит	Вспомогательные данные	Критерий принятия решения	Вероятность ошибки I рода	Вероятность ошибки II рода
[21]	256	Не требуются	Порог расстояния Хэмминга (32 бита)	0,071	0
[30]	до 252	Требуются	Точное совпадение после коррекции	0,070	0
[31]	65	Требуются	Точное совпадение после коррекции	0,104	0
[32]	Не приводится	Требуются	Точное совпадение после коррекции	Не приводится	Не приводится
[33]	105 (энтропия)	Требуются	Точное совпадение после коррекции	0,080	Не приводится
Настоящая работа	512	Не требуются	Точное побитовое совпадение	0,107	0

* Работы выполнены на разных наборах данных и по разным протоколам, поэтому сопоставление вероятностей ошибок носит ориентировочный характер. Строго сопоставимы столбцы «Вспомогательные данные» и «Критерий принятия решения». Для настоящей работы приведены показатели экземпляров, допущенных к эксплуатации; процедура допуска описана далее.

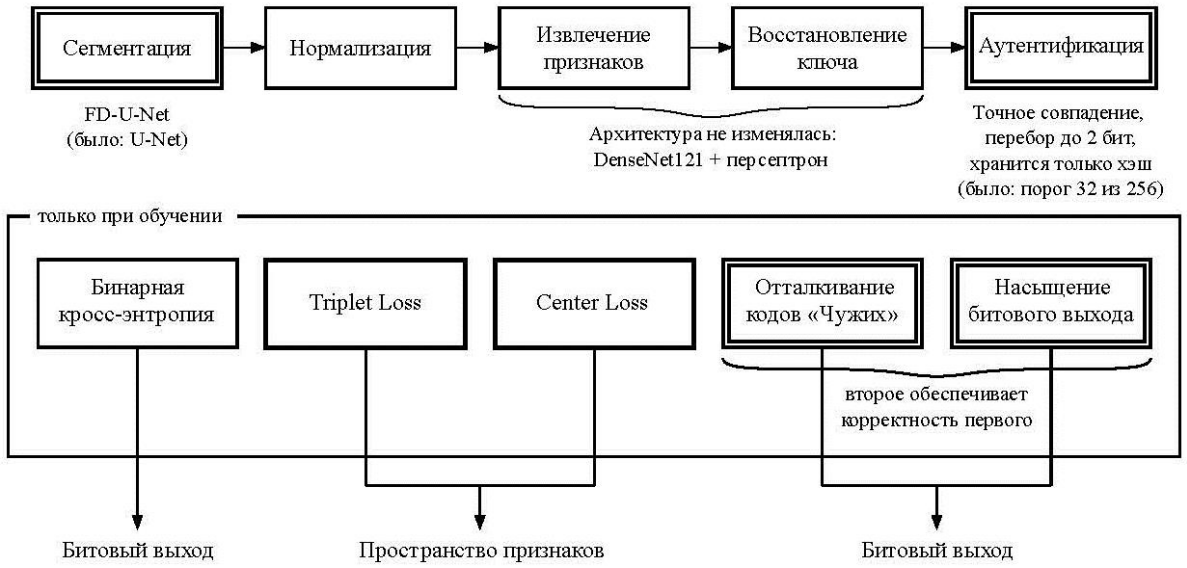


Рис. 1. Схема предлагаемого алгоритма. Двойной рамкой выделены этапы, изменённые по сравнению с решением, принятым за основу [21]; в нижней части приведены слагаемые функции потерь, используемые только при обучении

За основу принято решение [21]; изменения затрагивают выбор архитектуры сегментации, состав функции потерь при обучении и правило принятия решения при проверке, тогда как архитектура извлечения признаков и восстановления ключа оставлена без изменений. Слагаемые функции потерь, показанные в нижней части рисунка, используются только при обучении и в процедуре аутентификации не участвуют.

Выборка разделена по субъектам на пять непересекающихся групп. Первые 50 субъектов отведены под обучение модели сегментации и исключены из последующих этапов. Ещё 50 субъектов зарегистрированы в качестве легитимных пользователей – «Своих». Оставшиеся субъекты отнесены к «Чужим» и разделены на три группы: 250 субъектов используются как отрицательные примеры при обучении, 250 субъектов образуют группу подбора порога, по которой выбирается порог бинаризации и отбирается состояние модели, а 400 субъектов составляют тестовую выборку, по которой подтверждается допуск экземпляра к эксплуатации и вычисляются приводимые далее показатели. Тестовая выборка не участвует ни в обучении, ни в отборе. Состав групп приведён в табл. 3.

Изображения «Своих» разделялись на обучающую и тестовую части в отношении 8 к 2 для каждого пользователя, что обеспечивает одинаковое число отложенных снимков у всех зарегистрированных пользователей; пересечение между частями отсутствует. Изображения «Чужих» по снимкам не разделялись – каждый глаз целиком относился к одной из трёх групп. Аугментация данных применялась только при обучении модели сегментации.

Для сегментации были исследованы четыре архитектуры: U-Net с энкодером ResNet18, IRU-Net, FD-U-Net и Lightweight U-Net. Использовались не точные копии моделей из рассмотренных исследований, а решения, построенные на лежащих в их основе идеях с унифицированными числом базовых фильтров и глубиной.

Такая унификация необходима для сопоставимости: приводимые в первоисточниках значения F1-меры получены на других наборах данных и с другой разметкой, поэтому прямое сравнение с ними некорректно, тогда как задачей табл. 4 является сравнение архитектурных решений между собой при одинаковых условиях обучения.

Таблица 3

Разбиение выборки

Группа	Субъектов	Классов	Изображений
Обучение сегментации	50	100	1 000
«Свои»	50	100	1 000
«Чужие», обучение	250	500	5 000
«Чужие», подбор порога	250	500	5 000
«Чужие», тестирование	400	800	8 000
Итого	1 000	2 000	20 000

Все модели обучались на изображениях размером 640 на 480 пикселей.

Сравнительный анализ (см. табл. 4) показал, что наилучшие значения F1-меры и коэффициента Жаккара получены у модели FD-U-Net (F1 = 0,9712, IoU = 0,9442). Различие с моделью U-Net (ResNet18) не превышает 0,0001 и лежит в пределах воспроизводимости обучения, поэтому при выборе архитектуры учитывалось также число параметров: FD-U-Net достигает того же качества при 10,1 млн параметров против 14,3 млн у U-Net (ResNet18). Данная модель использована в итоговом алгоритме.

После выделения области радужной оболочки выполнялась нормализация: использовался метод кадрирования, который, как показано в [21], обеспечивает более высокое качество последующей аутентификации по сравнению с классической нормализацией по методу Даугмана. Кадрирование представляет собой обрезку радужной оболочки по её границам (границам маски сегментации) и приведение полученного изображения к фиксированному размеру, которое в рамках данной работы составляет 256 на 256 пикселей.

Далее нормализованное изображение поступает на этап извлечения признаков и восстановления ключа. Извлечение признаков производилось предобученной свёрточной сетью DenseNet121, восстановление криптографического ключа – многослойным персептроном (MLP). Данные модели были выбраны, так как удовлетворяют требованиям стандартов ГОСТ Р 52633.0-2006 и ГОСТ Р 52633.5-2011. Модели были объединены для совместного обучения.

Полная архитектура представлена в табл. 5.

Функция потерь строится последовательно: базовую составляющую, отвечающую за восстановление ключа, образует бинарная кросс-энтропия; к ней добавляются метрические слагаемые, формирующие признаковое пространство, и два слагаемых, действующих непосредственно в битовом выходе модели.

Для формирования максимально строгого пространства признаков, необходимого для качественного восстановления ключа, использовалась комбинированная функция потерь, включающая Triplet Loss и Center Loss.

Triplet Loss направлена на построение пространства, в котором расстояние между объектами одного класса меньше, чем между объектами разных классов, на заданную величину отступа α (margin).

Обучение происходит на тройках изображений: опорное изображение, пример того же класса и пример другого класса.

Таблица 4

Результаты сегментации с помощью различных моделей

Модель	Количество параметров	Accuracy	Precision	Recall	F1-мера	IoU
U-Net (ResNet18)	14 321 937	0,9969	0,9699	0,9726	0,9711	0,9441
FD-U-Net	10 126 241	0,9969	0,9706	0,9719	0,9712	0,9442
IRU-Net	2 772 385	0,9963	0,9611	0,9716	0,9661	0,9347
Lightweight U-Net	147 873	0,9947	0,9426	0,9606	0,9505	0,9072

Таблица 5

Архитектура модели для восстановления ключа

Модуль	Тип слоя / Операция	Входная размерность	Выходная размерность
Feature Extractor	DenseNet121	(3, *H*, *W*)	(1 024, *h*, *w*)
Pooling	AdaptiveAvgPool2d	(1 024, *h*, *w*)	(1 024, 1, 1)
Flatten	Flatten	(1 024, 1, 1)	1 024
Head Layer 1	Linear + BatchNorm + LeakyReLU + Dropout	1 024	768
Head Layer 2	Linear + BatchNorm + LeakyReLU + Dropout	768	512
Output Layer	Linear	512	512

Минимизация данной функции заставляет модель приближать векторы одного пользователя и отдалять векторы чужих пользователей, работая непосредственно с евклидовыми расстояниями [34]:

$$L_{\text{triplet}} = \frac{1}{|\mathbf{A}|} \sum_{a \in \mathbf{A}} \max(0, \|\mathbf{v}_a - \mathbf{v}_{p(a)}\|_2 - \|\mathbf{v}_a - \mathbf{v}_{n(a)}\|_2 + \alpha), \quad (9)$$

где $\mathbf{v}_a$ – вектор признаков опорного изображения, $\mathbf{v}_{p(a)}$ – вектор признаков положительного примера того же класса, $\mathbf{v}_{n(a)}$ – вектор признаков негативного примера, выбираемого из объединения признаков остальных «Своих» в мини-батче и признаков «Чужих», $\mathbf{A}$ – множество валидных якорей, α – параметр отступа, $\|\cdot\|_2$ – евклидова норма.

Параллельно применяется функция потерь Center Loss [35], которая штрафует модель за удаление признаков конкретного образца от центра его класса, минимизируя внутриклассовую вариацию:

$$L_{\text{center}} = \frac{1}{2B} \sum_{i=1}^{B} \|\mathbf{v}_i - \mathbf{c}_{y_i}\|_2^2, \quad (10)$$

где B – количество примеров в текущем мини-батче, $\mathbf{v}_i$ – вектор признаков i-го примера в батче, y_i – метка класса (пользователя) для i-го примера, $\mathbf{c}_{y_i}$ – текущий вектор-центр класса y_i.

Метрические слагаемые действуют в пространстве признаков и не гарантируют, что сам битовый выход не окажется близок к одному из ключей базы: при насыщении функций потерь Triplet Loss и Center Loss восстановленный код «Чужого» может подойти к базе вплотную. Поэтому введены два слагаемых – L_{rep} (отталкивание, repulsion) и L_{bin} (бинаризация, binarization), работающих непосредственно в битовом выходе.

Первое отталкивает коды «Чужих» от базы ключей. Для каждого «Чужого» из обучающей выборки вычисляется мягкое расстояние Хэмминга от вероятностей на выходе модели до каждого ключа базы, берётся минимальное, и, если оно меньше заданного запаса, начисляется штраф, пропорциональный недостающей величине. В отличие от метрических слагаемых, оно действует в той же величине, в которой измеряется минимальное расстояние Хэмминга, поэтому позволяет удерживать его на заданном уровне.

Второе насыщает битовый выход: штрафуется величина $p(1-p)$, где p – вероятность на выходе модели; она максимальна при p = 0,5 и обращается в нуль при значениях, близких к нулю и единице. Это слагаемое необходимо для корректности первого: при ненасыщенных выходах мягкое расстояние Хэмминга расходится с фактическим расстоянием после бинаризации, и код «Чужого» может совпасть с ключом базы, тогда как мягкое расстояние остаётся большим и отталкивание не срабатывает. Дополнительно насыщение делает устойчивее выбор порога бинаризации. Оба слагаемых используются только при обучении: при аутентификации они не применяются и никаких вспомогательных данных не сохраняется.

Общая функция потерь L_{key} для совместного обучения модели корректному восстановлению ключа и формирования детерминированного признакового пространства формируется как взвешенная сумма компонент:

$$L_{\text{key}} = L_{\text{BCEWithLogitsLoss}} + \lambda_{\text{t}} \cdot L_{\text{triplet}} + \lambda_{\text{c}} \cdot L_{\text{center}} + \\ + \lambda_{\text{rep}} \cdot L_{\text{rep}} + \lambda_{\text{bin}} \cdot L_{\text{bin}}, \quad (11)$$

где $L_{\text{BCEWithLogitsLoss}}$ – бинарная кросс-энтропия с логитами, объединяющая в себе сигмоидное преобразование и вычисление бинарной кросс-энтропии в одной численно устойчивой операции. Данная функция потерь напрямую минимизирует ошибку восстановления ключа; λ_{t}, λ_{c}, λ_{rep}, $\lambda_{\text{bin}} > 0$ – гиперпараметры, регулирующие вклад остальных компонент.

Значения весовых коэффициентов определялись с учётом масштаба соответствующих слагаемых. Поскольку векторы признаков нормируются, значения функций потерь Triplet Loss и Center Loss лежат в сопоставимых пределах, что допускает использование коэффициентов одного порядка; их итоговые значения ($\lambda_{\text{t}} = 1{,}0$, $\lambda_{\text{c}} = 0{,}5$, $\lambda_{\text{rep}} = 0{,}3$, $\lambda_{\text{bin}} = 0{,}1$) подбирались вручную по нескольким пробным обучениям; отдельного автоматического поиска не проводилось.

Анализ эффективности предлагаемого алгоритма

Отбор состояния модели и допуск экземпляра к эксплуатации разнесены по разным группам «Чужих». В ходе обучения сохраняется то состояние, у которого на группе подбора порога отсутствуют ложные допуски, а минимальное расстояние Хэмминга между восстановленными кодами «Чужих» и базой ключей составляет не менее 18 бит. Отобранное состояние допускается к эксплуатации, если на тестовой выборке ложные допуски также отсутствуют, а минимальное расстояние составляет не менее пяти бит; экземпляр, не удовлетворяющий этому условию, отбраковывается, и обучение повторяется с другой инициализацией. Величина запаса задана до серии

обучений как кратно превышающая требуемые пять бит. Достаточность выбранного значения проверена отдельным расчётом без переобучения, в котором при фиксированном состоянии модели порог подбирался заново для разных значений запаса: у обоих допущенных экземпляров требование пяти бит на тестовой выборке выполняется, начиная с запаса 18 бит, тогда как при 16 битах минимальное расстояние у одного из них падает до трёх. Расчёт носит апостериорный характер и на отбор состояний не влиял. Среди допустимых конфигураций предпочтение отдаётся той, у которой доля истинного принятия максимальна, а при различии в пределах разброса между независимыми обучениями – более воспроизводимой.

Переход к комбинированной функции потерь, включающей Triplet Loss и Center Loss, позволил получить состояния, удовлетворяющие введённому критерию. В качестве рабочей принята конфигурация с восемью обучаемыми блоками свёрточной сети; из трёх независимых обучений допуск к эксплуатации получили два. Для допущенных экземпляров доля истинного принятия составила 0,870 и 0,880, вероятность ошибки первого рода – 0,130 и 0,120, а вероятность ошибки второго рода оказалась нулевой: ни одного ложного допуска на 8000 предъявлений «Чужих» из тестовой выборки. Минимальное расстояние Хэмминга между восстановленными кодами «Чужих» и базой ключей составило 10 и 7 бит при медиане 201 и 227 бит; ближе двадцати бит к базе ключей оказались 4 и 3 предъявления из восьми тысяч. Влияние числа обучаемых блоков свёрточной сети приведено в табл. 6.

Качество восстановления ключа монотонно возрастает с увеличением числа обучаемых блоков и выходит на насыщение начиная с восьми блоков, где доля истинного принятия составляет 0,873±0,006 против 0,877±0,050 при десяти и 0,895±0,074 при двенадцати блоках. Различия между этими тремя конфигурациями лежат внутри разброса между независимыми обучениями. Доля обучений, получивших допуск к эксплуатации, при этом не возрастает и составляет два из трёх при восьми и десяти блоках и одно из трёх при двенадцати.

Из протоколов обучения видно, что с увеличением числа размороженных блоков модель точнее подгоняет коды зарегистрированных пользователей под назначенные им ключи, тогда как расстояние от восстановленных кодов «Чужих» до базы ключей при этом не возрастает. Условию отбора на группе подбора порога удовлетворяет подавляющее большинство эпох; количественно это показано далее. Следовательно, различие между конфигурациями проявляется не в доступности допустимых состояний, а в том, сохраняется ли достигнутый при отборе запас на тестовой выборке.

Выбор восьми блоков обоснован тем, что при статистически неразличимых значениях доли истинного принятия алгоритм даёт наименьший разброс между независимыми обучениями: 0,006 против 0,050 при десяти блоках и 0,074 при двенадцати. Тот же порядок сохраняется и для разброса вероятности ошибки первого рода при переборе в окрестности радиуса два бита: 0,003 против 0,044 при десяти блоках и 0,053 при двенадцати. Порог бинаризации при восьми блоках изменяется в более широких пределах: от 0,82 до 0,92, тогда как при десяти блоках он принимает значения от 0,84 до 0,86; на итоговые показатели это не влияет, поскольку порог подбирается для каждого обученного экземпляра отдельно и между обучениями не переносится.

Подтверждение допуска на тестовой выборке оказалось заметно более строгим, чем отбор на группе подбора порога. Условию отбора с запасом 18 бит удовлетворяло от 76 до 98% эпох в каждом обучении (при восьми и более обучаемых блоках – от 95 до 98%), тогда как требование пяти бит на тестовой выборке сохранилось в девяти обучениях из двадцати одного. Нарушение вызывается небольшим числом отдельных предъявлений: в обучениях, не получивших допуска, условию не удовлетворяли от одного до шести снимков, принадлежавших одному-четырём субъектам, при том, что медиана расстояний до базы ключей оставалась в пределах от 200 до 220 бит. Таким образом, распределение расстояний в целом не ухудшается, нарушается его левый край.

Увеличение запаса при подборе порога положения не исправляет. У экземпляров, получивших допуск, требование пяти бит выполняется начиная с запаса 10 и 18 бит, а с дальнейшим ростом запаса минимальное расстояние на тестовой выборке монотонно увеличивается и при запасе 40 бит достигает 32 и 18 бит.

Таблица 6

Влияние числа обучаемых блоков свёрточной нейронной сети (100 «Своих», по три независимых обучения на конфигурацию)*

Обучаемых блоков	Доля истинного принятия	Вероятность ошибки первого рода	Обучений, получивших допуск	Минимальное расстояние Хэмминга по обучениям, бит
0	0,062 ± 0,010	0,938 ± 0,010	2 из 3	2, 11, 6
2	0,275 ± 0,048	0,725 ± 0,048	0 из 3	0, 0, 1
4	0,627 ± 0,066	0,373 ± 0,066	1 из 3	1, 13, 0
6	0,775 ± 0,041	0,225 ± 0,041	1 из 3	1, 0, 13
8	0,873 ± 0,006	0,127 ± 0,006	2 из 3	10, 0, 7
10	0,877 ± 0,050	0,123 ± 0,050	2 из 3	10, 0, 7
12	0,895 ± 0,074	0,105 ± 0,074	1 из 3	2, 22, 0

*Отбор на группе подбора порога прошли все обучения; в четвёртом столбце указано, в скольких из них требование минимального расстояния сохранилось на тестовой выборке. У допущенных экземпляров ложных допусков не зафиксировано.

У экземпляра, не получившего допуска, оно остаётся нулевым при любом запасе в диапазоне от 4 до 40 бит: код одного из «Чужих» совпадает с ключом базы при всех значениях порога бинаризации, и выбором порога это совпадение не устраняется. Следовательно, запас улучшает калибровку исправного экземпляра, но не заменяет проверку на отдельной группе, а объём самой группы влияет на результат самостоятельно: минимум по выборке не возрастает с ростом её объёма, поэтому оценка по 5000 предъявлений не может быть строже оценки по 8000. Определение необходимого объёма группы, в том числе за счёт синтетического порождения изображений «Чужих», составляет предмет дальнейшей работы.

Снижение вероятности ложного отказа за счёт перебора в окрестности Хэмминга

Несмотря на нулевую оценку вероятности ложного допуска, в рассматриваемой системе сохраняется проблема ложных отказов, обусловленная требованием точного побитового совпадения восстановленного криптографического ключа с эталонным. Даже при высокой разделимости признакового пространства отдельные биты ключей восстанавливаются с ошибками, что приводит к отклонению легитимного пользователя при строгом критерии полного совпадения.

С учётом того, что ошибки восстановления ключа для одного и того же пользователя, как правило, носят локальный характер и затрагивают ограниченное число позиций, был рассмотрен подход, основанный на проверке принадлежности восстановленного ключа окрестности в пространстве Хэмминга. В рамках данного подхода при аутентификации анализируется не только исходный предсказанный ключ, но и множество ключей, находящихся на расстоянии не более k по метрике Хэмминга. Таким образом, задача проверки сводится к поиску совпадения в ограниченной области вокруг восстановленного вектора бинарного кода.

Размер этой области определяется параметром k. При $k=1$ выполняется проверка всех вариантов, отличающихся ровно в одном бите, что соответствует 512 дополнительным кандидатам для ключа длиной 512 бит. При $k=2$ число рассматриваемых комбинаций возрастает до $1{,}3\cdot10^5$, что существенно увеличивает вычислительную нагрузку, но остаётся допустимым при ограниченном числе попыток аутентификации.

Выбор допустимого радиуса поиска определяется структурой межклассовых расстояний в пространстве ключей. В проведённых экспериментах минимальное расстояние между восстановленными кодами «Чужих» и базой ключей у допущенных к эксплуатации экземпляров составило 7 и 10 бит. В этих условиях увеличение радиуса до $k=1$ и $k=2$ позволяет компенсировать наиболее вероятные ошибки восстановления, не приводя к заметному ухудшению разделимости ключей. Наименьшее из наблюдавшихся значений формально допускает радиус до шести бит, однако число проверяемых кандидатов растёт комбинаторно и при трёх битах превышает $2{,}2\cdot10^7$, поэтому радиус ограничен двумя битами.

Перебор в окрестности Хэмминга прямо предусмотрен ГОСТ Р 52633.0–2006 [3]: п. 5.7 допускает оснащение средства высоконадёжной биометрической аутентификации системой перебора возможных состояний наиболее нестабильных бит выходного ключа и разрешает перебор в пределах 7% бит ключа. Использованный в работе радиус два бита при длине ключа 512 бит соответствует 0,39% бит, т.е. лежит внутри допустимого предела с большим запасом. При этом предлагаемая схема не формирует и не хранит сведений о положении нестабильных разрядов: перебор ведётся по всей окрестности, поэтому требование стандарта об уничтожении такой информации к ней не относится. В традиционных системах подобные задачи решаются либо введением порога на расстояние Хэмминга, либо применением кодов коррекции ошибок и схем нечёткой экстракции признаков [21, 29]. В отличие от этих методов, в данной работе корректировка ошибок выполняется на этапе проверки путём локального поиска в пространстве бинарных ключей без ослабления требования полного совпадения ключа, что допускает хэширование, и без использования дополнительной вспомогательной информации.

Экспериментальные результаты показывают, что применение данного подхода позволяет снизить вероятность ложного отказа. Для рабочей конфигурации при разрешении одной битовой ошибки среднее значение вероятности ошибки первого рода уменьшается с 0,125 до 0,115, а при увеличении радиуса поиска до двух бит – до 0,107. Ложных допусков при этом не появляется: радиус поиска остаётся строго меньше минимального расстояния от восстановленных кодов «Чужих» до базы ключей, составляющего у допущенных экземпляров 7 и 10 бит.

Распределение расстояний Хэмминга между восстановленными кодами «Своих» и базой ключей носит двухмодальный характер: подавляющая часть снимков даёт точное совпадение либо отличается на один-два бита, тогда как оставшиеся отклонения составляют десятки и сотни бит. Соответственно, перебор в окрестности Хэмминга устраняет примерно седьмую часть отказов (те из них, которые вызваны единичными и парными искажениями битов), но не влияет на отказы второй группы, которые не могут быть компенсированы никаким практически выполнимым перебором. Дальнейшее снижение вероятности ложного отказа связано не с увеличением радиуса поиска, а с контролем качества предъявляемого изображения.

Дополнительно рассмотрен режим, допускающий вторую попытку аутентификации. ГОСТ Р 52633.0–2006 [3] (п. 5.6) допускает предоставление пользователю нескольких попыток аутентификации и указывает, что их число может быть сопоставимо с числом примеров образа «Свой», на которых обучалась система. Предположение о независимости попыток, однако, для биометрической аутентификации не

выполняется: отказы одного и того же пользователя порождаются устойчивыми свойствами его изображений, а не независимым шумом. Поэтому вероятность отказа при двух попытках не вычислялась как произведение вероятностей, а измерялась непосредственно.

Для измерения использовалось разбиение, при котором у каждого зарегистрированного пользователя отложены ровно два снимка, не участвовавших в обучении; таким образом, каждая пара отложенных снимков представляет собой реализацию режима двух попыток. Для рабочей конфигурации отказ по обеим попыткам наблюдался у 1 пользователя из 100 при требовании точного совпадения ключа в каждом из двух допущенных обучений; при переборе в окрестности радиуса два бита он сохранился лишь в одном из них. Единичное событие не позволяет судить о зависимости попыток по отдельному обучению, поэтому она оценивалась по совокупности обучений в области насыщения. На 900 парах, полученных в девяти обучениях при восьми, десяти и двенадцати блоках, двойной отказ наблюдался 25 раз при 13,8, ожидаемых в предположении независимости. Значимость проверена точным перестановочным критерием: в каждом обучении число отказавших снимков фиксировалось, отказы случайным образом распределялись по снимкам, и доля перестановок, дающих 25 и более двойных отказов, оказалась меньше 0,001. Условная вероятность отказа второй попытки при отказе первой составила 0,235 против безусловной 0,118 по этим же девяти обучениям. Следовательно, оценка по произведению вероятностей занижает вероятность двойного отказа примерно в 1,8 раза и в качестве расчётной приниматься не может.

Влияние повторной попытки на вероятность ложного допуска оценивалось сверху. По неравенству Буля вероятность ложного допуска хотя бы в одной из k попыток не превышает произведения k на вероятность ложного допуска при одной попытке; данная оценка справедлива при любой зависимости между попытками и не требует дополнительных предположений. У допущенных к эксплуатации экземпляров ложных допусков на тестовой выборке не зафиксировано, и вторая попытка этого вывода не меняет: она удваивает число предъявлений, но не сокращает расстояние от восстановленных кодов «Чужих» до базы ключей, минимум которого составляет 7 и 10 бит при радиусе поиска два бита.

Заключение

В ходе выполнения работы модифицирован и экспериментально исследован нейросетевой алгоритм высоконадёжной биометрической аутентификации по радужной оболочке глаза. Основные результаты работы могут быть резюмированы в следующих положениях:

1. Проведён анализ современных методов сегментации, извлечения признаков и аутентификации по радужной оболочке глаза. Установлено, что нейросетевые подходы обеспечивают существенно более высокую точность выделения радужной оболочки по сравнению с граничными методами, а метрические функции потерь позволяют формировать признаковое пространство, пригодное для последующего восстановления криптографического ключа. Обосновано также, что сиамские архитектуры и сквозные преобразователи, успешные в задачах идентификации, для высоконадёжной аутентификации неприменимы по требованиям ГОСТ Р 52633.

2. Предложен и реализован алгоритм, включающий сегментацию с помощью сети FD-U-Net, нормализацию кадрированием, извлечение признаков сетью DenseNet121 и восстановление криптографического ключа многослойным персептроном. Выбор FD-U-Net обоснован наилучшим значением F1-меры (0,9712) среди исследованных архитектур при числе параметров меньшем, чем у ближайшей по качеству модели.

3. Экспериментально подтверждена эффективность комбинированной функции потерь, включающей Triplet Loss и Center Loss, в сочетании с двумя дополнительными членами, действующими в битовом выходе модели. Для экземпляров, допущенных к эксплуатации, аутентификация по точному побитовому совпадению 512-битного криптографического ключа обеспечена при нулевой вероятности ложного допуска (ни одного ложного допуска на 8000 предъявлений «Чужих» из тестовой выборки, минимальное расстояние до базы ключей 10 и 7 бит) и вероятности ошибки первого рода 0,130 и 0,120, снижаемой до 0,110 и 0,105 перебором в окрестности радиуса два бита. Измеренная вероятность отказа при двух попытках составила 0,010 при требовании точного совпадения ключа и 0,005 при переборе в окрестности радиуса два бита. Судить о зависимости попыток по одному обучению это число не позволяет; по совокупности обучений двойной отказ происходит примерно в 1,8 раза чаще, чем предсказывает предположение о независимости, поэтому расчёт по произведению вероятностей для данной задачи неприменим. Установлено также, что запас, обеспечиваемый на группе подбора порога из 250 субъектов, переносится на тестовую выборку из 400 субъектов лишь частично, что определяет объём этой группы как самостоятельный параметр системы; влияние объёма группы «Чужих», используемой при обучении, предстоит оценить отдельно.

Практическая значимость работы заключается в том, что предложенный алгоритм позволяет отказаться от хранения криптографических ключей и биометрических эталонов в открытом виде: проверка выполняется по хэшу точно восстановленного ключа. Ключ формируется криптографическим генератором случайных чисел, не выводится из биометрических данных и имеет длину 512 бит при требовании энтропии не менее 128 бит; преобразователь «биометрия-код» реализован многослойным персептроном, допускающим послойное обучение.

Предложенные решения могут быть использованы при разработке систем аутентификации для критически важных информационных ресурсов.

Литература

1. Information security methods – modern research directions / A. Shelupanov, O. Evsyutin, A. Konev, E. Kostyuchenko, D. Kruchinin, D. Nikiforov // Symmetry. – 2019. – Vol. 11, No. 2. – P. 150.

2. Nguyen K. Deep learning for iris recognition: a survey / K. Nguyen, H. Proença, F. Alonso-Fernandez // ACM Computing Surveys. – 2024. – Vol. 56, No. 9. – P. 1–35.

3. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации. – М.: Стандартинформ, 2020. – 19 с.

4. Müller D. Towards a guideline for evaluation metrics in medical image segmentation / D. Müller, I. Soto-Rey, F. Kramer // BMC Research Notes. – 2022. – Vol. 15, No. 1. – P. 210.

5. Mayya A. Iris and palmprint decision fusion to enhance human recognition / A. Mayya, M. Saii // International Journal of Computer Science Trends and Technology. – 2017. – Vol. 5. – P. 42–46.

6. Семёнов М.С. Сравнение методов сегментации изображения радужной оболочки глаза / М.С. Семёнов, Е.В. Мясников // Информационные технологии и нанотехнологии (ИТНТ-2018): сб. трудов IV Междунар. конф. и молодёжной школы, Самара, 24–27 апреля 2018 г. – Самара: Самар. нац. исслед. ун-т им. акад. С.П. Королева, 2018. – С. 814–821.

7. Daugman J.G. High confidence visual recognition of persons by a test of statistical independence // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1993. – Vol. 15, No. 11. – P. 1148–1161.

8. Чигринский В.В. Быстрый алгоритм поиска границ зрачка и радужной оболочки глаза / В.В. Чигринский, Ю.С. Ефимов, И.А. Матвеев // Машинное обучение и анализ данных. – 2016. – Т. 2, № 2. – С. 159–172.

9. Ганеева Ю.Х. Сегментация радужной оболочки глаза на изображении с помощью свёрточной нейронной сети архитектуры U-Net / Ю.Х. Ганеева, Е.В. Мясников // Информационные технологии и нанотехнологии (ИТНТ-2020): сб. трудов VI Междунар. конф. и молодёжной школы, Самара, 26–29 мая 2020 г. – Самара: Самар. нац. исслед. ун-т им. акад. С.П. Королева, 2020. – Т. 2. – С. 57–63.

10. Ronneberger O. U-Net: convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: proc. of the 18th Int. Conf., Munich, Germany, October 5–9, 2015. – Cham: Springer, 2015. – P. 234–241.

11. Lozej J. End-to-end iris segmentation using U-Net / J. Lozej, B. Meden, V. Struc, P. Peer // 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI), July 18–20, 2018. – San Carlos, Costa Rica: IEEE, 2018. – P. 1–6. DOI: 10.1109/IWOBI.2018.8464213.

12. A robust iris segmentation scheme based on improved U-Net / W. Zhang, X. Lu, Y. Gu, Y. Liu, X. Meng, J. Li // IEEE Access. – 2019. – Vol. 7. – P. 85082–85089.

13. Wu X. Study on iris segmentation algorithm based on dense U-Net / X. Wu, L. Zhao // IEEE Access. – 2019. – Vol. 7. – P. 123959–123968.

14. Robust iris segmentation algorithm in non-cooperative environments using interleaved residual U-Net / Y.H. Li, W.R. Putri, M.S. Aslam, C.C. Chang // Sensors. – 2021. – Vol. 21, No. 4. – P. 1434.

15. Efficient and robust eye images iris segmentation using a lightweight U-Net convolutional network / C. Miron, A. Pasarica, V. Manta, R. Timofte // Multimedia Tools and Applications. – 2022. – Vol. 81, No. 11. – P. 14961–14977.

16. Farmanifard P. Iris-SAM: iris segmentation using a foundation model / P. Farmanifard, A. Ross // Pattern Recognition and Artificial Intelligence: proc. of the 4th Int. Conf. ICPRAI 2024, Jeju Island, South Korea, July 3–6, 2024. – Singapore: Springer Nature Singapore, 2024. – P. 394–409.

17. Uncertainty-aware bilateral transformer for accurate and reliable iris segmentation / J. Wei, X. Gao, Y. Wang, R. He, Z. Sun // IEEE Transactions on Information Forensics and Security. – 2025. – Vol. 20. – P. 5919–5933.

18. Jiang J. SAM-Iris: a SAM-based iris segmentation algorithm / J. Jiang, Q. Zhang, C. Wang // Electronics. – 2025. – Vol. 14, No. 2. – P. 246.

19. IrisST-Net for iris segmentation and contour parameters extraction / Y. Liu, W. Shen, D. Wu, J. Shao // Applied Intelligence. – 2023. – Vol. 53, No. 9. – P. 11267–11281.

20. Ganeeva Y.K. Identifying persons from iris images using neural networks for image segmentation and feature extraction / Y.K. Ganeeva, E.V. Myasnikov // Computer Optics. – 2022. – Vol. 46, No. 2. – P. 308–316.

21. Штука С.Е. Разработка нейросетевого алгоритма высоконадёжной биометрической аутентификации на основе радужной оболочки глаза / С.Е. Штука, А.Е. Боршевников // Региональная информатика и информационная безопасность: сб. трудов XIV Санкт-Петербургской межрегион. конф., Санкт-Петербург, 29–31 октября 2025 г. – СПб.: СПОИСУ, 2025. – С. 204–206.

22. Demirel H. Iris recognition system using combined histogram statistics / H. Demirel, G. Anbarjafari // 2008 23rd International Symposium on Computer and Information Sciences (ISCIS), October 27–29, 2008. – Istanbul, Turkey: IEEE, 2008. – P. 1–4.

23. Monro D.M. DCT-based iris recognition / D.M. Monro, S. Rakshit, D. Zhang // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2007. – Vol. 29, No. 4. – P. 586–595.

24. An effective model for the iris regional characteristics and classification using deep learning Alex network / T. Balashanmugam, K. Sengottaiyan, M.S. Kulandairaj, H. Dang // IET Image Processing. – 2023. – Vol. 17, No. 1. – P. 227–238.

25. Iris recognition based on local circular Gabor filters and multi-scale convolution feature fusion network / J. Sun, S. Zhao, Y. Yu, X. Wang, L. Zhou // Multimedia Tools and Applications. – 2022. – Vol. 81, No. 23. – P. 33051–33065.

26. Li Y. A survey on Siamese network: methodologies, applications, and opportunities / Y. Li, C.L.P. Chen, T. Zhang // IEEE Transactions on Artificial Intelligence. – 2022. – Vol. 3, No. 6. – P. 994–1014.

27. Latif S.A. Hybrid CNN-ViT integration into Siamese networks for robust iris biometric verification / S.A. Latif, K.A. Sidek, A.H.A. Hashim // Bulletin of Electrical Engineering and Informatics. – 2025. – Vol. 14, No. 4. – P. 3101–3109.

28. ГОСТ Р 52633.5-2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия–код доступа. – М.: Стандартинформ, 2018. – 16 с.

29. Quantifying privacy and security of biometric fuzzy commitment / X. Zhou, A. Kuijper, R. Veldhuis, C. Busch // 2011 International Joint Conference on Biometrics (IJCB), Washington, DC, USA, October 11–13, 2011. – IEEE, 2011. – P. 1–8.

30. Kausar F. Iris based cancelable biometric cryptosystem for secure healthcare smart card // Egyptian Informatics Journal. – 2021. – Vol. 22, No. 4. – P. 447–453.

31. Matveev I. From iris image to embedded code: system of methods / I. Matveev, I. Safonov // Algorithms. – 2023. – Vol. 16, No. 2. – Art. 87.

32. Lin K.-C. A high-security-level iris cryptosystem based on fuzzy commitment and soft reliability extraction / K.-C. Lin, Y.-M. Chen // IEEE Transactions on Dependable and Secure Computing. – 2024. – Vol. 21, No. 4. – P. 1770–1784.

33. Fuzzy extractors are practical: cryptographic strength key derivation from the iris / A. Shukla, L. Demarest, B. Fuller, S. Ahmad, C. Manicke, A. Russell, S. Chen // Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS '25), Taipei, Taiwan, October 13–17, 2025. – New York: ACM, 2025. – P. 3605–3619.

34. Schroff F. FaceNet: a unified embedding for face recognition and clustering / F. Schroff, D. Kalenichenko, J. Philbin // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7–12, 2015. – Boston, MA, USA: IEEE, 2015. – P. 815–823.

35. A discriminative feature learning approach for deep face recognition / Y. Wen, K. Zhang, Z. Li, Y. Qiao // Computer Vision – ECCV 2016: proc. of the 14th European Conf., Amsterdam, The Netherlands, October 11–14, 2016. – Cham: Springer, 2016. – P. 499–515.

Штука Станислав Евгеньевич
Студент департамента информационной безопасности
Дальневосточного федерального университета (ДВФУ)
п. Аякс, о. Русский, 10, г. Владивосток, Россия, 690922
ORCID: 0009-0007-0659-6166
Тел.: +7-914-712-25-67
Эл. почта: stanislav.shtuka@mail.ru

Боршевников Алексей Евгеньевич
Ст. аналитик АО «Промышленные криптосистемы»
(АО «ПроКСи»)
Отрадная ул., 2Б, стр. 1, г. Москва, Россия, 127273
ORCID: 0000-0002-7383-1820
Тел.: +7-924-131-67-97
Эл. почта: borshevnikov.ae@dvfu.ru

Поступила в редакцию: 29.05.2026.
Принята к публикации: 12.08.2026.

Shtuka S.E., Borshevnikov A.E.

Modification of a neural network algorithm for highly reliable iris-based biometric authentication

Relevance. The growing number of systems processing confidential data calls for authentication mechanisms that remain secure when the stored reference data is compromised. ***Purpose of the study.*** To modify a neural network algorithm for highly reliable iris-based biometric authentication, enabling the cryptographic key recovery with an exact bit-wise match. ***Methods.*** A convolutional neural network is used to extract the iris region, followed by crop-based normalization. Joint training of a convolutional model for feature extraction and a multilayer perceptron is performed using a combined loss function comprising metric terms and two terms acting directly on the bit-wise output of the model. The operating point is selected according to an admissibility criterion, and Hamming-neighborhood search is employed at the verification stage. ***Scientific novelty.*** Unlike known schemes, exact key recovery is achieved without auxiliary data or external error correction, allowing only the hash of the key to be stored. ***Results.*** For 100 enrolled users, the genuine acceptance rate among the instances admitted for operation was 0.875 with a zero false acceptance probability over 8000 impostor presentations. Searching within a radius of two bits reduces the type I error probability to 0.107 without changing the type II error probability. Rejection on both attempts under exact key match was observed for 1 out of 100 users. It is shown that the margin secured on the threshold-selection group transfers to the test set only in part. ***Practical significance.*** The algorithm is applicable in systems where the storage of biometric reference data and cryptographic keys in plaintext is unacceptable.

Keywords: biometric authentication, iris recognition, image segmentation, neural networks, cryptographic key, Triplet Loss, Center Loss, biometric cryptography.
DOI: 10.21293/1818-0442-2026-29-2-148-159

References

1. Shelupanov A., Evsyutin O., Konev A., Kostyuchenko E., Kruchinin D., Nikiforov D. Information security methods – modern research directions. *Symmetry*, 2019, vol. 11, no. 2, art. 150.

2. Nguyen K., Proença H., Alonso-Fernandez F. Deep learning for iris recognition: a survey. *ACM Computing Surveys*, 2024, vol. 56, no. 9, pp. 1–35.

3. GOST R 52633.0-2006. *Zashchita informatsii. Tekhnika zashchity informatsii. Trebovaniya k sredstvam vysokonadezhnoy biometricheskoy autentifikatsii* [Information protection. Information protection technique. Requirements for means of high-reliability biometric authentication]. Moscow, Standartinform Publ., 2020 (in Russ.).

4. Müller D., Soto-Rey I., Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*, 2022, vol. 15, no. 1, art. 210.

5. Mayya A., Saii M. Iris and palmprint decision fusion to enhance human recognition. *International Journal of Computer Science Trends and Technology*, 2017, vol. 5, no. 5, pp. 42–46.

6. Semyonov M.S., Myasnikov E.V. *Sravnenie metodov segmentatsii izobrazheniya raduzhnoy obolochki glaza* [Comparison of iris image segmentation methods]. Informatsionnye tekhnologii i nanotekhnologii (ITNT-2018) [*Information Technologies and Nanotechnologies*]. Proc. of the IV Int. Conf., Samara, 24–27 April 2018, Samara, Samara National Research University named after Academician S.P. Korolev, 2018, pp. 814–821 (in Russ.).

7. Daugman J.G. High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, vol. 15, no. 11, pp. 1148–1161.

8. Chigrinskiy V.V., Efimov Yu.S., Matveev I.A. *Bystryy algoritm poiska granits zrachka i raduzhnoy obolochki glaza* [A fast algorithm for locating pupil and iris boundaries]. *Mashinnoe obuchenie i analiz dannykh*, 2016, vol. 2, no. 2, pp. 159–172 (in Russ.).

9. Ganeeva Yu.Kh., Myasnikov E.V. *Segmentatsiya raduzhnoy obolochki glaza na izobrazhenii s pomoshchyu svertochnoy neyronnoy seti arkhitektury U-Net* [Iris segmentation in an image using a U-Net convolutional neural network]. Informatsionnye tekhnologii i nanotekhnologii (ITNT-2020). [*Information Technologies and Nanotechnologies*] Proc. of the VI Int. Conf., Samara, 26–29 May 2020, Samara, Samara National Research University named after Academician S.P. Korolev, 2020, vol. 2, pp. 57–63 (in Russ.).

10. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation. Proc. of the 18th Int. Conf.: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Munich, Germany, Cham, Springer, 2015, pp. 234–241.

11. Lozej J., Meden B., Struc V., Peer P. End-to-end iris segmentation using U-Net. 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI), San Carlos, Costa Rica, IEEE, 2018, pp. 1–6. DOI: 10.1109/IWOBI.2018.8464213.

12. Zhang W., Lu X., Gu Y., Liu Y., Meng X., Li J. A robust iris segmentation scheme based on improved U-Net. *IEEE Access*, 2019, vol. 7, pp. 85082–85089.

13. Wu X., Zhao L. Study on iris segmentation algorithm based on dense U-Net. *IEEE Access*, 2019, vol. 7, pp. 123959–123968.

14. Li Y.H., Putri W.R., Aslam M.S., Chang C.C. Robust iris segmentation algorithm in non-cooperative environments using interleaved residual U-Net. *Sensors*, 2021, vol. 21, no. 4, p. 1434.

15. Miron C., Pasarica A., Manta V., Timofte R. Efficient and robust eye images iris segmentation using a lightweight U-Net convolutional network. *Multimedia Tools and Applications*, 2022, vol. 81, no. 11, pp. 14961–14977.

16. Farmanifard P., Ross A. Iris-SAM: iris segmentation using a foundation model. Pattern Recognition and Artificial Intelligence. Proc. of the 4th Int. Conf. ICPRAI 2024, Jeju Island, South Korea, Singapore, Springer Nature Singapore, 2024, pp. 394–409.

17. Wei J., Gao X., Wang Y., He R., Sun Z. Uncertainty-aware bilateral transformer for accurate and reliable iris segmentation. I*EEE Transactions on Information Forensics and Security*, 2025, vol. 20, pp. 5919–5933.

18. Jiang J., Zhang Q., Wang C. SAM-Iris: a SAM-based iris segmentation algorithm. *Electronics*, 2025, vol. 14, no. 2, p. 246.

19. Liu Y., Shen W., Wu D., Shao J. IrisST-Net for iris segmentation and contour parameters extraction. *Applied Intelligence*, 2023, vol. 53, no. 9, pp. 11267–11281.

20. Ganeeva Y.K., Myasnikov E.V. Identifying persons from iris images using neural networks for image segmentation and feature extraction. *Computer Optics*, 2022, vol. 46, no. 2, pp. 308–316.

21. Shtuka S.E., Borshevnikov A.E. *Razrabotka neyrosetevogo algoritma vysokonadezhnoy biometricheskoy autentifikatsii na osnove raduzhnoy obolochki glaza* [Development of a neural network algorithm for high-reliability biometric authentication based on the iris]. Regionalnaya informatika i informatsionnaya bezopasnost. [*Regional Informatics and Information Security*] Proc. of the XIV St. Petersburg Interregional Conf., St. Petersburg, 29–31 October 2025, St. Petersburg, SPOISU Publ., 2025, pp. 204–206 (in Russ.).

22. Demirel H., Anbarjafari G. Iris recognition system using combined histogram statistics. 2008 23rd International Symposium on Computer and Information Sciences (ISCIS), Istanbul, Turkey, IEEE, 2008, pp. 1–4.

23. Monro D.M., Rakshit S., Zhang D. DCT-based iris recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, vol. 29, no. 4, pp. 586–595.

24. Balashanmugam T., Sengottaiyan K., Kulandairaj M.S., Dang H. An effective model for the iris regional characteristics and classification using deep learning Alex network. *IET Image Processing*, 2023, vol. 17, no. 1, pp. 227–238.

25. Sun J., Zhao S., Yu Y., Wang X., Zhou L. Iris recognition based on local circular Gabor filters and multi-scale convolution feature fusion network. *Multimedia Tools and Applications*, 2022, vol. 81, no. 23, pp. 33051–33065.

26. Li Y., Chen C.L.P., Zhang T. A survey on Siamese network: methodologies, applications, and opportunities. *IEEE Transactions on Artificial Intelligence*, 2022, vol. 3, no. 6, pp. 994–1014.

27. Latif S.A., Sidek K.A., Hashim A.H.A. Hybrid CNN-ViT integration into Siamese networks for robust iris biometric verification. *Bulletin of Electrical Engineering and Informatics*, 2025, vol. 14, no. 4, pp. 3101–3109.

28. GOST R 52633.5-2011. *Zashchita informatsii. Tekhnika zashchity informatsii. Avtomaticheskoe obuchenie neyrosetevykh preobrazovateley biometriya-kod dostupa* [Information protection. Information protection technique. Automatic training of neural network converters for biometrics-access code]. Moscow, Standartinform Publ., 2018 (in Russ.).

29. Zhou X., Kuijper A., Veldhuis R., Busch C. Quantifying privacy and security of biometric fuzzy commitment. 2011 International Joint Conference on Biometrics (IJCB), Washington, DC, USA, IEEE, 2011, pp. 1–8.

30. Kausar F. Iris based cancelable biometric cryptosystem for secure healthcare smart card. *Egyptian Informatics Journal*, 2021, vol. 22, no. 4, pp. 447–453.

31. Matveev I., Safonov I. From iris image to embedded code: system of methods. *Algorithms*, 2023, vol. 16, no. 2, art. 87.

32. Lin K.-C., Chen Y.-M. A high-security-level iris cryptosystem based on fuzzy commitment and soft reliability extraction. *IEEE Transactions on Dependable and Secure Computing*, 2024, vol. 21, no. 4, pp. 1770–1784.

33. Shukla A., Demarest L., Fuller B., Ahmad S., Manicke C., Russell A., Chen S. Fuzzy extractors are practical: cryptographic strength key derivation from the iris. Proc. of the 2025 ACM SIGSAC Conf. on Computer and Communications Security (CCS '25), Taipei, Taiwan, New York, ACM, 2025, pp. 3605–3619.

34. Schroff F., Kalenichenko D., Philbin J. FaceNet: a unified embedding for face recognition and clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, IEEE, 2015, pp. 815–823.

35. Wen Y., Zhang K., Li Z., Qiao Y. A discriminative feature learning approach for deep face recognition. Computer Vision – ECCV 2016. Proc. of the 14th European Conf., Amsterdam, The Netherlands, Cham, Springer, 2016, pp. 499–515.

Stanislav E. Shtuka
Student, Department of Information Security,
Far Eastern Federal University (FEFU)
10, Ajax, Russky Island, Vladivostok, Russia, 690922
ORCID: 0009-0007-0659-6166
Phone: +7-914-712-25-67
Email: stanislav.shtuka@mail.ru

Aleksey E. Borshevnikov
Senior Analyst,
Industrial Cryptosystems JSC (ProCSy JSC)
2B, bld. 1, Otradnaya st., Moscow, Russia, 127273
ORCID: 0000-0002-7383-1820
Phone: +7-924-131-67-97
Email: borshevnikov.ae@dvfu.ru

Received: 29.05.2026.
Accepted: 12.08.2026.

УДК 004.056
Шифр специальности ВАК: 2.3.6

А.В. Бочков, А.Г. Сабанов, А.А. Шелупанов

Концепция формирования метрик безопасности прикладного ПО для объектов КИИ

Актуальность. Усиление требований нормативно-правовой базы к защите информации объектов критической информационной инфраструктуры приводит к необходимости создания системы количественных оценок безопасности разрабатываемого и используемого прикладного программного обеспечения. ***Цель исследования***. Разработка модернизированного метода количественных оценок безопасности прикладного программного обеспечения на основе методологии системы международных стандартов по качеству программного обеспечения. ***Методы.*** В работе комплексно применены методы системного анализа, формализации, синтеза, сравнения, вычислительного эксперимента и нормативной привязки. ***Научная новизна.*** Разработана концепция введения системы количественных оценок безопасности прикладного программного обеспечения и новой подхарактеристики безопасности программ «уровни безопасности программного обеспечения» как показателя качества разрабатываемых программ. Предложена трехуровневая система формирования уровней безопасности программного обеспечения, опирающаяся на двухконтурную модель агрегирования, в которой заградительные и градуируемые метрики играют математически различную функциональную роль. ***Результаты.*** Применимость предложенной концепции показана на примере расчета метрик безопасности. ***Практическая значимость.*** Предложенная система количественных оценок уточненных метрик безопасности может применяться в проектах создания и модернизации прикладного программного обеспечения объектов критической информационной инфраструктуры.

Ключевые слова: безопасность программного обеспечения, критическая информационная инфраструктура, SQuaRE, метрики безопасности, гармоническое среднее, заградительные метрики, модель угроз, уровни безопасности программного обеспечения.
DOI: 10.21293/1818-0442-2026-29-2-160-176

Введение

Ключевая роль в обеспечении безопасности информационных систем (ИС) различного назначения принадлежит программному обеспечению (ПО). Применительно к ПО ИС объектов *критической информационной инфраструктуры (КИИ)*, согласно ст. 11 Федерального закона от 26.07.2017 № 187-ФЗ и приказу ФСТЭК России от 25 декабря 2017 г. № 239, существенным требованием является обеспечение безопасности ПО в соответствии с основными положениями стандарта [1].

Пункт 3.1 указанного стандарта трактует термин «безопасность ПО» как «программное обеспечение, разработанное в ходе реализации совокупности процессов (мер), направленных на предотвращение появления и устранение недостатков программы». Под недостатками ПО понимается наличие ошибок, известных уязвимостей, нежелательных конструкций, недекларированных возможностей (НДВ), недостаточной защиты от несанкционированного доступа (НСД) и скрытых каналов. При этом подразумевается, что нарушение безопасности ПО может привести к реализации угроз безопасности информации с использованием недостатков ПО, направленных на систему, для которой данное ПО разрабатывается, поскольку само по себе ПО гораздо реже является предметом непосредственной цели злоумышленников. Другими словами, в этой парадигме ПО может являться источником угроз или путем их реализации.

В целях нейтрализации связанных с ПО угроз, согласно [1], ПО должно оставаться безопасным на всем жизненном цикле (ЖЦ) – на этапах проектирования, разработки, внедрения, эксплуатации и вывода из эксплуатации [2].

В стандарте [1] степень безопасности ПО предлагается оценивать по строгости выполнения требований к организации процесса разработки, а также по полноте проведения различных видов анализа кода на всех этапах ЖЦ. Например, наличие или отсутствие уязвимостей в коде программы следует проверять на основе статического и динамического анализа, композиционного анализа, тестирования на проникновение и экспертизы кода, включая анализ самих инструментов тестирования, используемых анализаторами баз уязвимостей и настроек тестирования. Обратим внимание, что даже после выполнения перечисленного комплекса различных видов тестирования никто не подпишется, что ПО гарантированно не содержит уязвимостей, поскольку через непродолжительное время в заимствованном или собственном коде могут быть обнаружены новые уязвимости и не обнаруженные ранее другие недостатки. Поэтому в отчетах по результатам анализа кода часто присутствует сопровождающая фраза, например, «известных на данный момент уязвимостей в данном ПО не обнаружено».

Рассмотренный пример показывает, что о безопасности ПО можно судить по косвенным признакам, в данном случае – по мере соблюдения требований положений стандарта [1] на всем ЖЦ ПО, а именно в затраченных усилиях (приобретение и настройка анализаторов кода, подготовка и проведение тестирования, анализ отчетов, устранение уязвимостей, повторный анализ для появления уверенности в

отсутствии известных уязвимостей и т.д.). Применительно к безопасности ПО для объектов КИИ ФСТЭК России установил требования к процессу разработки и поиску уязвимостей на всех этапах ЖЦ, документации, порядку обновления ПО [3].

В недавно принятом методическом документе ФСТЭК России [4] операторам (владельцам) ИС рекомендуется проведение оценок защищенности систем, при этом предполагается выполнение требований стандарта [1] и проведение в основном качественного анализа безопасности ПО, в первую очередь, на отсутствие уязвимостей. Однако в практических работах все чаще возникает необходимость получения не только качественной, но и количественной оценки безопасности разрабатываемого заказного или приобретаемого ПО, особенно для объектов КИИ [5].

Одним из путей решения указанной задачи является использование подхода хорошо развитой количественной оценки качества ПО в системе стандартов ISO/IEC 25000, более известной как SQuaRE [6–12], поскольку безопасность является одной из характеристик качества ПО. Качеству ПО в указанной системе международных стандартов посвящено более 30 документов, значительная часть которых переведена на русский язык и введена в систему стандартов ГОСТа в виде идентичных (прямой перевод).

Система стандартов ISO/IEC 25000 развивается с 1991 г. по настоящее время и имеет ряд преимуществ, которые привлекают к себе для их применения для оценки качества программного обеспечения прозрачностью, измеримостью, повторяемостью и простотой оценивания метрик, позволяющих в цифровом выражении получать результат оценки качества ПО. Оценка качества ПО в указанной системе международных стандартов проводится с помощью трехуровневой системы: введены 8 основных характеристик качества (в число которых входит безопасность), каждая характеристика на втором уровне определяется рядом подхарактеристик, которые на третьем уровне имеют достаточно просто измеримые метрики, позволяющие получать их численные значения. Именно это свойство системы стандартов ISO/IEC 25000 привлекает для попытки ее использования в оценке безопасности ПО.

Для ПО систем объектов КИИ необходимость использования исключительно безопасного ПО утверждена в приказе ФСТЭК России [3]. Следовательно, для применения стандартов ISO/IEC 25000 к созданию системы количественных метрик безопасности ПО для объектов КИИ требуется коррекция подходов SQuaRE, обусловленная требованиями [3], что является целью данной работы.

Таким образом, в настоящей статье приводится обоснование введения модернизированных подхарактеристик безопасности ПО, учитывающих современное состояние теории защиты информации и определяющих причинно-следственную связь между процессами безопасной разработки ПО и характеристиками его безопасности. С использованием подходов к формированию уровней доверия [13, 14] предлагается введение уровней доверия к безопасности ПО. На основе приведенного ниже анализа научных работ и стандартов, определяющих качество ПО, разработана методика количественной оценки безопасности разрабатываемого ПО как показателя (характеристики) качества ПО. В настоящей статье рассматривается прикладное ПО, поскольку оно наиболее часто разрабатывается для объектов КИИ.

Обзор научных работ по оценке безопасности ПО и анализ недостатков SQuaRE

Отечественные исследования. Из работ отечественных авторов выделим фундаментальные исследования. Вопросы исследования безопасности ПО рассматривались в работах И.Б. Шубинского, в том числе применительно к АСУ ТП железнодорожного транспорта [15–17]. В работе [15] рассмотрены основные модели надежности функционирования ПО с использованием стандарта качества ПО [9]. Также установлено, что для объектов КИИ требования к качеству ПО определяются рисками. В частности, показано, что для критических систем чрезвычайно важно обеспечение целостности и доступности ПО.

В работе [16] показано, что комплексная безопасность ПО в АСУ ТП должна включать в себя качество и функциональную безопасность, отсутствие НДВ и защищенность от НСД. В работе [17] рассмотрены требования к информационной безопасности ПО, которые не должны содержать вредоносного кода, вирусов, НДВ, нарушений целостности ПО.

Значительный вклад в развитие теории безопасного ПО и методов оценки безопасности ПО внес коллектив исследователей под руководством А.С. Маркова. Так, в работе [18] рассмотрены математические модели полноты тестирования, сложности ПО, временные модели роста надежности и отладочные модели. К сожалению, модели оценки безопасности ПО в данной работе не рассматривались. В работе [19] сформулировано определение БПО: безопасным ПО называется ПО, разработанное с использованием совокупности мер, направленных на предотвращение появления и устранение уязвимостей программы. Также показано, что в оценках безопасности ПО много субъективного от экспертов-оценщиков.

В работе [20] отмечено, что наиболее актуальными моделями представления программ с позиции безопасности являются абстрактное синтаксическое дерево и абстрактный синтаксический граф. Методика основывается на поиске программных дефектов в программном коде путем сопоставления фрагментов кода с образцами из базы данных шаблонов (эвристических правил) дефектов безопасности. Разработана метамодель, позволяющая формализовать описание эвристик поиска дефектов на различных уровнях представления кода, что обеспечивает повышение унификации выявления сложных уязвимостей. Показано, что эвристический анализ может служить базой при использовании различных техник аудита БПО.

Оценка безопасности ПО затронута в работах [21, 22], при этом, как правило, использовались или модернизировались актуальные на время написания работ модели качества системы стандартов SQuaRE без акцента на метрики безопасности ПО.

Зарубежные исследования. За рубежом в последние годы развивается подход SSA (Software Security Assurance). Обеспечение гарантий безопасности программного обеспечения – это процесс, направленный на то, чтобы ПО работало с уровнем безопасности, соответствующим потенциальному ущербу, который может быть вызван потерей, неточностью, изменением, недоступностью или неправомерным использованием данных и ресурсов, находящихся под контролем и защитой ПО [23]. В процессе разработки ПО это выполнение требований безопасности, начиная с ранних стадий проектирования, при кодировании и внедрении ПО. Известны работы Кэрола Вуди (Carol Woody), развивающего подход SSA на основе анализа угроз безопасности, связанных с ПО [24].

Одним из продвигаемых на рынке решений SSA является защищенная платформа разработки веб-приложений Рефлектиз [25]. Тем не менее в условиях роста кибератак, согласно отчету State of Web Exposure 2026 [26], реальная защищенность веб-ресурсов снижается. Так, 64% сторонних приложений получают доступ к конфиденциальным данным через Web-доступ по сравнению с 51% в 2024 г. Между тем опрос руководителей служб безопасности выявил серьезный пробел: хотя 81% специалистов по безопасности считают веб-атаки приоритетной угрозой, только 39% из них используют специальные решения для парирования этих угроз.

Другим развивающимся направлением исследований является Dep (Defect Prediction in Software – прогнозирование недостатков ПО). Модели Dep создаются с использованием двух подходов: с использованием измеримых свойств программной системы, называемых программными метриками, и с использованием данных о недостатках ПО из аналогичного программного проекта. Изначально в Dep использовались только статические методы анализа кода, но затем для отслеживания поведения ПО на разных стадиях ЖЦ стали интенсивно применяться методы машинного обучения. Например, некоторые авторы использовали дерево решений (DT) [27] и искусственную нейронную сеть (ANN) [28] для построения моделей Dep с использованием объектно-ориентированных (ОО) метрик [29].

Стратегия Zero Trust основывается на принципе «никогда не доверяй, всегда проверяй» (never trust, always verify), при котором ни один пользователь, устройство, пакет данных или сетевой интерфейс по умолчанию не считаются надёжными, а любой элемент инфраструктуры или учётная запись рассматриваются как потенциально небезопасные и подлежат строгой верификации. Данная концепция была в 2010 г. сформулирована Д. Киндервагом, аналитиком Forrester Research [30]. Этот подход стал весьма актуален и распространен в годы эпидемии ковид-19 в условиях массового перехода на удалённую работу, чтобы сотрудники пользовались корпоративными ресурсами с личных устройств.

Анализ модели качества SQuaRE и её недостатков

Наиболее известным и традиционным подходом является оценка ПО и безопасности как одной из характеристик модели качества согласно стандартам SQuaRE. Так, в работе [31] показано, что общепринятое определение качества программного обеспечения является не всегда адекватным и должно быть реализовано с использованием модели качества, которая уточняет характеристики в соответствии с рисками и ожиданиями заинтересованных сторон в отношении ПО в их индивидуальном контексте. Некоторые из последних результатов исследований подходов к оценке качества ПО представлены Champion et al. [32]. В этой работе выполнен обзор 75 релевантных научных статей, в котором выделено пять различных подходов, в которых качество рассматривается как эвристика, как ремонтопригодность, как целостная концепция, как структурные особенности ПО и как надежность его функционирования. При этом, как показано в данном обзоре, система оценки характеристик качества ПО, принятая в стандартах SQuaRE, упоминается практически во всех рассмотренных работах.

Ключевые проблемы существующих метрик безопасности SQuaRE

Проведённый анализ позволяет выделить следующие фундаментальные недостатки текущей системы метрик безопасности в стандартах SQuaRE, которые однозначно вытекают из обзора:

1. *Неполнота покрытия требований российских регуляторов.* Стандарты SQuaRE не учитывают требования ГОСТ Р 56939 [1] в части обязательности проверки ПО на наличие НДВ (недекларированных возможностей) и нежелательных конструкций. Отсутствуют метрики, оценивающие наличие утверждённой модели угроз (МУ), что является обязательным требованием для объектов КИИ согласно [3].

2. *Отсутствие метрик безопасности самого кода ПО.* В стандарте [11] безопасность ПО представлена подхарактеристиками целостности, конфиденциальности, неотказуемости, отслеживаемости и аутентичности. Однако отсутствует явная подхарактеристика, оценивающая безопасность исходного и исполняемого кода: наличие уязвимостей, результатов статического (SAST) и динамического (DAST) анализа, анализа состава кода (SCA), тестов на проникновение. Как показано в работах [19, 20], именно качество кода является первичным источником угроз.

3. *Использование среднего арифметического для агрегирования.* В стандарте [11] и большинстве работ используется среднее арифметическое для свёртки метрик [21]. Как убедительно показано в работе [32], это приводит к эффекту маскировки (компенсации), когда критический провал по одной метрике (например, $K_1 = 0$ – отсутствие модели угроз) компенсируется высокими значениями других мет-

рик, давая формально приемлемую оценку. Это недопустимо для объектов КИИ.

4. *Недостаточность метрик неотказуемости и аутентификации.* В стандарте [11] для неотказуемости предусмотрена только одна метрика – «использование цифровой подписи». Однако этого недостаточно – требуются также метрики идентификации, аутентификации согласно стандартам [13, 14], журналирования, меток времени.

5. *Отсутствие учёта приоритетов и критичности.* Метрики конфиденциальности, целостности и доступности имеют разный вес для разных систем, однако в SQuaRE они равнозначны. Как показано в [15, 16], для АСУ ТП приоритетом является доступность и целостность, а для государственных информационных систем – конфиденциальность.

6. *Отсутствие уровней безопасности и чётких порогов.* Стандарты SQuaRE не устанавливают градацию безопасности ПО по уровням (низкий/средний/высокий), что затрудняет принятие решений о допуске ПО к эксплуатации на объектах КИИ.

Таким образом, из проведённого обзора однозначно следует, что существующая система метрик безопасности в SQuaRE требует существенной модернизации с учётом российских требований для КИИ, введения новых подхарактеристик, изменения метода агрегирования и установления уровней безопасности. Рассмотрим характеристику «безопасность» в данной системе стандартов подробнее.

Характеристика «безопасность» в стандартах по качеству ПО

Первый международный стандарт ISO/IEC 9126, содержащий модель качества ПО, появился в системе национальных стандартов России в 1993 г. [6]. Модель качества, представленная в первой части этого стандарта, классифицирует качество ПО в шести основных характеристиках (функциональность, надежность, практичность, эффективность, сопровождаемость, мобильность), которые, в свою очередь, детализированы подхарактеристиками. Вопрос безопасности (security) в данном стандарте входит в число подхарактеристик первой характеристики качества – функциональности ПО.

В 2001 г. стандарт [6] был развит в систему из четырёх взаимосвязанных стандартов, включающих модель качества, внешние и внутренние метрики, а также метрики качества при использовании. В 2002 г. опубликован стандарт по классификации программных средств [7], в котором для выполнения требований безопасности ПО должно классифицироваться по уровню защиты от несанкционированного доступа, контрольному следу (контрольной суммой, позволяющей судить о целостности) и обеспечению живучести (устойчивости к внешним воздействиям). При этом могут быть установлены дополнительные классы требований защиты, в стандарте приведен пример: слабая, средняя, сильная.

Таким образом, в период с 1993 по 2002 г. в международных стандартах была сформирована трехуровневая модель качества ПО, связанная с безопасностью информации: безопасность являлась подхарактеристикой функциональности и характеризовалась степенью защиты от НСД, контрольным следом и обеспечением живучести.

В 2015 г. стандарт [6] был заменен на стандарт [9]. По сравнению с версией [6] к основным характеристикам качества ПО как характеристика, а не подхарактеристика функциональности, в данном стандарте была добавлена «безопасность» с подхарактеристиками «конфиденциальность», «целостность», «безотказность», «отслеживаемость» и «подлинность» [10, 11].

Приведем показатели характеристики «безопасность» одного из недавних стандартов SQuaRE [12].

Показатели конфиденциальности (доступность только легальным пользователям):

- контроль доступа,
- корректность шифрования данных,
- сложность криптографических алгоритмов.

Показатели целостности:

- целостность данных,
- защита данных от повреждения,
- защита от переполнения буфера.

Показатели неотказуемости от выполненных действий:

- использование цифровой подписи.

Показатели отчетности:

- полнота журнала действий пользователей,
- хранение системного журнала.

Показатели аутентификации:

- корректность механизма аутентификации (число реализуемых правил к определенным к реализации),
- полнота реализации правил аутентификации (число реализованных к необходимым реализовать).

Анализ приведенных показателей показывает, что перечень подхарактеристик в стандарте нуждается в совершенствовании с точки зрения современного состояния нормативно-правовой базы РФ и теории защиты информации. Если в стандарте [6], который является базовым стандартом по качеству ПО, показатель защищенности был указан в конце перечисления атрибутов из состава характеристики «функциональные возможности», то в настоящее время в условиях непрерывного роста кибератак показатели безопасности и надежности функционирования ПО выходят на первый план. В частности, еще одна подхарактеристика должна характеризовать отсутствие в ПО уязвимостей, НДВ, выполнением требований защиты от НСД и других недостатков (в терминах ГОСТ Р 56939).

Эта подхарактеристика должна отвечать на вопрос о собственно безопасности исходного и исполняемого кода программ. Также, на наш взгляд, необходимо внести уточнения в первый показатель конфиденциальности, показатели целостности и неотказуемости.

В следующем разделе в виде концепции предложена обновленная система подхарактеристик.

Концепция уточнения подхарактеристик безопасности ПО для систем КИИ

При наличии утвержденной МУ основной метрикой предложено считать степень парирования угроз безопасности, связанных с ПО. В отсутствие МУ возможны два подхода к оценке безопасности ПО.

Первый подход состоит в оценке безопасности собственно программы на разных этапах жизненного цикла разработки [2]. Второй подход учитывает возможности внешнего и внутреннего нарушителя использовать уязвимости и другие недостатки ПО с целью НСД к данным и информационным ресурсам системы, для которой это ПО разрабатывалось. При данном подходе количество метрик существенно больше, чем в первом. Реализуем второй подход. Уточнение подхарактеристик безопасности ПО проведем с минимальными изменениями применяемых на практике национальных стандартов системы SQuaRE, но с дополнениями метрик на основе существующих требований нормативно-правовой базы и анализа научных работ.

Унифицированная система подхарактеристик и метрик

Введём шесть подхарактеристик с уникальными индексами метрик, устраняющими путаницу в обозначениях, которая имела место в предшествующих работах.

Конфиденциальность (С)

Для этой метрики п. 8.7 стандарта [11] предлагает следующую функцию измерения для ответа на вопрос «какая часть конфиденциальных элементов данных защищена от НСД?»:

$$X = 1 - A / B, \quad (1)$$

где A – число конфиденциальных элементов данных, доступ к которым разрешен без авторизации; B – число конфиденциальных элементов данных, доступ к которым ограничен.

Такая система оценки, на наш взгляд, не совсем отражает задачу контроля доступа в рамках обеспечения конфиденциальности защищаемой информации, сформулированной в ст. 2 № 149-ФЗ «Об информации, информационных технологиях и о защите информации»: «Конфиденциальность информации – обязательное для выполнения лицом, получившим доступ к определенной информации, требование не передавать такую информацию третьим лицам без согласия ее обладателя». В случае использования функций контроля доступа средствами операционной системы (ОС) защищаемой информацией в процессах разработки является исходный код ПО, для которого формула (1) применима.

В случае встроенных в разрабатываемое ПО механизмов управления доступом в функции измерения подхарактеристики качества «контроль доступа» предлагается учитывать следующие метрики:

C_1 – метрика соответствия реализации встроенной в ПО модели управления доступом утвержденной политике безопасности, измеряемая как отношение реализованных функций к заданным политикой безопасности. Важно отметить, что C_1 оценивает именно модель управления доступом (разграничение прав субъектов доступа к объектам), в то время как K_1 оценивает наличие МУ безопасности информации в целом. Эти метрики имеют разное смысловое назначение: C_1 отвечает на вопрос «как реализовано разграничение доступа?», а K_1 – на вопрос «какие угрозы учтены при проектировании системы защиты?». МУ (K_1) является исходным документом, определяющим перечень актуальных угроз, а модель управления доступом (C_1) – одной из мер парирования этих угроз. Следовательно, K_1 имеет более высокий приоритет и является заградительной метрикой.

- C_2 – метрика журналирования событий доступа;
- C_3 – периодический аудит прав доступа;
- C_4 – анализ аномальной активности пользователей;
- C_5 – защита от НСД и учётных записей, где $C_i \in [0,1]$ – значения предложенных метрик.

Безопасность кода ПО (К)

Следующей подхарактеристикой безопасности в стандарте [11] является корректность шифрования данных. В случае использования для защиты информации, содержащейся в информационных системах, шифровальных (криптографических) средств защиты информации должны применяться требования, установленные ФСБ России в соответствии с ч. 5 ст. 16 Федерального закона от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации», и требования приказа [5].

Для систем объектов КИИ, согласно 187-ФЗ, применение шифрования в явном виде не является обязательным, а зависит от категории объекта и обрабатываемой в системе информации. Для значимых объектов (ЗО) КИИ обязательным является применение сертифицированных на соответствие требованиям по безопасности систем защиты информации и шифрование каналов связи при передаче конфиденциальных сведений из защищенного периметра через сети общего пользования (перечисленные требования обычно выполняются подсистемой защиты информации), а также использование в ЗО КИИ безопасного ПО в соответствии с [1].

Если шифрование данных встроено в прикладное ПО, для оценки корректности шифрования можно воспользоваться функцией измерения данной подхарактеристики из стандарта [11], которая построена на формуле (1). В случае, когда функция шифрования не встроена в разрабатываемое ПО, а исполняется подсистемой безопасности, логично вместо подхарактеристики «корректность шифрования данных» для разрабатываемого ПО ввести подхарактеристику «безопасность кода ПО», поскольку требования [1] построены на парадигме устранения (блокирования) угроз безопасности, связанных непосредственно с разработанным для системы объекта КИИ ПО: наличия в программе уязвимостей, НДВ,

несовершенной защитой от НСД и другими недостатками ПО, могущими повлечь в процессах разработки и эксплуатации реализацию угроз безопасности информации.

В функции измерения подхарактеристики качества «безопасность (security) кода ПО» предлагается учитывать следующие метрики:

• K_1 – наличие утвержденной МУ безопасности информации, связанной с использованием ПО; при составлении перечня угроз безопасности и их описания рекомендуется учитывать положения ГОСТ Р 58412, а также угрозы безопасности информации Банка данных угроз безопасности информации ФСТЭК России, других источников (например, методологии STRIDE, Open Web Application Security Project (OWASP), DREAD и пр.). В МУ рекомендуется указывать использованную при моделировании методологию, в том числе в случае ее собственной разработки. При отсутствии утвержденной МУ коэффициент $K_1 = 0$. Нейтрализация всех выявленных угроз из утвержденной МУ приведет к значению $K_1 = 1$;

• K_2 – с применением SAST выявлены синтаксические ошибки, стилевые нарушения, простые логические дефекты, устранены все известные на момент анализа уязвимости, в том числе находящиеся в банке данных БДУ ФСТЭК России, а также в базах данных SVE, OWASP;

• K_3 – с применением анализаторов состава исходного кода (SCA) в ПО выявлен весь заимствованный код, который может содержать уязвимости, НДВ и другие недостатки ПО;

• K_4 – с применением динамических анализаторов скомпилированного кода, в том числе фаззинг-анализа, выявлены проблемы производительности, утечки памяти, проблемы многопоточности, недостатки безопасности при взаимодействии с окружением;

• K_5 – проведение тестов на проникновение (пентестов) среды разработки с целью выявления уязвимостей инфраструктуры, используемой в процессе создания ПО (системы контроля версий, сборочные стенды, репозитории артефактов, среды тестирования). Результаты пентестов должны быть задокументированы, а выявленные уязвимости устранены или приняты с оформлением акцептованных рисков. При проведении пентестов и устранении критических уязвимостей $K_5 = 1$; при частичном проведении или наличии неустранённых уязвимостей среднего уровня – $0 \le K_5 < 1$; при отсутствии пентестов – $K_5 = 0$. Важное уточнение: K_5 оценивает безопасность среды разработки, за которую ответствен разработчик ПО. Безопасность среды функционирования (эксплуатационной инфраструктуры заказчика) предлагается оценивать отдельно на этапе приёмочных испытаний с использованием метрики E_1, которая не входит в предложенную систему и должна определяться эксплуатантом на основе его модели угроз и политик безопасности.

Такое разделение устраняет конфликт ответственности: разработчик не может и не должен отвечать за безопасность инфраструктуры, которая ему не принадлежит и находится вне его контроля. Здесь $K_i \in [0,1]$ – значения предложенных метрик. Как показано выше, метрики K_1 и K_2 являются заградительными и учитываются отдельно от агрегированной оценки подхарактеристики Y_K, которая вычисляется по «графируемым» метрикам K_3, K_4, K_5.

Целостность (I)

В функции измерения подхарактеристики качества «целостность» предлагается учитывать следующие метрики:

• I_1 – обеспечение целостности эталонных версий ПО, включающее в себя управление и контроль целостности версий методом вычисления контрольных сумм; при регулярном вычислении и последующей проверке контрольных сумм $I_1 = 1$; при нерегулярном $0 \le I_1 < 1$;

• I_2 – применение электронной подписи (ЭП) эталонных версий ПО для подтверждения целостности артефактов сборки; при применении ЭП к артефактам сборки $I_2 = 1$, когда ЭП не применяется, $I_2 = 0$;

• I_3 – резервирование эталонных версий ПО; при регулярном резервировании $I_3 = 1$;

• I_4 – контроль целостности данных, обрабатываемых ПО; при регулярном контроле целостности данных $I_4 = 1$;

• I_5 – защита данных от повреждения и модификации, в том числе с помощью минимизации доступа; при отсутствии защиты $I_5 = 0$, при наличии $I_5 = 1$.

Аналогично метрикам K_1, K_2 метрика I_1 выделена в качестве заградительной и не входит в агрегированную оценку Y_I, которая вычисляется по метрикам $I_2 - I_5$.

Неотказуемость (N)

В стандарте [11] установлена только одна метрика – «использование цифровой подписи». Как показано в работе [43], одной указанной метрики для обеспечения неотказуемости от выполненного действия недостаточно. В функции измерения подхарактеристики качества «неотказуемость» введем следующие метрики:

• N_1 – обеспечение надежной идентификации и аутентификации пользователя согласно стандартам [13, 14]: $N_1 = 1$ при полном выполнении рекомендаций перечисленных стандартов; при неполном $0 \le N_1 < 1$;

• N_2 – использование меток времени; при регулярном обязательном использовании $N_2 = 1$, в случае, когда метки времени не используются, $N_2 = 0$;

• N_3 – журналирование всех действий пользователей $N_3 = 1$; при ненадлежащем или нерегулярном журналировании $0 \le N_3 < 1$;

• N_4 – применение электронной подписи (ЭП) для фиксации действий пользователей; в отличие от I_2, которая оценивает применение ЭП к артефактам сборки, N_4 оценивает применение ЭП к журналам действий пользователей или к подтверждению факта выполнения операций. При применении ЭП для подтверждения действий $N_4 = 1$; при неприменении $N_4 = 0$.

Отчётность (R)

Метрика «полнота журнала действий пользователей», согласно [11], оценивается как $X = A/B$, где A – количество получений доступа, зафиксированных во всех используемых журналах, B – общее количество получений доступа, осуществленных во время тестирования. Вместо этой метрики, согласно современному состоянию теории защиты информации, предлагается введение следующих метрик:

• R_1 – метрика статистики попыток аутентификации пользователей (отношение количества успешных попыток к количеству всех попыток);

• R_2 – метрика статистики изменения привилегий пользователей (отношение количества успешных попыток к количеству всех попыток);

• R_3 – метрика попыток доступа к объектам (отношение количества успешных попыток к количеству всех попыток);

• R_4 – метрика, определенная в [11] как «хранение системного журнала», также определена как отношение действительной продолжительности хранения системного журнала в стабильных хранилищах с ограниченным доступом к журналу к требуемому периоду хранения системного журнала в стабильных хранилищах.

Аутентификация (A)

В стандарте [11] установлены две метрики: «корректность механизма аутентификации (число реализуемых правил к определенным к реализации)» и «полнота реализации правил аутентификации (число реализованных к необходимым реализовать)». Предлагается расширить количество метрик показателей аутентификации до четырех:

• A_1 – метрика соответствия организации процесса аутентификации политике безопасности по доступу;

• A_2 – метрика соответствия уровня строгости криптографии, применяемой в процессе аутентификации, требованиям нормативных требований и стандартов;

• A_3 – метрика корректности реализации механизма аутентификации (действительно разные одновременно применяемые факторы, ограничение числа неуспешных попыток аутентификации и др.), измеряемая как количество реализованных правил к количеству определенных к реализации;

• A_4 – метрика полноты реализации правил аутентификации (число реализованных правил к количеству необходимых к реализации).

Двухконтурная модель агрегирования: заградительные и градуируемые метрики

Из 27 предложенных метрик три – K_1, K_2 и I_1 – представляют собой нормативно обоснованные требования первого уровня (наличие МУ, статический анализ кода без критических уязвимостей, контроль целостности эталонных версий). Для объектов КИИ их невыполнение не может компенсироваться никакими другими показателями. Остальные 24 метрики являются градуируемыми: они описывают степень (полноту, зрелость) выполнения требования и содержательно допускают промежуточные значения без фатальных последствий для итоговой оценки.

Чтобы эти два класса метрик действительно играли разную роль в итоговой оценке, а не только на уровне терминологии, ниже предлагается пересмотренная схема агрегирования, состоящая из двух независимых контуров: контура качества (градуируемые метрики) и контура допуска (заградительные метрики).

Контур качества

Заградительные метрики K_1, K_2, I_1 исключаются из подхарактеристик K и I и не участвуют в вычислении Y_K, Y_I: подхарактеристика K вычисляется по метрикам $\{K_3, K_4, K_5\}$, подхарактеристика I – по метрикам $\{I_2, I_3, I_4, I_5\}$. Внутри каждой из шести подхарактеристик по-прежнему используется гармоническое среднее (формула (5)) – это сохраняет устойчивость к маскировке провала внутри одного направления безопасности (например, если из трёх метрик безопасности кода одна равна нулю, подхарактеристика Y_K обнуляется, что непосредственно указывает на наличие критического дефекта в данной подхарактеристике.

Однако на верхнем уровне – при объединении шести подхарактеристик $Y_C, Y_K, Y_I, Y_N, Y_R, Y_A$ в базовую оценку S_0 – гармоническое среднее уже неприменимо: система из шести содержательно разнородных направлений безопасности не должна обнуляться целиком из-за провала в одном второстепенном направлении, если этот провал не относится к нормативно закреплённым заградительным требованиям. Поэтому для верхнего уровня используется взвешенное среднее арифметическое

$$S_0 = \sum_{j=1}^{6} w_j Y_j, \quad \sum_{j=1}^{6} w_j = 1 \ , \qquad (2)$$

где при отсутствии дополнительного обоснования весов для конкретного класса систем рекомендуется $w_j = 1/6$ (равнозначность подхарактеристик, см. ниже разд. «Заградительные метрики»).

Контур допуска

Заградительные метрики образуют независимую бинарную (точнее, полубинарную) проверку допуска

$$G = \mathbb{I}(K_1 > 0) \cdot \mathbb{I}(K_2 > 0) \cdot \mathbb{I}(I_1 > 0) \in \{0,1\}. \quad (3)$$

Итоговая оценка

Контур допуска действует как жёсткий множитель поверх контура качества:

$$S = G \cdot S_0 = \mathbb{I}(K_1 > 0 \wedge K_2 > 0 \wedge I_1 > 0) \cdot \sum_{j=1}^{6} w_j Y_j . \quad (4)$$

Рисунок 1 иллюстрирует эту двухконтурную схему.

Почему это устраняет прежнее противоречие

В формуле (4) обнуление любой из 24 градуируемых метрик снижает соответствующую Y_j (вплоть до нуля, если внутри подхарактеристики нет ни одной положительной метрики), что пропорционально своему весу w_j уменьшает S_0, но не обнуляет её целиком, если остальные пять подхарактеристик положительны. Обнуление же любой из трёх заградительных метрик K_1, K_2, I_1 обнуляет G и, следовательно, S – независимо от значения S_0. Данная схема обеспечивает содержательную, а не только формальную асимметрию между двумя классами метрик.

Рисунок 2 показывает эту асимметрию на числовом примере: провал одной градуируемой метрики (например, $C_3 = 0$, отмечено на левой паре столбцов) снижает S примерно на величину её веса, тогда как провал заградительной метрики (например, $K_1 = 0$, правая пара столбцов) обнуляет S полностью.

Заградительные метрики

1. K_1 – *наличие утверждённой модели угроз безопасности информации (МУ).*

Обоснование. Согласно требованиям стандарта [1] и методических рекомендаций ФСТЭК России, модель угроз является фундаментальным документом, определяющим перечень актуальных угроз безопасности информации, поверхность атаки, а также обоснование выбора и достаточности мер защиты. Без МУ невозможно корректно определить, какие угрозы должны быть парированы, какие метрики являются приоритетными и какие средства защиты требуются.

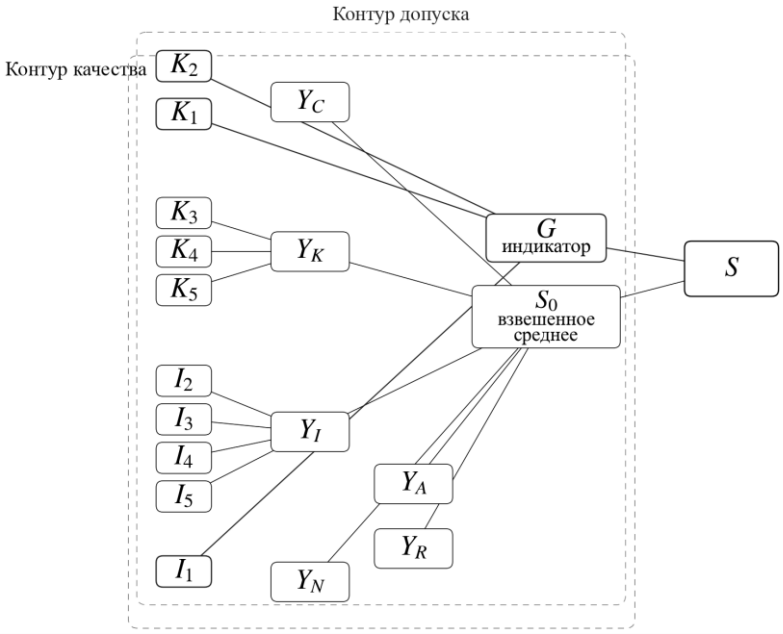


Рис. 1. Двухконтурная схема агрегирования: градуируемые метрики формируют базовую оценку S_0 через гармоническое среднее внутри подхарактеристик и взвешенное среднее между ними; заградительные метрики образуют независимый контур допуска G, действующий как жёсткий множитель

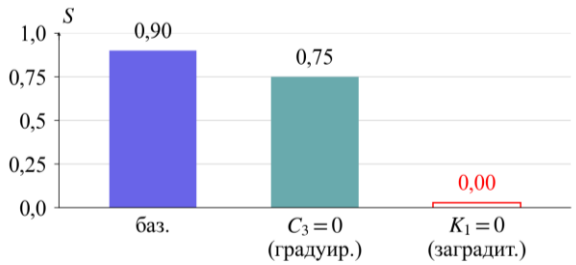


Рис. 2. Реакция итоговой оценки S на провал разных классов метрик при прочих равных: обнуление обычной (градуируемой) метрики снижает S пропорционально её весу, а обнуление любой заградительной метрики обнуляет S полностью

Как показано в работах [15, 16], для критических систем проектирование защиты должно начинаться именно с анализа угроз. Отсутствие МУ означает, что система защиты проектируется «вслепую», что недопустимо для объектов КИИ согласно ст. 11 Федерального закона 187-ФЗ и приказу [3].

2. K_2 – *проведение статического анализа исходного кода (SAST) с устранением всех выявленных уязвимостей уровней Critical и High.*

Обоснование. Статический анализ кода является обязательным требованием ГОСТ Р 56939 [1] для всех этапов разработки безопасного ПО. Как пока-

зано в работах [19, 20], именно недостатки кода (уязвимости, НДВ, нежелательные конструкции) являются первичным источником угроз безопасности информации. Отсутствие SAST-анализа или наличие неустранённых критических уязвимостей в коде означает, что ПО содержит заведомо известные недостатки, что делает его неприемлемым для эксплуатации на объектах КИИ. Согласно приказу [3] и п. 5.3 стандарта [1], все выявленные уязвимости должны быть устранены до ввода ПО в эксплуатацию.

3. I_1 – *обеспечение целостности эталонных версий ПО (регулярное вычисление и проверка контрольных сумм).*

Обоснование. Целостность эталонных версий ПО является критическим требованием для объектов КИИ, поскольку нарушение целостности может быть следствием внедрения вредоносного кода, НДВ или несанкционированных модификаций на этапах разработки, сборки или поставки. Как показано в работе [17], для критических систем целостность ПО является одним из ключевых факторов обеспечения как информационной, так и функциональной безопасности.

Согласно [1], управление конфигурациями и контроль целостности версий являются обязательными требованиями к процессу безопасной разработки. Отсутствие такого контроля делает невозможным гарантировать, что эксплуатируемое ПО соответствует эталонному образцу, прошедшему все проверки.

Правило применения заградительных метрик

В отличие от предыдущей редакции методики, где заградительный статус K_1, K_2, I_1 был заявлен, но не выражен в формулах явно (поскольку эти метрики оставались внутри гармонического среднего подхарактеристик Y_K и Y_I наравне с градуируемыми метриками, а это уже само по себе обнуляло всю систему при любом нуле, в формуле (4) заградительные метрики выведены из контура качества и образуют самостоятельный контур допуска G. Это даёт содержательно корректное правило: нулевое значение K_1, K_2 или I_1 блокирует положительное заключение о безопасности ПО независимо от значений всех остальных 24 метрик, тогда как нулевое значение любой из 24 градуируемых метрик снижает итоговую оценку S пропорционально весу соответствующей подхарактеристики, но не обнуляет её автоматически.

В табл. 1 представлена сводка заградительных метрик.

Унифицированное агрегирование метрик

Для каждой подхарактеристики вычисляется агрегированная оценка $Y_j \in [0,1]$ по формуле гармонического среднего, применяемой к градуируемым метрикам данной подхарактеристики (для K – это $\{K_3, K_4, K_5\}$, для I – это $\{I_2, \ldots, I_5\}$, для остальных четырёх подхарактеристик – полный состав их метрик):

$$Y_j = \begin{cases} 0, \text{ если } \exists i : m_i = 0; \\ \dfrac{n}{\sum_{i=1}^{n} \dfrac{1}{m_i}}, \text{ в ином случае,} \end{cases} \qquad (5)$$

где n – число градуируемых метрик в подхарактеристике, m_i – значение i-й метрики.

Таблица 1

Заградительные метрики безопасности ПО для объектов КИИ

Метрика	Содержание	Нормативное обоснование	Приоритет
K_1	Наличие утверждённой МУ	ГОСТ Р 56939, Приказ ФСТЭК №239, ст. 11 ФЗ 187-ФЗ	1 (высший)
K_2	Проведение SAST-анализа с устранением уязвимостей Critical и High	ГОСТ Р 56939, п. 5.3	2
I_1	Обеспечение целостности эталонных версий ПО (контрольные суммы)	ГОСТ Р 56939, п.5.4	3

Итоговая оценка безопасности ПО S вычисляется согласно формуле (4) как произведение индикатора допуска G на взвешенное среднее арифметическое шести подхарактеристик S_0.

Обоснование иерархической свёртки

Применение двухуровневой схемы агрегирования (гармоническое среднее внутри подхарактеристик, взвешенное арифметическое среднее между ними) является сознательным методологическим выбором, а не математической необходимостью. Этот выбор обусловлен тремя соображениями.

1. *Содержательная неоднородность подхарактеристик*. Каждая из шести подхарактеристик (конфиденциальность, безопасность кода, целостность, неотказуемость, отчётность, аутентификация) описывает качественно различный аспект безопасности ПО. Свёртка всех градуируемых метрик в одну оценку без промежуточной группировки математически возможна, но содержательно неоправданна, поскольку стирает границы между различными классами требований.

2. *Равноправность подхарактеристик*. Для объектов КИИ все шесть подхарактеристик являются обязательными в силу требований стандарта [1] и приказа [3]. Ни одна из них не может быть признана второстепенной с точки зрения самого факта необходимости её контроля. Взвешенное среднее с $w_j = 1/6$ обеспечивает равный вклад каждой подхарактеристики в базовую оценку S_0.

3. *Разный характер требований внутри и между подхарактеристиками*. Метрики внутри одной подхарактеристики (например, K_3, K_4, K_5 – SCA, DAST, пентест среды разработки) являются взаимодополня-

ющими техниками контроля одного и того же аспекта безопасности кода, и провал любой из них действительно не должен маскироваться остальными – отсюда гармоническое среднее. Подхарактеристики же между собой описывают content-независимые направления (конфиденциальность vs. отчётность и т.д.), и полное взаимное сглаживание между ними столь же нежелательно, как и его отсутствие: значимый провал в одном направлении должен ощутимо снижать S_0 (что и обеспечивает вес $w_j = 1/6$ при провале целой подхарактеристики), но не должно полностью обнулять результаты по остальным пяти направлениям, если ни одно из них не относится к нормативно закреплённым заградительным требованиям.

Для количественной оценки поведения новой схемы на данных модельного примера (см. раздел ниже «Модельный пример расчёта») при сценарии после устранения критического дефекта ($K_1 = 1$, $G = 1$) получается $S_0 = 0{,}861$ (см. табл. 5), что практически совпадает с прежним значением 0,867, полученным по формуле с гармоническим средним верхнего уровня и включением K_1, K_2, I_1 в состав Y_K, Y_I – за счёт исключения K_1, K_2 из Y_K и использования более мягкого агрегирования на верхнем уровне. Такое поведение ожидаемо и корректно: часть чувствительности к провалам теперь сосредоточена именно в контуре допуска G, а не размазана по формуле гармонического среднего верхнего уровня, как это было прежде.

Такой подход обеспечивает:

• высокую чувствительность к низким значениям градуируемых метрик *внутри* каждой подхарактеристики (гармоническое среднее);

• умеренную, пропорциональную весу реакцию S_0 на провал целой подхарактеристики, вызванный градуируемыми метриками;

• полное и безусловное обнуление итоговой оценки S при провале любой из трёх нормативно обоснованных заградительных метрик – независимо от значения S_0;

• согласованность методологии на всех уровнях и, что важно, реальное (а не только терминологическое) различие между заградительными и градуируемыми метриками.

Уровни безопасности разрабатываемого программного обеспечения

На основе предложенных метрик безопасности и с учётом требований нормативных документов предлагается введение уровней безопасности разрабатываемого ПО (табл. 2).

Обоснование пороговых значений. Пороговые значения 0,5 и 0,75 выбраны на основе следующего обоснования. В проанализированных нормативных документах [1, 4, 5] отсутствует прямое указание на конкретные числовые пороги 50 и 75% для оценки безопасности ПО. Приведённые значения являются результатом экспертного обобщения практики применения указанных документов, а также сложившейся отраслевой традиции, согласно которой, выполнение менее 50% требований не позволяет сделать вывод о контролируемости процесса разработки, а выполнение более 75% требований соответствует высокому уровню зрелости процессов.

Аналогичные пороговые значения (0,5 и 0,75) используются в оценочных шкалах стандарта [9] для дифференциации уровней качества, а также в методиках оценки зрелости процессов разработки (например, CMMI, SPICE). Таким образом, предлагаемые пороги являются обоснованными по аналогии, а не прямым нормативным предписанием и могут быть скорректированы для конкретных классов систем на основе анализа рисков.

При $S = 0$ оформляется мотивированное заключение о невозможности использования ПО до устранения причин, приведших к обнулению контура допуска G (т.е. до устранения нулевого значения хотя бы одной заградительной метрики).

В зависимости от уровня рисков реализации выявленных угроз безопасности информации количество уровней безопасности и границы их определения могут варьироваться.

Например, для крупных государственных холдингов, делающих первые шаги к обеспечению безопасности разрабатываемого заказного ПО, могут быть введены начальные уровни безопасности. Допускается уточнение пороговых значений интерпретации общей оценки для конкретных классов или категорий значимости объектов КИИ на основе анализа рисков.

Обоснование выбора гармонического среднего

Для обоснования выбора в табл. 3 приведено сравнение гармонического среднего с другими методами на характерных для КИИ наборах метрик (сравнение относится к агрегированию внутри подхарактеристики, где действительно требуется полная чувствительность к нулю).

Таблица 2

Уровни безопасности ПО для объектов КИИ

Значение S	Уровень безопасности	Интерпретация	Нормативное обоснование
$S = 0$	Недопустимо	Использование ПО запрещено до устранения критических дефектов	Заградительные метрики (Приказ № 239, ГОСТ Р 56939)
$0 < S < 0{,}5$	Низкий	Использование ПО не допускается до устранения критических дефектов	Выполнение < 50% требований (экспертная интерпретация)
$0{,}5 \le S < 0{,}75$	Средний	Допускается с ограничениями, требуется план повышения	Выполнение 50–75% требований, требуется доработка
$0{,}75 \le S \le 1$	Высокий	Допускается для объектов КИИ без ограничений	Выполнение > 75% требований (ISO/IEC 25010, отраслевая практика)

Таблица 3

Сравнение методов агрегирования

Тип свёртки	Пример: [0; 0,9; 0,9; 0,9]	Пример: [0,5; 0,9; 0,9; 0,9]	Пример: [0,9; 0,9; 0,9; 0,9]
Среднее арифметическое	0,68	0,80	0,90
Медиана	0,90	0,90	0,90
Минимум	0	0,50	0,90
Перемножение	0	0,36	0,66
Гармоническое	0	0,71	0,90

Как видно из табл. 3:

• среднее арифметическое маскирует критический ноль, давая 0,68 при наличии нулевой метрики;

• медиана полностью игнорирует низкие значения;

• минимум избыточно жесткий при отсутствии нуля;

• перемножение чрезмерно наказывает за некритические снижения;

• гармоническое среднее обеспечивает оптимальную чувствительность: обнуляется при наличии нуля, существенно снижается при наличии низкого значения (0,5) и практически не меняется при высоких значениях.

Именно эти свойства делают гармоническое среднее уместным *внутри* подхарактеристики (где метрики – взаимодополняющие проверки одного аспекта), но неуместным на верхнем уровне между подхарактеристиками, что и обосновывает переход к взвешенному арифметическому среднему в формуле (4).

Модельный пример расчёта

Рассмотрим гипотетический проект по разработке прикладного ПО для объекта КИИ. Значения метрик (0–1) получены экспертным путём на основе результатов тестирования и анализа (табл. 4).

Для визуального анализа исходных данных построены типовые карты метрик. Такая форма представления позволяет быстро обнаруживать не только средний уровень выполнения требований, но и локальные провалы внутри отдельных подхарактеристик.

Рисунок 3 показывает пример определения всех 27 метрик для исходного сценария (с $K_1 = 0$): заградительные метрики выделены рамкой, единственная критическая ячейка – K_1.

Расчёт по формуле гармонического среднего для градуируемых метрик подхарактеристик:

$$Y_C = \frac{5}{1/0{,}8+1/0{,}9+1/0{,}7+1/0{,}6+1/1{,}0} = 0{,}774;$$

$$Y_K = \frac{3}{1/0{,}8+1/0{,}7+1/0{,}9} = 0{,}792,$$

(по K_3, K_4, K_5, без учёта заградительных K_1, K_2);

$$Y_I = \frac{4}{1/1{,}0+1/0{,}8+1/0{,}9+1/1{,}0} = 0{,}917,$$

(по $I_2, \ldots, I_5$, без учёта заградительного I_1);

$$Y_N = \frac{4}{1/1{,}0+1/0{,}9+1/1{,}0+1/0{,}8} = 0{,}917;$$

$$Y_R = \frac{4}{1/0{,}9+1/0{,}9+1/0{,}8+1/1{,}0} = 0{,}894;$$

$$Y_A = \frac{4}{1/0{,}9+1/0{,}8+1/0{,}9+1/0{,}9} = 0{,}873.$$

Контур допуска:

$G = \mathbb{I}(K_1 = 0{,}0 > 0) \cdot \mathbb{I}(K_2 = 0{,}9 > 0) \cdot \mathbb{I}(I_1 = 0{,}9 > 0) = 0$,

поскольку $K_1 = 0$.

Таблица 4

Исходные данные для модельного примера

C	K (в т.ч. K_1, K_2)	I (в т.ч. I_1)	N	R	A
[0,8; 0,9; 0,7; 0,6; 1,0]	[$K_1 = 0{,}0$; $K_2 = 0{,}9$; 0,8; 0,7; 0,9]	[$I_1 = 0{,}9$; 1,0; 0,8; 0,9; 1,0]	[1,0; 0,9; 1,0; 0,8]	[0,9; 0,9; 0,8; 1,0]	[0,9; 0,8; 0,9; 0,9]

C	0,8	0,9	**0,7**	**0,6**	1,0
K	**0,0**	**0,9**	0,8	**9,7**	0,9
I	**0,9**	1,0	0,8	0,9	1,0
N	1,0	0,9	1,0	0,8	
R	0,9	0,9	0,8	1,0	
A	0,9	0,8	0,9	0,9	
	1	2	3	4	5

Рис. 3. Пример определения метрик исходного сценария

Таким образом, отсутствие утверждённой модели угроз ($K_1 = 0$) обнуляет контур допуска G и, следовательно, итоговую оценку $S = G \cdot S_0 = 0$, даже при высоких значениях S_0 по всем шести подхарактеристикам. Это соответствует логике заградительного контроля, когда фундаментальное требование (наличие модели угроз) является обязательным.

Описание результатов расчёта

Расчёт демонстрирует независимую работу двух контуров модели. Контур качества по всем шести подхарактеристикам показывает высокие значения: конфиденциальность – 0,774, безопасность кода (по градуируемым метрикам) – 0,792, целостность (по градуируемым метрикам) – 0,917, неотказуемость – 0,917, отчётность – 0,894, аутентификация – 0,873 (табл. 5).

Таблица 5

Агрегированные оценки подхарактеристик безопасности ПО (контур качества)

Подхарактеристика	Оценка Y_j	Вклад в S_0 ($w_j Y_j$, $w_j = 1/6$)
Конфиденциальность	0,774	0,129
Безопасность кода (градуир.)	0,792	0,132
Целостность (градуир.)	0,917	0,153
Неотказуемость	0,917	0,153
Отчётность	0,894	0,149
Аутентификация	0,873	0,146
S_0 (сумма)	0,861	

Несмотря на это, контур допуска $G = 0$ из-за $K_1 = 0$, поэтому итоговая оценка $S = 0$: ПО не может быть допущено к применению на объекте КИИ, неза-

висимо от того, насколько хорошо выполнены остальные 24 метрики.

Дополнительно рассмотрим чувствительность итоговой оценки к устранению критического дефекта.

Если после разработки и утверждения МУ значение метрики K_1 будет повышено до 1,0 (индикатор $\mathbb{I}(K_1 > 0)$ становится равным 1, $G = 1$) при сохранении остальных метрик безопасности кода на исходном уровне, оценка подхарактеристики безопасности кода Y_K не меняется (поскольку K_1 более не входит в её расчёт), а итоговая оценка становится равной $S = S_0 = 0{,}861$. Такой результат переводит проект из состояния недопустимости в высокий уровень безопасности.

Для сравнения приведём также реакцию на провал *градуируемой* (не заградительной) метрики: если вместо K_1 обнулить, например, $C_3 = 0$ при $K_1 = 1$ (МУ утверждена), то $Y_C = 0$ (гармоническое среднее внутри подхарактеристики C обнуляется), контур допуска остаётся $G = 1$ (поскольку K_1, K_2, I_1 по-прежнему положительны), и итоговая оценка составит

$$S = 1 \cdot \left(0 \cdot \frac{1}{6} + 0{,}792 \cdot \frac{1}{6} + 0{,}917 \cdot \frac{1}{6} + 0{,}917 \cdot \frac{1}{6} + \right.$$
$$\left. + 0{,}894 \cdot \frac{1}{6} + 0{,}873 \cdot \frac{1}{6} \right) = 0{,}732.$$

В табл. 6 сведены все три сценария и наглядно показана искомая асимметрия: провал заградительной метрики блокирует оценку полностью, провал градуируемой метрики – существенно, но не полностью.

Содержательно полученные результаты позволяют выделить три практических вывода.

Во-первых, наиболее сильными сторонами рассматриваемого проекта являются целостность, неотказуемость и отчётность: все три оценки находятся выше 0,89.

Во-вторых, конфиденциальность имеет наименьшее ненулевое значение среди градуируемых подхарактеристик (0,774), поэтому после устранения критического дефекта именно это направление целесообразно рассматривать как ближайший резерв повышения итогового уровня.

Таблица 6

Сравнение сценариев: реакция *S* на разные классы дефектов

Сценарий оценки	*G*	*S*
Исходное состояние: заградительная метрика $K_1 = 0$	0	0,000
После утверждения модели угроз: все заградительные метрики > 0	1	0,861
Провал одной градуируемой метрики ($C_3 = 0$) при исправной модели угроз	1	0,732

В-третьих, двухконтурная модель впервые в данной методике действительно различает последствия провала заградительной и градуируемой метрики: первое полностью блокирует допуск, второе понижает уровень (в примере – со «высокого» до «среднего»), но не блокирует его автоматически.

С учётом приведённых иллюстраций итоговая интерпретация результата следующая: ПО не может быть допущено к применению на объекте КИИ до устранения критического дефекта, связанного с отсутствием утверждённой МУ (табл. 7). После разработки и утверждения МУ требуется повторная оценка метрики K_1, пересчёт контура допуска G и повторное вычисление итоговой оценки безопасности S.

Таблица 7

Интерпретация общей оценки безопасности ПО

Диапазон	Уровень	Пример сценария
$S = 0$	Недопустимо	Исходное состояние ($K_1 = 0$, $G = 0$)
$0 < S < 0{,}5$	Низкий уровень	Требуется доработка
$0{,}5 \le S < 0{,}75$	Средний уровень	Провал градуируемой метрики C_3 ($S = 0{,}732$)
$0{,}75 \le S \le 1{,}0$	Высокий уровень	После устранения K_1 все метрики в норме ($S = 0{,}861$)

Ограничения метода

Предложенный подход имеет следующие ограничения, которые необходимо учитывать при практическом применении:

1. *Субъективность значений метрик.* Значения метрик $C_i, K_i, I_i, N_i, R_i, A_i$ выставляются экспертами на основе анализа документации, результатов тестирования и аудита кода. Это вносит элемент субъективизма, особенно для метрик, не имеющих чётких количественных критериев (например, C_4 – анализ аномальной активности). Рекомендуется формализация критериев для каждой метрики с указанием пороговых значений (например, $K_2 = 1$ только при покрытии тестами SAST ≥ 80% кода и устранении всех уязвимостей уровнем Critical и High).

2. *Статичность оценки.* Методика даёт точечную оценку на момент проверки, но безопасность ПО меняется в процессе эксплуатации: появляются новые уязвимости, изменяется среда функционирования, обновляются версии компонентов. Требуется периодический пересчёт метрик (например, ежеквартально или после каждого обновления ПО).

3. *Неполнота покрытия для специфических классов ПО.* Набор из 27 метрик не является исчерпывающим для всех классов ПО. Например, для вебприложений могут потребоваться дополнительные метрики по защите от OWASP Top 10 (инъекции, нарушение аутентификации, чувствительные данные и т.д.), для встраиваемых систем – метрики по защите от аппаратных атак. Предложенная система является базовой и может быть расширена под конкретный класс систем.

4. *Отсутствие весовых коэффициентов приоритетности между подхарактеристиками.* В базовой версии методики все шесть подхарактеристик

считаются равнозначными ($w_j = 1/6$), однако для разных классов систем приоритеты могут различаться. Например, для АСУ ТП приоритетом являются доступность и целостность, а для государственных информационных систем – конфиденциальность. Формула (4) допускает неравномерные веса w_j, но их обоснование требует отдельного анализа рисков для конкретной системы.

5. *Трудозатратность сбора данных.* Для вычисления всех 27 метрик требуется проведение значительного объёма работ: статический и динамический анализ кода, анализ состава, тесты на проникновение, аудит документации, проверка журналов и т.д. Для небольших организаций это может быть экономически нецелесообразно.

6. *Зависимость от актуальности баз уязвимостей.* Метрики K_2 и K_3 зависят от актуальности используемых баз данных уязвимостей (БДУ ФСТЭК, SVE, OWASP, CVE). Если базы не обновляются, оценка может не отражать реального уровня безопасности.

7. *Остаточный риск veto-логики внутри подхарактеристик.* Хотя переход к двухконтурной схеме устранил основной практический риск – полное обнуление S из-за одной случайно не измеренной градуируемой метрики, гармоническое среднее внутри каждой подхарактеристики по-прежнему обнуляет Y_j при любом нулевом значении входящей в неё метрики. Если, например, метрика C_4 (анализ аномальной активности) в организации попросту не отслеживается, это законно должно снижать Y_C, но объективно не всегда отличимо от ситуации «метрика не измерялась» от ситуации «требование провалено». Рекомендуется, как и прежде, предусмотреть отдельное состояние «не измерено» (с последующим требованием обязательного измерения до итоговой аттестации), а не автоматически трактовать отсутствие данных как $m_i = 0$.

8. *Разграничение зон ответственности.* Предложенная система метрик оценивает безопасность ПО с точки зрения разработчика. Безопасность инфраструктуры эксплуатации (серверы, сеть, СЗИ, ОС) требует отдельной оценки и не входит в предметную область данной работы.

Рекомендуется при практическом применении методики чётко фиксировать границы ответственности: разработчик предоставляет оценку по предлагаемым метрикам (включая K_5 для среды разработки), а заказчик на основе этих данных и собственного анализа эксплуатационной среды принимает решение о допуске ПО.

9. *Для АСУ ТП и других технологических систем* рекомендуется пересмотреть весовые коэффициенты w_i в формуле (4) с учётом приоритета доступности и целостности над конфиденциальностью, а также предусмотреть дополнительные метрики, оценивающие время восстановления после сбоев и устойчивость к инцидентам.

Заключение

Проведен краткий анализ научных работ и стандартов, определяющих качество ПО применительно к характеристике «безопасность». Показано, что на текущий момент подхарактеристики показателя безопасности ПО как одной из основных характеристик качества ПО, установленные в системе стандартов SQuaRE, нуждаются в совершенствовании с привлечением специалистов по информационной безопасности. Предложена концепция введения системы количественных оценок безопасности ПО и новой подхарактеристики безопасности ПО «уровни безопасности ПО» как показателя качества разрабатываемого ПО.

Предложена простая трехуровневая система формирования уровней безопасности ПО, опирающаяся на двухконтурную модель агрегирования, в которой заградительные и градуируемые метрики играют математически различную функциональную роль.

Предложенная концепция уровней безопасности нуждается в дальнейшей проработке с точки зрения развития теории защиты информации, в частности, уточнения требований и характеристик уровней безопасности [33]. В частности, установлены, но подлежат дальнейшему дополнению и уточнению следующие показатели безопасности ПО:

- обязательность проверки на наличие известных уязвимостей и их устранение;
- отсутствие в ПО НДВ и нежелательных конструкций;
- обеспечение целостности ПО;
- необходимость наличия в ПО функций разграничения доступа;
- наличие в ПО встроенных функций идентификации и аутентификации, различающихся по уровням доверия, установленным в национальных стандартах [7, 9];
- применение в программном обеспечении криптографических функций согласно требованиям при необходимости использования сервисов безопасности на базе инфраструктуры открытых ключей;
- обеспечение функциональной надежности ПО;
- наличие в ПО функций резервирования (архивирования) и восстановления информации.

Предложенная в данной работе система количественных оценок безопасности разрабатываемого ПО для ИС объектов КИИ может быть расширена на другие классы систем в зависимости от решения владельца (оператора). Для каких-то систем обеспечение конфиденциальности, доступности и целостности может иметь разные приоритеты. Например, для некоторых государственных систем приоритеты могут быть выстроены в порядке: доступность, целостность, конфиденциальность. Заданные уровни безопасности для таких систем могут корректироваться весовыми коэффициентами w_j при вычислении S_0.

В продолжение работы планируется исследовать вопросы оценки остаточных рисков ПО, достигнутого уровня безопасности ПО [34, 35], соотношения требований функциональной безопасности и без-

опасности информации при использовании ПО, а также провести сравнение требований безопасности и надежности ПО.

Литература

1. ГОСТ Р 56939–2024. Защита информации. Разработка безопасного программного обеспечения. Общие требования. – М.: Стандартинформ, 2024.

2. ГОСТ Р 12207–2010. Информационная технология. Системная и программная инженерия. Процессы жизненного цикла программных средств. – М.: Стандартинформ, 2010.

3. Об утверждении пребований по обеспечению безопасности значимых объектов КИИ Российской Федерации. Утв. приказом ФСТЭК России от 25 декабря 2017 г. № 239 [Электронный ресурс]. – URL: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239 (дата обращения: 09.10.2026).

4. Методический документ ФСТЭК России. Состав и содержание мероприятий и мер по защите информации, содержащейся в информационных системах. Утв. ФСТЭК России 12 апреля 2026 г. [Электронный ресурс]. – URL: https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-12-aprelya-2026-g (дата обращения: 09.10.2026).

5. Приказ ФСТЭК России от 11.04.2025 № 117 «Об утверждении Требований о защите информации, содержащейся в государственных информационных системах, иных информационных системах государственных органов, государственных унитарных предприятий, государственных учреждений» (Зарегистрирован в Минюсте России 16.06.2025 № 82619) // Официальный интернет-портал правовой информации. – URL: http://publication.pravo.gov.ru/document/0001202506170011 (дата обращения: 09.10.2026).

6. ГОСТ Р ИСО/МЭК 9126-93. Информационная технология. Оценка программной продукции. Характеристика качества и руководства по их применению. – Введ. 1993-12-28. – М.: ИПК «Изд-во стандартов», 2003.

7. ГОСТ Р ИСО/МЭК ТО 12182–2002. Информационная технология. Классификация программных средств. Введ. 2002-06-11. – М.: ИПК «Изд-во стандартов», 2002.

8. ГОСТ Р ИСО/МЭК 25000–2021. Системная и программная инженерия. Требования и оценка качества систем и программных средств (SQuaRE). Руководство. Введ. 2021–10–25. – М.: ФГБУ «РСТ», 2021.

9. ГОСТ Р ИСО/МЭК 25010–2015. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов. Введ. 2015-05-29. – М.: Стандартинформ, 2015.

10. ГОСТ Р ИСО/МЭК 25022–2021. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Измерение качества при использовании. Введ. 2021-10-25. – М.: ФГБУ «РСТ», 2021.

11. ГОСТ Р ИСО/МЭК 25023–2021. Системная и программная инженерия. Требования и оценка качества систем и программной продукции (SQuaRE). Измерения качества системы и программной продукции. Введ. 2021-11-19. – М.: ФГБУ «РСТ», 2021.

12. ГОСТ Р 70921-2023. Системная и программная инженерия. Требования и оценка качества систем и программной продукции (SQuaRE). Концепция требований к качеству. Введ. 2023-09-14. – М.: Российский институт стандартизации, 2023.

13. ГОСТ Р 70262.1–2022. Защита информации. Идентификация и аутентификация. Уровни доверия к идентификации. – М.: Стандартинформ, 2022.

14. ГОСТ Р 70262.2–2025. Защита информации. Идентификация и аутентификация. Уровни доверия к аутентификации. – М.: Стандартинформ, 2025.

15. Шубинский И.Б. Функциональная надежность информационных систем. Методы анализа. – Ульяновск: Печатный двор, 2012. – 296 с.

16. Шубинский И.Б. Надежные отказоустойчивые информационные системы. Методы синтеза. – Ульяновск: Печатный двор, 2016. – 544 с.

17. Шубинский И.Б. Функциональная безопасность систем управления на железнодорожном транспорте / И.Б. Шубинский, Е.Н. Розенберг. – Москва; Вологда: Инфра-Инженерия, 2023. – 360 с.

18. Марков А.С. Модели оценки и планирования испытаний программных средств по требованиям безопасности информации // Вестник МГТУ им. Н.Э. Баумана. – Сер.: Приборостроение. – 2011. – S.1. – С. 90–103.

19. Марков А.С. Систематика уязвимостей и дефектов безопасности программных ресурсов / А.С. Марков, А.А. Фадин // Защита информации. Инсайд. – 2013. – № 3 (51). – С. 56–61.

20. Марков А.С. Эвристический анализ безопасности программного кода / А.С. Марков, В.А. Матвеев, А.А. Фадин, В.Л. Цирлов // Вестник МГТУ им. Н.Э. Баумана. – Сер.: Приборостроение. – 2016. – № 1 (106). – С. 98–111. DOI: 10.18698/0236-3933-2016-1-98-111.

21. Мерзлякова Е.Ю. Обзор методов верификации и оценки качества программного обеспечения / Е.Ю. Мерзлякова, Е.В. Янченко // Вестник СибГУТИ. – 2023. – Т. 17, № 1. – С. 92–106. DOI: 10.55648/1998-6920-2023-17-1-92-106.

22. Кулямин В.В. Технологии программирования. Компонентный подход. – Лекция 5: Качество программного обеспечения и методы его контроля [Электронный ресурс]. – URL: https://studfile.net/preview/993777/page:14/ (дата обращения: 22.07.2026).

23. Goertzel K.M. Software Security Assurance: A State-of-Art Report (SAR) / K.M. Goertzel, Th. Winograd, H.L. McKinley et al. – Fort Belvoir, VA: Information Assurance Technology Analysis Center, Data and Analysis Center for Software. – 2007.

24. Woody C. Exploring the Use of Metrics for Software Assurance (SEI Report CMU/SEI-2018-TN-004) / C. Woody, R.J. Ellison, Ch. Ryan. – Carnegie Mellon University, USA: Software Engineering Institute, 2020. https://doi.org/10.1184/R1/12366842.v1.

25. Платформа Рефлектиз [Электронный ресурс]. – URL: https://www.reflectiz.com/product/platform-overview (дата обращения: 21.07.2026).

26. State of Web Exposure 2026 [Электронный ресурс]. – URL: https://www.reflectiz.com/learning-hub/web-exposure-2026-research. (дата обращения: 22.07.2026).

27. Systematic literature review of machine learning based software development effort estimation models / S. Wen, Z. Li, Y. Lin, C. Hu, C. Huang // Information and Software Technology. – 2012. – Vol. 54, No. 1. – P. 41–59. DOI: 10.1016/j.infsof.2011.09.002.

28. Naeem S. Predicting Faults in High Assurance Software / S. Naeem, J. Khoshgoftaar, J. Van Hulse // 12th IEEE High Assurance Systems Engineering Symposium, 3–4 November 2010. – San Jose, CA, USA: IEEE, pp. 26–34. DOI: 10.1109/HASE.2010.29.

29. Hassan A.E. The top ten List: Dynamic Fault Prediction / A.E. Hassan, R.C. Holt // Proceedings of the 21st IEEE International Conference on Software Maintenance, 26–29 September 2005. – Budapest, Hungary: IEEE, 2005. – P. 263–272. DOI: 10.1109/ICSM.2005.91.

30. Kindervag J. (2010-11-05). Build Security Into Your Network's DNA: The Zero Trust Network Architecture. Forrester Research [Электронный ресурс]. – URL: https://www.virtualstarmedia.com/downloads/Forrester_zero_trust_DNA.pdf (дата обращения: 22.07.2026).

31. Témolé F. An Integrated Approach to Managing Software Quality in Complex Systems / F. Témolé, D. Ata-nasova // American Journal of Software Engineering and Applications. – 2025. – Vol. 13, No. 1. – P. 1–17. DOI: 10.11648/j.ajsea.20251301.11.

32. Champion K. Qualities of Quality: A Tertiary Review of Software Quality Measurement Research / K. Champion, S. Khatri and B. M. Hill // arXiv preprint. – 2021. – arXiv: 2107.13687. DOI: 10.48550/arXiv.2107.13687.

33. Деркач Е.А. Оценка нулевого доверия на этапах жизненного цикла создания программных продуктов / Е.А. Деркач, А.А. Шелупанов, С.Н. Новиков, Д.С. Брагин // Доклады ТУСУР. – 2025. – Т. 28, № 4. – С. 144–151. DOI: 10.21293/1818-0442-2025-28-4-144-151.

34. Баранов В.В. Методика и алгоритмы расчета защищенности элементов распределенных информационных систем в условиях деструктивного воздействия / В.В. Баранов, А.А. Шелупанов // Доклады ТУСУР. – 2022. – Т. 25, № 4. – С. 88–100. DOI: 10.21293/1818-0442-2022-25-4-88-100.

35. Деркач Е.А. Методологические подходы к обеспечению безопасности программных продуктов / Е.А. Деркач, А.А. Шелупанов // Доклады ТУСУР. – 2024. – Т. 27, № 1. – С. 37–43. DOI: 10.21293/1818-0442-2024-27-1-37-43.

Бочков Александр Владимирович
Д-р техн. наук, учёный секретарь,
АО «Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте» (АО «НИИАС»)
Нижегородская ул., 27, стр. 1, г. Москва, Россия, 109029
ORCID: 0000-0003-0211-8320
Тел.: +7-916-234-40-32
Эл. почта: a.bochkov@vniias.ru

Сабанов Алексей Геннадьевич
Д-р техн. наук, гл. эксперт АО «НИИАС»
Нижегородская ул., 27, стр. 1, г. Москва, Россия, 109029
ORCID: 0000-0002-2231-0240
Эл. почта: a.sabanov@vniias.ru
Тел.: +7-985-924-52-09

Шелупанов Александр Александрович
Д-р техн. наук, член-корр. РАН, проф., президент Томского государственного университета систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0003-2393-6701
Тел.: +7 (382-2) 55-55-55
Эл. почта: saa@fb.tusur.ru

Поступила в редакцию: 30.07.2026.
Принята к публикации: 31.08.2026.

Bochkov A.V., Sabanov A.G., Shelupanov A.A.

A Concept for Developing Security Metrics for Application Software in Critical Information Infrastructure Facilities

Relevance. The increasing regulatory and legal requirements for information protection in critical information infrastructure (CII) facilities necessitate the development of a system of quantitative measures for assessing the security of application software being developed and used. ***Purpose of the study.*** Development of an advanced method for quantitative security assessment of application software based on the methodology of the international standards for software quality. ***Methods.*** The study comprehensively employs methods of system analysis, formalization, synthesis, comparative analysis, computational experiment, and regulatory alignment. ***Scientific novelty.*** A concept has been developed for introducing a system of quantitative security assessments of application software and a new software security sub-characteristic, «software security levels», as a quality indicator for software under development. A three-level framework for establishing software security levels is proposed, based on a two-contour aggregation model in which blocking and gradable metrics play mathematically distinct functional roles. ***Results.*** The applicability of the proposed concept is demonstrated through a case study of calculating security metrics. ***Practical significance.*** The proposed system of quantitative assessments for refined security metrics can be applied in projects involving the development and modernization of application software for critical information infrastructure facilities.

Keywords: software security, critical information infrastructure, SQuaRE, security metrics, harmonic mean, blocking metrics, threat model, software security levels.
DOI: 10.21293/1818-0442-2026-29-2-160-176

References

1. GOST R 56939–2024. *Zashchita informatsii. Razrabotka bezopasnogo programmnogo obespecheniya. Obshchie trebovaniya* [Information protection. Development of secure software. General requirements]. Moscow, Standartinform, 2024 (in Russ.).

2. GOST R 12207–2010. *Informatsionnaya tekhnologiya. Sistemnaya i programmnaya inzheneriya. Protsessy zhiznennogo tsikla programmnykh sredstv* [Information technology. System and software engineering. Software life cycle processes]. Moscow, Standartinform, 2010 (in Russ.).

3. *Ob utverzhdenii Trebovanii po obespecheniyu bezopasnosti znachimykh ob'ektov KII Rossiiskoi Federatsii* [On the approval of Requirements for ensuring the security of significant objects of critical information infrastructure of the Russian Federation]. Order No. 239 of the FSTEC of Russia dated December 25, 2017 (in Russ.). Available at: URL: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239 (accessed: 09 October 2026).

4. *Metodicheskii dokument FSTEK Rossii. Sostav i soderzhanie meropriyatii i mer po zashite informacii, soderzhashikhsya v informatsionnykh sistemah* [Methodological document of the FSTEC of Russia. Composition and content of measures and activities to protect information contained in information systems]. Approved April 12, 2026 (in Russ.). Available at: https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-12-aprelya-2026-g (accessed: 09 October 2026).

5. Order of the FSTEC of Russia No. 117 dated 11.04.2025 "*On Approval of the Requirements for the Protection of Information Contained in State Information Systems and Other Information Systems of State Bodies, State Unitary Enterprises, and State Institutions*" (Registered with the Ministry of Justice of Russia on 16.06.2025 No. 82619) // Official Internet Portal of Legal Information. URL: http://publication.pravo.gov.ru/document/0001202506170011 (accessed: 09.10.2026).

6. GOST R ISO/IEC 9126–93. *Informatsionnaya tekhnologiya. Otsenka programmnoi produktsii. Kharakteris-*

tika kachestva i rukovodstva po ikh primeneniyu [Information technology. Evaluation of software products. Quality characteristics and guidelines for their use]. Introduced 1993-12-28. Moscow, «Izdatel'stvo standartov» Publ. house, 2003 (in Russ.).

7. GOST R ISO/IEC TO 12182–2002. *Informatsionnaya tekhnologiya. Klassifikatsiya programmnykh sredstv* [Information technology. Classification of software]. Introduced 2002-06-11. Moscow, «Izdatel'stvo standartov» Publ. house, 2002 (in Russ.).

8. GOST R ISO/IEC 25000–2021. *Sistemnaya i programmnaya inzheneriya. Trebovaniya i otsenka kachestva sistem i programmnykh sredstv* (SQuaRE). Rukovodstvo [System and software engineering. Systems and software quality requirements and evaluation (SQuaRE). Guide]. Introduced 2021-10-25. Moscow, Russian Institute of Standardization, 2021 (in Russ.).

9. GOST R ISO/IEC 25010-2015. *Sistemnaya i programmnaya inzheneriya. Trebovaniya i otsenka kachestva sistem i programmnogo obespecheniya (SQuaRE). Modeli kachestva sistem i programmnykh produktov* [System and software engineering. Systems and software quality requirements and evaluation (SQuaRE). System and software product quality models]. Introduced 2015-05-29. Moscow, Standartinform, 2015 (in Russ.).

10. GOST R ISO/IEC 25022-2021. *Sistemnaya i programmnaya inzheneriya. Trebovaniya i otsenka kachestva sistem i programmnogo obespecheniya (SQuaRE). Izmerenie kachestva pri ispol'zovanii* [System and software engineering. Systems and software quality requirements and evaluation (SQuaRE). Measurement of quality in use]. Introduced 2021-10-25. Moscow, Russian Institute of Standardization, 2021 (in Russ.).

11. GOST R ISO/IEC 25023–2021. Sistemnaya i programmnaya inzheneriya. Trebovaniya i otsenka kachestva sistem i programmnoi produktsii (SQuaRE). Izmereniya kachestva sistemy i programmnoi produktsii [System and software engineering. Systems and software quality requirements and evaluation (SQuaRE). Measurement of system and software product quality]. Introduced 2021-11-19. Moscow, Russian Institute of Standardization, 2021 (in Russ.).

12. GOST R 70921–2023. *Sistemnaya i programmnaya inzheneriya. Trebovaniya i otsenka kachestva sistem i programmnoi produktsii (SQuaRE). Kontseptsiya trebovanii k kachestvu* [System and software engineering. Systems and software quality requirements and evaluation (SQuaRE). Concept of quality requirements]. Introduced 2023-09-14. Moscow, Russian Institute of Standardization, 2023 (in Russ.).

13. GOST R 70262.1-2022. *Zashchita informatsii. Identifikatsiya i autentifikatsiya. Urovni doveriya k identifykatsii* [Information protection. Identification and authentication. Levels of trust in identification] Moscow, Russian Institute of Standardization, 2022 (in Russ.).

14. GOST R 70262.2-2025. *Zashchita informatsii. Identifikatsiya i autentifikatsiya. Urovni doveriya k autentifikatsii* [Information protection. Identification and authentication. Levels of trust in authentication] Moscow, Russian Institute of Standardization, 2025 (in Russ.).

15. Shubinskii, I.B. Funktsional'naya nadezhnost' informatsionnykh sistem. Metody analiza [Functional reliability of information systems. Methods of analysis]. Ulyanovsk, Regional Printing House «Pechatnyi dvor», 2012. 295 p. (in Russ.)

16. Shubinskii I.B. *Nadezhnye otkazoustoichivye informatsionnye sistemy. Metody sinteza* [Reliable fault-tolerant information systems. Methods of synthesis]. Ulyanovsk, Regional Printing House «Pechatnyi dvor», 2016, 544 p. (in Russ.).

17. Shubinskii I.B., Rozenberg E.N. *Funktsional'naya bezopasnost' sistem upravleniya na zheleznodorozhnom transporte* [Functional safety of control systems in railway transport]. Moscow, Vologda, Infra-Inzheneriya, 2023, 360 p. (in Russ.)

18. Markov A.S. *Modeli otsenki i planirovaniya ispytanii programmnykh sredstv po trebovaniyam bezopasnosti informatsii* [Models for evaluation and planning of software testing for information security requirements]. *Herald of the Bauman Moscow State Technical University. Series Instrument Engineering*, 2011, Special issue, pp. 90–103 (in Russ.).

19. Markov A.S., Fadin, A.A. *Sistematika uyazvimostei i defektov bezopasnosti programmnykh resursov* [Taxonomy of vulnerabilities and security defects of software resources]. Zaŝita informacii. Inside, 2013, No. 3, pp. 56–61 (in Russ.).

20. Markov A.S., Matveev V.A., Fadin, A.A., Tsirlov, V.L. [Heuristic analysis of source code security]. *Herald of the Bauman Moscow State Technical University. Series Instrument Engineering*, 2016, no. 1 (106), pp. 98–111. DOI: 10.18698/0236-3933-2016-1-98-111 (in Russ.)

21. Merzlyakova E.Yu., Yanchenko E.V. [Review of Software Quality Verification and Evaluation Methods]. *Vestnik SibGUTI*, vol. 17, 2023, no. 1, pp. 92–106. DOI: 10.55648/1998-6920-2023-17-1-92-106 (in Russ.).

22. Kulyamin V.V. *Tekhnologii programmirovaniya. Komponentnyy podkhod. Lektsiya 5. Kachestvo programmnogo obespecheniya i metody ego kontrolya* [Programming Technologies. Component-Based Approach. Lecture 5. Software quality and methods of its control] (in Russ.). Available: https://studfile.net/preview/993777/page:14/ (accessed: 22 July 2026).

23. Goertzel K.M., Winograd Th., McKinley H.L. et al. Software Security Assurance: A State-of-Art Report (SAR). Fort Belvoir, VA: Information Assurance Technology Analysis Center, Data and Analysis Center for Software, 2007.

24. Woody C., Ellison R.J., Ryan Ch. Exploring the Use of Metrics for Software Assurance (SEI Report CMU/SEI-2018-TN-004). Carnegie Mellon University, USA, Software Engineering Institute, 2020. https://doi.org/10.1184/R1/12366842.v1.

25. *Platforma Reflektiz* [Reflectiz Platform] (in Russ.). Available at: https://www.reflectiz.com/product/platform-overview, free (accessed: 21 July 2026).

26. State of Web Exposure 2026. URL: https://www.reflectiz.com/learning-hub/web-exposure-2026-research (accessed: 22 July 2026).

27. Wen S., Li Z., Lin Y., Hu C., Huang C. Systematic literature review of machine learning based software development effort estimation models. *Information and Software Technology*, 2012, 54, no. 1, pp. 41–59. DOI: 10.1016/j. infsof.2011.09.002.

28. Naeem S., Khoshgoftaar J., Van Hulse J. Predicting Faults in High Assurance Software. Proceedings of the 12th IEEE International Symposium on High Assurance Systems Engineering, November 3-4, 2010, San Jose, CA, USA, IEEE, pp. 26–34. DOI: 10.1109/HASE.2010.29.

29. Hassan A.E., Holt R.C. The top ten List: Dynamic Fault Prediction. Proceedings of the 21st IEEE International Conference on Software Maintenance, 26-29 September 2005, Budapest, Hungary, IEEE, 2005, pp. 263–272. DOI: 10.1109/ICSM.2005.91.

30. Kindervag J. (2010-11-05). Build Security Into Your Network's DNA: The Zero Trust Network Architecture. Forrester Research. Available at: https://www.virtual starmedia.com/downloads/Forrester_zero_trust_DNA.pdf (accessed: 22 July 2026).

31. Témolé F., Atanasova D. An Integrated Approach to Managing Software Quality in Complex Systems. *American Journal of Software Engineering and Applications*, 2025, vol. 13, no. 1, pp. 1–17. DOI: 10.11648/j.ajsea.20251301.11.

32. Champion K., Khatri S., Hill B.M. Qualities of Quality: A Tertiary Review of Software Quality Measurement Research. arXiv preprint, 2021, arXiv: 2107.13687. DOI: 10.48550/arXiv.2107.13687.

33. Derkach E.A., Shelupanov A.A., Novikov S.N., Bragin D.S. [Zero Trust assessment across the software product development lifecycle]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2025, vol. 28, no. 4, pp. 144–151. DOI: 10.21293/1818-0442-2025-28-4-144-151 (in Russ.).

34. Baranov V.V., Shelupanov A.A. [Cognitive model for assessing the security of information systems for various purposes]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, 2022, vol. 25, no. 4, pp. 88–100. DOI: 10.21293/1818-0442-2022-25-4-88-100 (in Russ.).

35. Derkach E.A., Shelupanov A.A. [Domestic and foreign experience in security testing]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radio-elektroniki*, 2024, vol. 27, no. 1, pp. 37–43. DOI: 10.21293/ 1818-0442-2024-27-1-37-43 (in Russ.).

Alexander V. Bochkov
Doctor of Technical Sciences, Scientific Secretary,
ЭJoint-Stock Company «Research and Design Institute
for Informatization, Automation and Communication
in Railway Transport» (JSC «NIIAS»)
27/1, Nizhegorodskaya St., Moscow, Russia, 109029
ORCID: 0000-0003-0211-8320
Phone: +7-916-234-40-32
Email: a.bochkov@vniias.ru

Alexey G. Sabanov
Doctor of Technical Sciences, Chief Expert, JSC «NIIAS»
27/1, Nizhegorodskaya St., Moscow, Russia, 109029
ORCID: 0000-0002-2231-0240
Phone: +7-985-924-52-09
Email: a.sabanov@vniias.ru

Alexander A. Shelupanov
Doctor of Technical Sciences,
Corresponding Member of the Russian Academy of Sciences
President, Tomsk State University
of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0003-2393-6701
Phone: +7 (382-2) 55-55-55
Email: saa@tusur.ru

Received: 30.07.2026.
Accepted: 31.08.2026.

ЭЛЕКТРОТЕХНИКА

В разделе статьи по следующим специальностям (технические науки)

2.4.1. Теоретическая и прикладная электротехника.
2.4.2. Электротехнические комплексы и системы.

УДК 621.314.04
Шифр специальности ВАК 2.4.1

Е.С. Полынцев, И.Ю. Кодорова, А.И. Бартенев, А.А. Аксенов, М.С. Воробьев, В.А. Кагадей

Силовая GaN-микросборка на основе AlN-DBC-платы с металлизированными переходными отверстиями

Актуальность. Благодаря высокой скорости переключения и низким потерям проводимости, GaN-транзисторы рассматриваются как наиболее перспективная электронная компонентная база в импульсных преобразователях электрической энергии. Однако их широкое внедрение сдерживается двумя ключевыми факторами: необходимостью минимизации паразитных индуктивностей и требованием к обеспечению эффективного теплоотвода. ***Цель исследования.*** Целью исследования является разработка и экспериментальная верификация силовой GaN-микросборки, в которой два GaN-транзистора, затворный драйвер и пассивные компоненты смонтированы на керамической AlN-DBC-плате с металлизированными переходными отверстиями, обеспечивающей низкие значения паразитной индуктивности межсоединений, а также теплового сопротивления микросборки. ***Методы.*** Схемотехническое моделирование электрических характеристик проводилось в среде LTSpice. Топологическое проектирование керамической платы выполнялось в Altium Designer. Электрическое и тепловое моделирование осуществлялось в системе инженерного анализа ANSYS Electronics Desktop: для экстракции паразитных индуктивностей межсоединений использовался модуль Q3D, а для расчета распределения температуры и определения теплового сопротивления сборки – модуль Icepak. Экспериментальная верификация включала тест двойным импульсом для оценки коммутационных перенапряжений и пиковых токов, а также инфракрасную термографию работающего образца. ***Научная новизна.*** Предложена оригинальная конструкция силовой GaN-микросборки на основе AlN-керамической платы с двусторонней медной металлизацией и переходными металлизированными отверстиями, характеризующаяся низкими величинами индуктивностей затворного и силового контуров GaN-транзисторов, а также теплового сопротивления микросборки. ***Результаты.*** Силовая GaN-микросборка имеет два независимых канала, каждый из которых обеспечивает ток до 7 А, рабочее напряжение не менее 100 В. Показано, что для данной конструкции микросборки индуктивность в затворном контуре обоих каналов имеет характерную величину 2,6 и 2,2 нГн, а добавочная паразитная индуктивность в силовой контур транзистора не превышает 0,5 нГн. Согласно результатам моделирования и экспериментальных исследований, перенапряжение на стоке транзистора не превышало 20 В при пиковом токе в индуктивной нагрузке 125 А, а тепловое сопротивление «активная область силового транзистора – окружающая среда» не превышает 36 К/Вт. ***Практическая значимость.*** GaN-микросборка предназначена для применения в составе высокоплотных низковольтных импульсных преобразователей, работающих на частотах выше 0,5 МГц, и позволяет поднять их удельную мощность.

Ключевые слова: силовая GaN-электроника, GaN-микросборка, экстракция паразитных параметров, DBC-керамика, металлизированные переходные отверстия, низкая паразитная индуктивность, гибридная интеграция.
DOI: 10.21293/1818-0442-2026-29-2-179-188

Введение

Силовые транзисторы на основе нитрида галлия (GaN) характеризуются улучшенным комплексом электрических параметров по сравнению с традиционными аналогами на основе кремния (Si) [1]. Ключевыми из них являются рекордно низкое сопротивление канала в открытом состоянии (единицы–десятки мОм), позволяющее существенно снизить статические потери на проводимость, а также экстремально малые времена переключения (единицы нс), обеспечивающие возможность работы на частотах до нескольких мегагерц с динамическими потерями, на порядок меньшими, чем у Si-приборов [2, 3]. Современные коммерчески доступные GaN-транзисторы и диоды Шоттки рассчитаны на рабочие напряжения от 40 до 650 В и токи до 150 А [4], при этом они демонстрируют высокую температурную стабильность характеристик и сохраняют работоспособность при температурах кристалла до 200 °С [5]. Однако столь выдающиеся параметры неизбежно сопровождаются повышенной удельной тепловой мощностью [6] и высокой чувствительностью к паразитным индуктивностям монтажа [7, 8], что предъявляет жёсткие требования к конструктивному и топологическому исполнению окружения силовых GaN-транзисторов. В этой связи научно-техническая задача по разработке новых конструкций силовых GaN-микросборок и модулей, обеспечивающих малые паразитные индуктивности и эффективный теплоотвод, является актуальной.

Из литературы [9, 10] известны конструкции, основанные на 3D-интеграции кристаллов силовых приборов в объем органической печатной платы. Такие конструкции обладают минимальной длиной проводников и, следовательно, низкой паразитной индуктивностью межсоединений (менее 1 нГн). Основным недостатком таких конструкций является значительное тепловое сопротивление, обусловленное низкой теплопроводностью материала печатной платы, которое приводит к интенсивному саморазогреву кристаллов силовых приборов в ходе их работы. Это ограничивает максимальную электриче-

скую мощность устройств на основе подобных конструкций.

Конструкции классических силовых модулей на основе кремниевых силовых приборов [11–13] основаны на применении высокотеплопроводных керамических подложек с двухсторонней металлизацией и проволочных межсоединений. Такие конструкции обладают низким тепловым сопротивлением, однако характеризуются повышенной паразитной индуктивностью силового контура $L_{\text{сил}}^{\text{пар}}$ (порядка 5–10 нГн), которая ограничивает КПД устройств на основе подобных конструкций на повышенных рабочих частотах.

Существуют также и гибридные конструкции, основанные на сочетании керамических металлизированных подложек и органических печатных плат [14–16], обеспечивающие как эффективный теплоотвод, так и низкие паразитные индуктивности межсоединений. В конструкциях подобного типа кристаллы силовых транзисторов монтируются на печатную плату, а керамическая плата крепится к тыльной стороне кристаллов для отвода тепла. Такая конструкция обеспечивает $L_{\text{сил}}^{\text{пар}}$ около 1 нГн, а также низкое тепловое сопротивление. Недостатком такого решения является невозможность организации эффективного двухстороннего теплоотвода вследствие низкой теплопроводности печатной платы.

Ведущий производитель силовых низковольтных GaN-транзисторов, компания EPC, рекомендует осуществлять монтаж кристаллов силовых транзисторов с бампами напрямую на поверхность печатных плат [17, 18]. Такой способ монтажа позволяет обеспечить малую величину $L_{\text{сил}}^{\text{пар}}$ порядка 0,4 нГн [17] за счет отказа от классического корпуса. Но из-за низкой теплопроводности печатных плат для этого способа характерны высокие значения теплового сопротивления [18].

В ранее опубликованных работах [19, 20] была предложена конструкция силовой GaN-микросборки, основанная на применении керамической платы с двусторонней медной металлизацией и переходными металлизированными отверстиями. Внешний вид данной конструкции в поперечном сечении представлен на рис. 1.

На лицевой стороне керамической платы методом перевернутого монтажа (flip-chip) [21] монтировались кристаллы силовых GaN-транзисторов, затворные драйверы и пассивные чип-компоненты. Обратной стороной керамическая плата монтировалась на поверхность органической печатной платы. Применение переходных металлизированных отверстий в керамической плате обеспечивало минимальную длину проводников, и как следствие малую величину паразитных индуктивностей в силовом контуре GaN-транзистора. Отвод тепла в данной конструкции может осуществляться в двух направлениях: через радиатор, расположенный на тыльной стороне кристалла силового GaN-транзистора, и через объем керамической платы, переходные тепловые отверстия в печатной плате на радиатор или основание корпуса.

На основании электрического и теплового моделирования, описанного в работах [19, 20], была показана перспективность новой конструкции силовых GaN-микросборок. Настоящая работа является продолжением ранее опубликованных работ, в которой будут представлены результаты моделирования и экспериментальной верификации силовой GaN-микросборки.

Методика

Проектирование силовой GaN-микросборки выполнялось с применением нескольких программных систем автоматизированного проектирования (САПР). Топология керамической платы проектировалась в Altium Designer. Разработка 3D-модели силовой GaN-микросборки осуществлялась в САПР для трехмерного моделирования FreeCAD. Электрическое и тепловое моделирование микросборки выполнялось методом конечных элементов в автоматизированной системе инженерного анализа ANSYS Electronics Desktop Student.

Электрическое моделирование силовой GaN-микросборки заключалось в экстракции паразитных параметров металлизации керамической платы и исследовании их влияния на динамические характеристики.

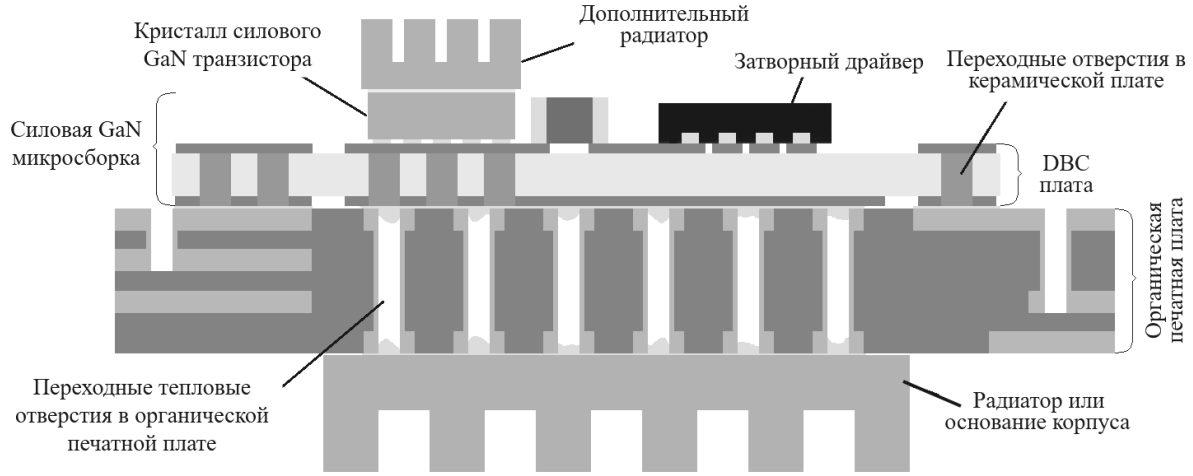


Рис. 1. Новая конструкция силовой GaN-микросборки

Экстракция паразитных параметров металлизации керамической платы выполнялась с использованием модуля Q3D. В качестве входных и выходных терминалов, через которые протекал электрический ток, использовались поверхности контактных площадок на керамической плате. При электрическом моделировании для используемых материалов задавались следующие физические параметры:

• удельная электрическая проводимость медных проводников керамической платы составляла 58×10^6 См/м;

• эквивалентная проводимость металлизированных отверстий в керамической плате составляла 34×10^6 См/м;

• диэлектрическая постоянная AlN-подложки составляла 8,5, а тангенс угла диэлектрических потерь 0,0003.

Исследование влияния паразитных параметров межсоединений на динамические характеристики микросборки выполнялось при помощи теста двойным импульсом [22]. Исследование проводилось на одном канале упрощенной имитационной модели микросборки в среде моделирования электронных схем LTSpice.

Имитационная модель канала *A* силовой GaN-микросборки, представленная на рис. 2, состояла из источника управляющих сигналов V_1, источника входного напряжения V_2, транзистора VT_1 EPC2016C, затворного резистора R_1 сопротивлением 20 Ом, совокупности паразитных индуктивностей металлизации керамической платы $L_1 - L_4$, паразитной индуктивности измерительной платы L_5 величиной 2 нГн, диода Шоттки VD_1 10L100SP, а также нагрузки, состоящей из резистивного шунта R_2 сопротивлением 0,1 Ом и индуктивности L_6 величиной 5 мкГн.

Из результатов электрического моделирования теста двойным импульсом канала *A* имитационной модели микросборки вычислялись величины коммутационных перенапряжений на затворе, $U_{\text{ЗИ}_{\text{сим}}^{\text{перенапр}}}$ и стоке транзистора, $U_{\text{СИ}_{\text{сим}}^{\text{перенапр}}}$ при максимальном входном напряжении и импульсном токе в индуктивной нагрузке, $I_{L_{\text{сим}}^{\text{макс}}}$.

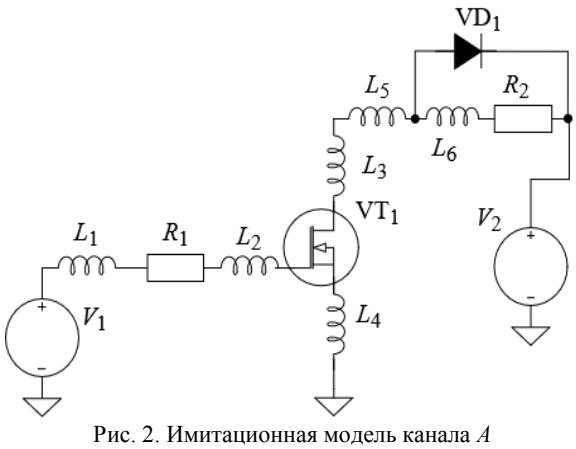


Рис. 2. Имитационная модель канала *A* силовой GaN-микросборки

Тепловое моделирование силовой GaN-микросборки заключалось в расчёте распределения температуры по объему микромодуля в момент достижения стационарного состояния при помощи модуля Icepak. При проведении теплового моделирования рассматривался одноканальный режим работы микросборки, при котором основным источником тепла являлся силовой GaN-транзистор, на котором рассеивался 1 Вт тепловой мощности. При тепловом моделировании для используемых материалов задавались тепловые характеристики, представленные в табл. 1.

В начальный момент времени температура силовой GaN-микросборки считалась равной температуре окружающей среды ($T_{\text{сред}} = 22$ °C). Значение скорости воздушного потока было задано 0,5 м/с, что соответствовало условиям естественной конвекции.

Эквивалентное тепловое сопротивление активная область силового транзистора – окружающая среда R_{th} микросборки рассчитывалось по формуле (1):

$$R_{\text{th}} = \frac{\left(T_{\text{крист}} - T_{\text{сред}}\right)}{P_{\text{расс}}}, \qquad (1)$$

где $T_{\text{крист}}$ – температура кристалла при достижении стационарного состояния; $T_{\text{сред}}$ – температура окружающей среды; $P_{\text{расс}}$ – рассеиваемая тепловая мощность.

Для сравнения с конструкций между собой, по формуле (2) рассчитывалась удельная величина теплового сопротивления между кристаллом силового транзистора и окружающей средой, нормированная на площадь кристалла $S_{\text{крист}}$:

$$R_{\text{th}}^{\text{уд}} = \frac{\left(T_{\text{крист}} - T_{\text{сред}}\right)\cdot S_{\text{крист}}}{P_{\text{расс}}}. \qquad (2)$$

Экспериментальная верификация силовой GaN-микросборки включала в себя измерение динамических электрических и тепловых характеристик. Принципиальная электрическая схема для проведения измерений приведена на рис. 3 и состоит из измерительной печатной платы, функционального генератора сигналов G_1 Rigol DG2102, источника питания V_1 GW Instek SPD-73606, нагрузки, а также осциллографа P_1 Rigol MSO8064.

Таблица 1

Перечень материалов и их тепловые характеристики

Структурный элемент	Материал	Теплопродность, Вт/(м·К)
Силовой GaN-транзистор	GaN/Si	150
Выводы силового GaN-транзистора	Олово	50
Припой силового GaN-транзистора	Kester SAC305	55
Верхний проводящий слой	Cu	400
Подложка	AlN	170
Нижний проводящий слой	Cu	400
Припой керамической платы	ПОС61	57
Переходные тепловые отверстия в печатной плате	Cu	400
Припой радиатора	ПОС61	57
Радиатор	Cu	400

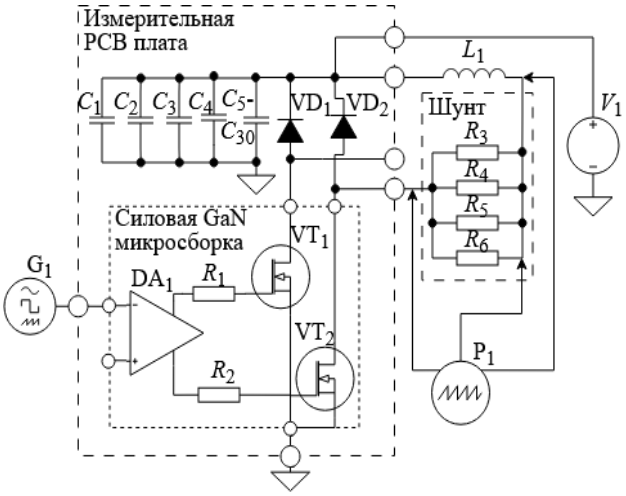


Рис. 3. Принципиальная электрическая схема для экспериментальной верификации силовой GaN-микросборки

На измерительной печатной плате размещались: силовая GaN-микросборка, магазин керамических конденсаторов (C_1 – C_{30}) с рабочим напряжением до 100 В, емкостью 10 мкФ и шунтирующие диоды VD_1 и VD_2 для каждого канала. В качестве шунтирующих диодов VD_1 и VD_2 был выбран диод Шоттки 10L100SP, имеющий следующие основные характеристики: максимальное обратное напряжение – 100 В, средний выпрямленный ток – 10 А, импульсный прямой ток – 200 А, прямое падение напряжения – 0,63 В при токе 10 А. Выбор данных диодов обусловлен их высокой скоростью переключения и низким прямым падением напряжения, что минимизирует влияние на коммутационные процессы. При измерении тепловых характеристик нагрузка представляла собой четыре параллельно соединенных резистора (R_3 – R_6) сопротивлением 1 Ом каждый, образующих измерительный шунт. При измерении динамических электрических характеристик нагрузка представляла собой измерительный резистивный шунт R_3 сопротивлением 0,1 Ом с последовательно соединенным дросселем L_1 индуктивностью 3 мкГн.

Измерение динамических электрических характеристик силовой GaN-микросборки заключалось в проведении теста двойным импульсом в одноканальном режиме. Из полученных экспериментальных данных вычислялись следующие динамические параметры:

• величина коммутационных перенапряжений на затворе, $U_{\text{ЗИ}_{\text{эксп}}^{\text{перенапр}}}$ и стоке, $U_{\text{СИ}_{\text{эксп}}^{\text{перенапр}}}$ транзистора при максимальном входном напряжении и импульсном токе в индуктивной нагрузке, $I_{L_{\text{эксп}}^{\text{макс}}}$;

• скорость нарастания импульса входного напряжения, $dU_{\text{СИ}}/dt$;

• скорость спада импульса управляющего напряжения, $dU_{\text{ЗИ}}/dt$.

Измерение тепловых характеристик заключалось в расчёте теплового сопротивления R_{th} микросборки по формуле (1) на основе результатов измерений температуры поверхности кристалла транзистора и выделяемой тепловой мощности. Измерения тепловой мощности проводились на частоте 10 кГц; ширина импульса определяла электрическую мощность, рассеиваемую на транзисторах. Температура поверхности кристалла измерялась при помощи тепловизора testo 875i. Коэффициент отражения для кристаллов силовых транзисторов составлял 0,77. Тепловая мощность представляла собой отношение полной мощности ко времени работы до достижения стационарного состояния.

Разработка силовой GaN-микросборки

Силовая GaN-микросборка разрабатывалась с учетом следующих технических требований:

• наличие двух выходных каналов;

• выходной ток в канале транзистора ≥ 7А;

• максимальное напряжение сток-исток силовых транзисторов ≥ 100 В;

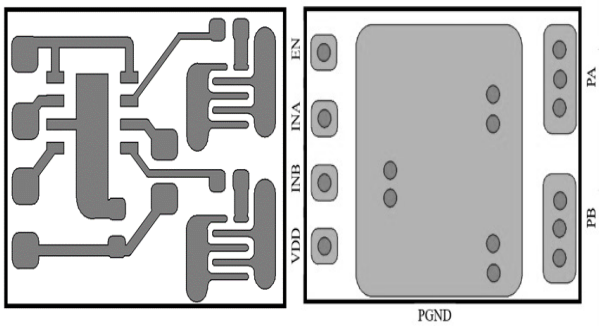
• сопротивление канала в открытом состоянии силового транзистора ≤ 16 мОм при 25 °С.

Схемотехническая реализация силовой GaN-микросборки детально описана в [19], она имела двухканальный затворный драйвер нижнего уровня UCC27524, который управлял двумя силовыми GaN-транзисторами EPC2016C. Силовая GaN-микросборка имела три силовых вывода PA, PB и PGND и четыре управляющих вывода EN, INA, INB и VDD.

Для изготовления керамической платы силовой GaN-микросборки использовалась подложка из AlN толщиной 380 мкм с двухсторонней медной DBC-металлизацией толщиной 200 мкм.

Технология DBC-металлизации обеспечивает высокую токовую и тепловую проводимость, а также хорошее рассеяние тепла через высокочистую медную металлизацию. Электрический контакт между металлизацией на лицевой и обратной сторонах керамической платы обеспечивался через глухие металлизированные переходные отверстия. Заполнение отверстий осуществлялось с помощью медного микрофиллера и серебросодержащей паяльной пасты WNB-217.

Внешний вид топологии лицевой и обратной сторон керамической платы приведен на рис. 4. Керамическая плата имела размер 10,6×8,2 мм.



а *б*

Рис. 4. Топология керамической платы силовой GaN-микросборки: *а* – топология лицевой стороны керамической платы; *б* – топология обратной стороны керамической платы

Изготовление силовой GaN-микросборки

Изготовление экспериментального образца силовой GaN-микросборки осуществлялось в три технологических этапа: 1) формирование топологии металлизации и отверстий в керамической плате; 2) металлизация отверстий; 3) монтаж компонентов на лицевую поверхность керамической платы.

Формирование топологии металлизации и глухих отверстий диаметром 500 мкм в керамической плате выполнялось методом лазерной абляции с использованием ЧПУ с пикосекундным лазером УФ-диапазона. Металлизация глухих отверстий в керамической плате, выполнялась с помощью медного микрофиллера диаметром 250 мкм и бессвинцовой серебросодержащей паяльной пасты WNB-217 с температурой плавления 217 °С, нанесенной методом дозирования.

На рис. 5 представлен общий вид изготовленной силовой GaN-микросборки.

Монтаж компонентов на лицевую поверхность керамической платы осуществлялся по технологии поверхностного и перевернутого монтажа с помощью паяльной пасты SAC305 с температурой плавления 220 °С.

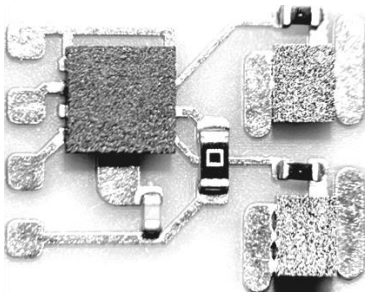


Рис. 5. Общий вид силовой GaN-микросборки

Исследование паразитных параметров

Наиболее критичными контурами силовой GaN-микросборки, в которых контролировались паразитные параметры, являлись затворный и силовой контуры. Затворный контур канала *A* образуется проводниками, соединяющими OUTA-терминал драйвера и, через резистор R_1, затворный терминал транзистора VT_1, а также обратный путь через PGND и формирующими индуктивности L_1 и L_2. Затворный контур канала *B* образуется проводниками, соединяющими OUTB-терминал драйвера, и через резистор R_2, затворный терминал транзистора VT_2, а также обратный путь через PGND и формирующими индуктивности L_3 и L_4.

Часть силового контура канала *A* образуется проводниками, соединяющими силовой терминал PA и стоковый терминал транзистора VT_1, а также истоковый терминал транзистора VT_1 и силовой терминал PGND и формирующими добавочные паразитные индуктивности L_5 и L_6. Часть силового контура канала *B* образуется проводниками, соединяющими силовой терминал PB и стоковый терминал транзистора VT_2, а также истоковый терминал транзистора VT_2 и силовой терминал PGND и формирующими добавочные паразитные индуктивности L_7 и L_8. Результаты экстракции приведены в табл. 2.

Таблица 2

Значения паразитных параметров керамической DBC-платы

Контур	Индуктивность	Индуктивность на постоянном токе, нГн		Индуктивность на частоте 1 МГц, нГн
Затворный контур канала *A*	L_1	2,4	2,6	1,4
	L_2	0,2		
Затворный контур канала *B*	L_3	2	2,2	1,15
	L_4	0,2		
Часть силового контура канала *A*	L_5	0,25	0,5	0,7
	L_6	0,25		
Часть силового контура канала *B*	L_7	0,25	0,5	0,3
	L_8	0,25		

По результатам моделирования, выполненного согласно описанной методике, суммарная паразитная индуктивность затворного контура составила 2,6 нГн для канала *A* и 2,2 нГн для канала *B*. Добавочная индуктивность к паразитной индуктивности силового контура каждого канала составила около 0,5 нГн.

На рис. 6 приведены результаты динамических измерений силовой GaN-микросборки с помощью теста двойным импульсом, полученные в результате моделирования, а также результаты, полученные экспериментально. Моделирование и эксперимент выполнялись по описанной ранее методологии со следующими параметрами: напряжение на входе микросборки – 80 В, амплитуда управляющих импульсов – 5,5 В, длительность первого и второго импульсов – 3 мкс, длительность паузы между импульсами – 3 мкс, время нарастания и спада фронта управляющего импульса – 9 нс.

Результаты моделирования. Согласно результатам моделирования теста двойным импульсом, при закрытии транзистора на стоке возникали коммутационные перенапряжения $U_{\text{СИ}_{\text{сим}}^{\text{перенапр}}}$, величина которых составляет 20 В при пиковом токе в индуктивности L_1 $I_{L_{\text{сим}}^{\text{макс}}}$ величиной 120 А. Они обусловлены наличием паразитных индуктивностей измерительной платы и добавочной паразитной индуктивности, вносимой элементами монтажа микросборки. При открытии транзистора на его затворе возникали коммутационные перенапряжения $U_{\text{ЗИ}_{\text{сим}}^{\text{перенапр}}}$ величиной 0,5 В, обусловленные наличием паразитных индуктивностей в затворном контуре канала *A* микросборки.

Экспериментальные результаты. Согласно результатам экспериментальных измерений, в индуктивности L_1 был достигнут пиковый ток $I_{L_{\text{эксп}}^{\text{макс}}}$ величиной 125 А. Из-за паразитных индуктивностей в измерительной плате и микросборке при открытии и закрытии транзистора возникали коммутационные перенапряжения $U_{\text{ЗИ}_{\text{эксп}}^{\text{перенапр}}}$ и $U_{\text{СИ}_{\text{эксп}}^{\text{перенапр}}}$. При величине $dU_{\text{ЗИ}}/dt$ на затворе транзистора в 0,56 В/нс коммутационное перенапряжение $U_{\text{ЗИ}_{\text{эксп}}^{\text{перенапр}}}$ достигало 0,6 В. На стоке транзистора при $dU_{\text{СИ}}/dt$ = 18 В/нс величина коммутационных перенапряжений $U_{\text{СИ}_{\text{эксп}}^{\text{перенапр}}}$ – 20 В.

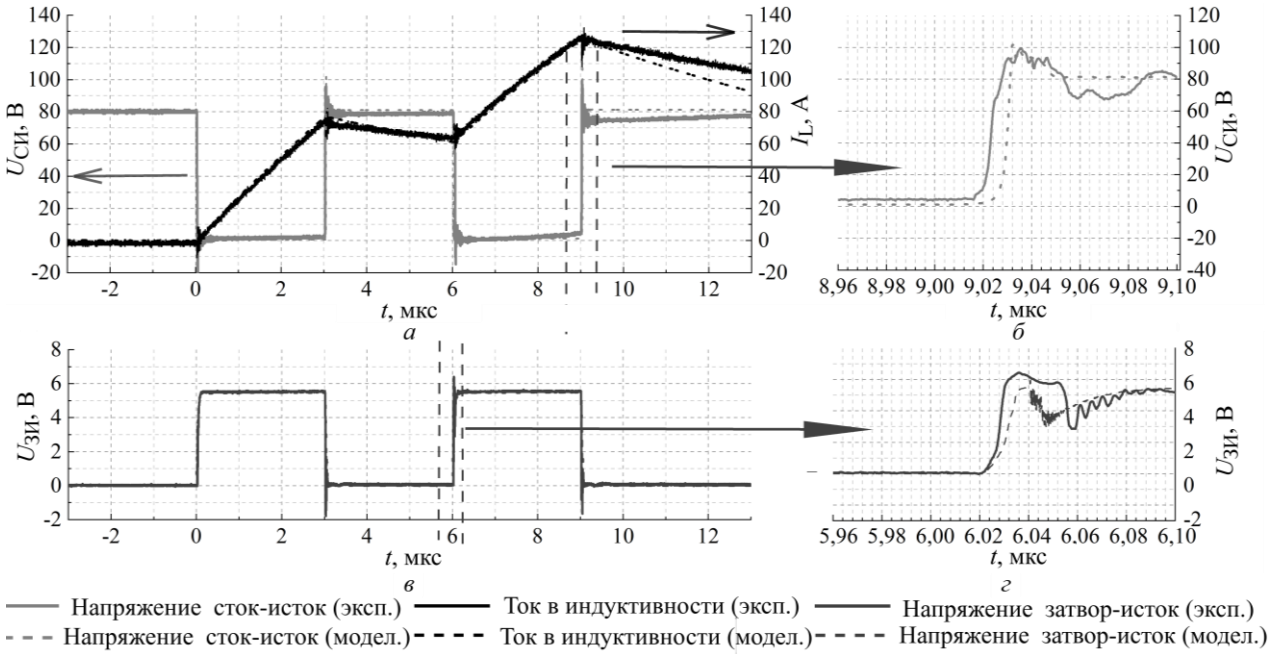


Рис. 6. Результаты моделирования и измерений силовой GaN-микросборки тестом двойным импульсом: *а* – напряжение сток-исток $U_{СИ}$ и ток в индуктивной нагрузке I_L; *б* – напряжение затвор-исток $U_{ЗИ}$; *в* – увеличенная форма напряжения $U_{СИ}$; *г* – увеличенная форма напряжения $U_{ЗИ}$

Экспериментальные значения перенапряжений и тока в индуктивной нагрузке соответствуют значениям, полученным в результате моделирования имитационной модели силовой GaN-микросборки.

Полученные экспериментальные и смоделированные значения коммутационных перенапряжений в силовой GaN-микросборке при максимальном пиковом токе в индуктивной нагрузке, вызванные паразитными индуктивностями межсоединений, не превышали значения максимального импульсного рабочего напряжения, что подтверждает ее работоспособность в режиме жесткой коммутации.

Исследование тепловых параметров

Распределение температуры по объему силовой GaN-микросборки, полученное в результате моделирования при достижении стационарного состояния, представлено на рис. 7.

В стационарном состоянии температура GaN-транзисторов не превышала 58 °С. Рассчитанное по формуле (1) эквивалентное тепловое сопротивление микросборки R_{th} составило 36 К/Вт. На рис. 8 приведена термограмма, отражающая распределение температуры по поверхности силовой GaN-микросборки в составе измерительной платы при достижении стационарного состояния.

Согласно термограмме силовой GaN-микросборки, полученной при помощи ИК-тепловизора, максимальная температура обратной стороны кристалла достигала значения 42 °С при средней тепловой мощности, рассеиваемой кристаллом равной 0,6 Вт. Таким образом, экспериментальное значение R_{th} микросборки составило 33 К/Вт, что достаточно точно согласуется со значением 36 К/Вт, полученным в результате моделирования.

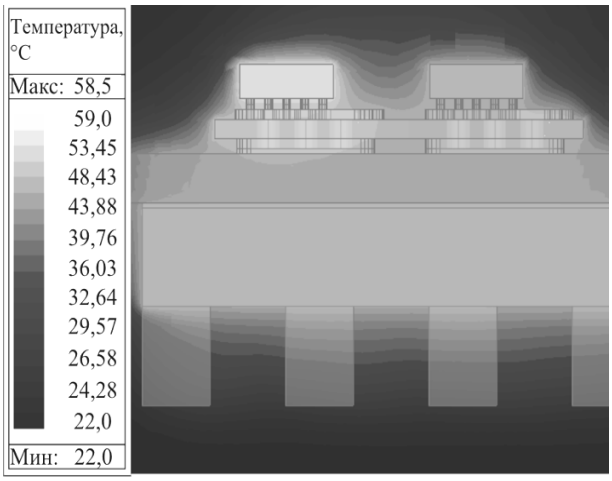


Рис. 7. Распределение температуры по объему силовой GaN-микросборки

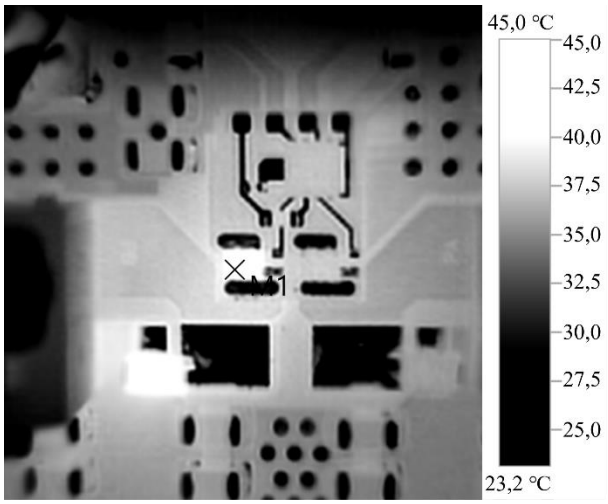


Рис. 8. Термограмма силовой GaN-микросборки

Сравнение с аналогами

На рис. 9 приведены результаты сравнения различных конструкций силовых GaN-микросборок и модулей по величинам удельного теплового сопротивления и паразитной индуктивности силового контура. Значения, приведенные на рис. 9, носят оценочный характер и были составлены на основе данных, опубликованных в работах [9–18], часть из которых были получены экспериментально, а часть – в результате моделирования, граничные условия которого также отличались. Расчет удельного теплового сопротивления $R_{th}^{уд}$ осуществлялся по формуле (2) в соответствии с методикой, описанной ранее.

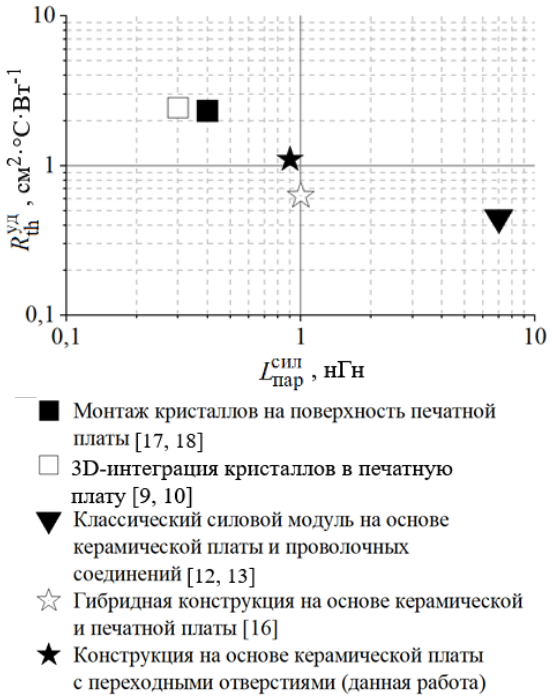


Рис. 9. Сравнительный анализ различных конструкций силовых микросборок и модулей по удельному тепловому сопротивлению и паразитной индуктивности силового контура [9–18]

В сравнении с конструкциями на основе керамических плат с проволочными межсоединениями [12, 13], где паразитная индуктивность силового контура достигает 5–10 нГн, разработанная новая конструкция микросборки, обеспечивала значительное снижение этого параметра до уровня, менее чем 1 нГн. Переходные металлизированные отверстия в керамической плате вносят дополнительную индуктивность около 0,5 нГн к базовому значению 0,4 нГн, характерному для монтажа на печатную плату [17, 18]; в результате суммарная индуктивность силового контура достигает 0,9 нГн, что сопоставимо с уровнем гибридных конструкций [14–16]. При этом новая конструкция обеспечивает более чем двукратное снижение теплового сопротивления по сравнению с аналогами на основе печатных плат [9, 10, 17, 18].

Таким образом, разработанная конструкция по величинам $L_{пар}^{сил}$ и $R_{th}^{уд}$ занимает промежуточное положение между классическими конструкциями силовых модулей [12, 13] на основе керамических плат с проволочными соединениями и конструкциями на основе печатных плат [9, 10, 17, 18], приближаясь к показателям, характерным для гибридных конструкций на основе сочетания керамических и печатных плат [16]. При этом, в отличие от данных гибридных конструкций, предложенная новая конструкция состоит из меньшего числа конструкционных элементов и имеет более простую технологию изготовления.

Заключение

Разработана, реализована и исследована оригинальная силовая GaN-микросборка на основе AlN-DBC-платы с металлизированными переходными отверстиями. Согласно результатам моделирования и измерений методом двойного импульса, перенапряжение на стоке транзистора не превышало 20 В при пиковом токе в индуктивной нагрузке 125 А, а добавочная паразитная индуктивность силового контура не превышала 0,5 нГн. Экспериментально измеренное значение теплового сопротивления R_{th} микросборки не превышало 33 К/Вт, что хорошо согласовывалось со значением 36 К/Вт, полученным в результате моделирования.

Результаты экспериментальных исследований и численного моделирования подтвердили, что разработанная силовая GaN-микросборка сочетает в себе низкую паразитную индуктивность, характерную для печатных плат, и эффективный двухсторонний теплоотвод, характерный для высокотеплопроводных керамических подложек. Это обеспечивает рекордные электрические и тепловые характеристики силовой GaN-микросборки по сравнению с уже известными техническими решениями. Силовые GaN-микросборки могут применяться в составе высокочастотных высокоплотных DC/DC-преобразователей и высокочастотных инверторах большой мощности.

Литература

1. Recent advances in GaN-based power HEMT devices / J. He, W.C. Cheng, Q. Wang, K. Cheng, H. Yu, Y. Chai // Advanced Electronic Materials. – 2021. – Vol. 7, No. 4. – P. 2001045.

2. Reliability, applications and challenges of GaN HEMT technology for modern power devices: A review / N. Islam, M.F.P. Mohamed, M.F.A.J. Khan, S. Falina, H. Kawarada, M. Syamsul // Crystals. – 2022. – Vol. 12, No. 11. – P. 1581.

3. Garg T. Recent developments, reliability issues, challenges and applications of GaN HEMT technology / T. Garg, S. Kale // IEEE Electron Devices Reviews. – 2024. – Vol. 1. – P. 16–30.

4. Udabe A. Gallium nitride power devices: A state of the art review / A. Udabe, I. Baraia-Etxaburu, D.G. Diez // IEEE Access. – 2023. – Vol. 11. – P. 48628–48650.

5. High-temperature electrical performances and physics-based analysis of p-GaN HEMT-device / S. Li, S. Liu, Y. Tian, C. Zhang, J. Wei, X. Tao, N. Li, L. Zhang, W. Sun // IET Power Electronics. – 2020. – Vol. 13, No. 3. – P. 420–425.

6. Thermal analysis of flip-chip bonding designs for GaN-power-HEMTs with an on-chip heat-spreading layer / K.B. Hong, C.Y. Peng, W.C. Lin, K.L. Chen, S.C. Chen, H.C. Kuo, E.Y. Chang, C.H. Lin // Micromachines. – 2023. – Vol. 14, No. 3. – P. 519.

7. Design of Low Parasitic Inductance GaN-HEMT-Flip-Chip Power Module / M.D. Rahman, T. Akter, A.S.M.K. Hasan, H.A. Mantooth, X. Song // 2025 IEEE Applied Power Electronics Conference and Exposition (APEC). – Atlanta, GA, USA: IEEE, 2025. – P. 1844–1849.

8. A highly integrated GaN power module with low parasitic inductance and high thermal performance / H. Kong, F. Yang, C. Yang, Y. Zhang, Z. Wang, Y. Yao, Y. Wang, L. Wang // 2022 IEEE Energy Conversion Congress and Exposition (ECCE). – Detroit, MI, USA: IEEE, 2022. – P. 1–7.

9. High power-density 3D integrated power supply module based on panel-level PCB embedded technology / F. Hou, X. Guo, Q. Wang, W. Wang, T. Lin, L. Cao, G.Q. Zhang, J.A. Ferreira // 2018 IEEE 68th Electronic Components and Technology Conference (ECTC). – San Diego, CA, USA: IEEE, 2018. – P. 1365–1370.

10. Mrad R. A PCB based package and 3D-assembly for high power density converters / R. Mrad, J. Morand, R. Perrin, S. Mollov // 2019 IEEE International Workshop on Integrated Power Packaging (IWIPP). – Toulouse, France: IEEE, 2019. – P. 73–77.

11. IGBT power module design for suppressing gate voltage spike at digital gate control / Z. Lou, T. Mamee, K. Hata, M. Takamiya, S.I. Nishizawa, W. Saito // IEEE Access. – 2023. – Vol. 11. – P. 6632–6640.

12. Automotive power module packaging: Current status and future trends / Y. Yang, L. Dorn-Gomba, R. Rodriguez, C. Mak, A. Emadi // IEEE Access. – 2020. – Vol. 8. – P. 160126–160144.

13. Смирнов В.И. Моделирование теплоэлектрических процессов в силовых модулях на MOSFET-транзисторах / В.И. Смирнов, А.М. Ходаков, А.А. Гавриков // РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии. – 2023. – Т. 15, № 2. – С. 117–124.

14. Implementation and switching behavior of a PCB-DBC-IGBT-module based on the power chip-on-chip 3-D-concept / J.L. Marchesini, P.O. Jeannin, Y. Avenas, J. Delaine, C. Buttay, R. Riva // IEEE Transactions on Industry Applications. – 2016. – Vol. 53, No. 1. – P. 362–370.

15. Embedded PCB and DBC-integrated design of 650 V GaN-power modules with enhanced thermal and electrical performance / A.Y. Choi, J. Jeon, J.H. Lim, S.W. Yoon // Microelectronics Reliability. – 2026. – Vol. 179. – P. 116053.

16. PCB-on-DBC GaN power module design with high-density integration and double-sided cooling / X. Tian, N. Jia, D. DeVoto, P. Paret, H. Bai, L.M. Tolbert, H. Cui // IEEE Transactions on Power Electronics. – 2023. – Vol. 39, No. 1. – P. 507–516.

17. Morganti T. PCB Layout Rules for Converters and Motor Drives // Efficient Power Conversion [Электронный ресурс]. – URL: https://epc-co.com/epc/Portals/0/epc/documents/presentations/pcb-layout-rules-04292025.pdf (дата обращения: 20.03.2026).

18. EPC2016C – eGaN Fet Datasheet // Efficient Power Conversion [Электронный ресурс]. – URL: https://epc-co.com/epc/portals/0/epc/documents/datasheets/EPC2016C_datasheet.pdf#page=1.00 (дата обращения: 20.03.2026).

19. Design and Optimization of 100 V GaN Multi-chip Power Micromodule Based on AlN DBC Substrate / E.S. Polyntsev, I.Y. Kodorova, A.A. Aksenov, V.A. Kagadey // 2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM). – Altai, Russian Federation: IEEE, 2024. – P. 1180–1184.

20. Kodorova I.Y. Thermal Analysis of Packaging Solution for GaN-Multi-chip-Power Micromodule / I.Y. Kodorova, E.S. Polyntsev, V.A. Kagadey // 2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM). – Altai, Russian Federation: IEEE, 2024. – P. 1230–1233.

21. Gan C.L. Advanced flip chip packaging / C.L. Gan, C.Y. Huang // Interconnect Reliability in Advanced Memory Device Packaging. – Cham: Springer International Publishing, 2023. – P. 67–94.

22. Levett D. Double pulse testing: The how, what and why / D. Levett, Z. Zheng, T. Frank // Infineon Technologies. – 2020. – P. 30–35 [Электронный ресурс]. – URL: https://ru.scribd.com/document/803783300/Infineon-Double-pulse-testing-Bodos-power-systems-Article-v01-00-EN (дата обращения: 20.03.2026).

Полынцев Егор Сергеевич
Аспирант каф. физической электроники (ФЭ)
Томского государственного университета
систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0000-0003-4368-8258
Тел.: +7 (382-2) 41-48-61
Эл. почта: egoleo@mail.ru

Кодорова Ирина Юрьевна
Инж.-электроник НПП «Радар ммс», НПК «Томское»
В. Высоцкого ул., 28/3, г. Томск, Россия, 634040
ORCID: 0009-0009-8521-8343
Тел.: +7 (382-2) 61-40-00
Эл. почта: irina_tusur@mail.ru

Бартенев Александр Иванович
Аспирант каф. промышленной электроники (ПрЭ) ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
ORCID: 0009-0001-2082-9313
Тел.: +7 (382-2) 41-48-61
Эл. почта: san4ubas@mail.ru

Аксенов Андрей Александрович
Аспирант Института сильноточной электроники
Сибирского отд. Российской академии наук (ИСЭ СО РАН)
Академический пр-т, 2/3, г. Томск, Россия, 634055
Тел.: +7-905-990-57-14
Эл. почта: aksenov@mail.tsu.ru

Воробьев Максим Сергеевич
Д-р. тех. наук. с.н.с.
Национального исследовательского
Томского политехнического университета (НИ ТПУ)
Ленина пр-т, 30, г. Томск, Россия, 634050
ORCID: 0000-0001-5136-5905
Тел.: +7-913-886-95-37
Эл. почта: vormas@tpu.ru

Кагадей Валерий Алексеевич
Д-р физ.-мат. наук, проф. каф. ФЭ ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 41-48-61
Эл. почта: vak@mail.tsu.ru

Поступила в редакцию: 29.04.2026.
Принята к публикации: 07.09.2026.

Polyntsev E.S., Kodorova I.Y., Bartenev A.I.,
Aksenov A.A., Vorobyov M.S., Kagadey V.A.

GaN-power micromodule based on an AlN-DBC-substrate with metallized vias

Relevance. The high switching speed and low conduction losses of GaN-transistors make them one of the most promising electronic component technologies for switched-mode power converters. However, their widespread adoption is constrained by two key factors: the need to minimize parasitic interconnection inductances and the requirement to ensure efficient heat dissipation. ***Purpose of the study.*** The study aims to develop and experimentally validate a GaN power micromodule in which two GaN transistors, a gate driver, and passive components are mounted on an AlN direct-bonded copper (DBC) substrate with metallized vias, providing low parasitic inductance of the interconnections and low thermal resistance of the micromodule. ***Methods.*** Circuit-level simulation of electrical characteristics was performed using LTspice. The ceramic substrate layout was designed in Altium Designer. Electrical and thermal modeling were performed in the Ansys Electronics Desktop engineering analysis platform. The Q3D module was used to extract the parasitic inductances of the interconnects, and the Icepak module was employed to calculate the temperature distribution and determine the thermal resistance of the module. Experimental validation included double-pulse testing to evaluate switching overvoltages and peak currents, along with infrared thermography of the operating prototype. ***Scientific Novelty.*** A novel GaN-power-micromodule based on an AlN ceramic substrate with double-sided copper metallization and plated vias is proposed. The proposed design is characterized by reduced gate-loop and power-loop parasitic inductances of the GaN transistors, as well as lower thermal resistance of the micromodule. ***Results.*** The presented micromodule incorporates two independent channels, each providing a current of up to 7 A, and an operating voltage of at least 100 V. It is shown that for the proposed micromodule design, the inductance in the gate loop of both channels has a characteristic value of 2.6 nH and 2.2 nH, while the additional parasitic inductance in the power loop of the transistor does not exceed 0.5 nH. According to the simulation and experimental results, the transistor drain overvoltage did not exceed 20 V at a peak current of 125 A in the inductive load. The junction-to-ambient thermal resistance was verified by simulation and experimental results achieving 36 K/W. ***Practical significance.*** The proposed GaN power micromodule is intended for use in high-density low‑voltage switched‑mode DC/DC-converters operating at frequencies above 0.5 MHz, enabling an increase in their power density.

Keywords: power GaN-electronics, GaN-multichip micromodule, parasitic parameters extraction, DBC-ceramic substrate, metallized vias, low parasitic inductance, hybrid integration.
DOI: 10.21293/1818-0442-2026-29-2-179-188

References

1. He J., Cheng W.C., Wang Q., Cheng K., Yu H., Chai Y. Recent advances in GaN-based power HEMT devices. *Advanced Electronic Materials*, 2021, vol. 7, no. 4, pp. 2001045.

2. Islam N., Mohamed M.F.P., Khan M.F.A.J., Falina S., Kawarada H., Syamsul M. Reliability, applications and challenges of GaN-HEMT-technology for modern power devices: A review. *Crystals*, 2022, vol. 12, no. 11, pp. 1581.

3. Garg T., Kale S. Recent developments, reliability issues, challenges and applications of GaN-HEMT-technology. *IEEE Electron Devices Reviews*, 2024, vol. 1, pp. 16–30.

4. Udabe A., Baraia-Etxaburu I., Diez D.G. Gallium nitride power devices: A state of the art review. *IEEE Access*, 2023, vol. 11, pp. 48628–48650.

5. Li S., Liu S., Tian Y., Zhang C., Wei J., Tao X., Li N., Zhang L., Sun W. High-temperature electrical performances and physics-based analysis of p-GaN-HEMT-device. *IET Power Electronics*, 2020, vol. 13, no. 3, pp. 420–425.

6. Hong K.B., Peng C.Y., Lin W.C., Chen K.L., Chen S.C., Kuo H.C., Chang E.Y., Lin C.H.Thermal analysis of flip-chip bonding designs for GaN-power-HEMTs with an on-chip heat-spreading layer. *Micromachines*, 2023, vol. 14, no. 3, pp. 519.

7. Rahman M.D., Akter T., Hasan A.S.M.K., Mantooth H.A., Song X. Design of Low Parasitic Inductance GaN-HEMT Flip-Chip Power Module. 2025 IEEE Applied Power Electronics Conference and Exposition (APEC), Atlanta, GA, USA, IEEE, 2025, pp. 1844–1849.

8. Kong H., Yang F., Yang C., Zhang Y., Wang Z., Yao Y., Wang Y., Wang L.A highly integrated GaN power module with low parasitic inductance and high thermal performance. 2022 IEEE Energy Conversion Congress and Exposition (ECCE), Detroit, MI, USA, IEEE, 2022, pp. 1–7.

9. Hou F., Guo X., Wang Q., Wang W., Lin T., Cao L., Zhang G.Q., Ferreira J.A. High power-density 3D-integrated power supply module based on panel-level PCB embedded technology. 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), San Diego, CA, USA, IEEE, 2018, pp. 1365–1370.

10. Mrad R., Morand J., Perrin R., Mollov S. A PCB based package and 3D assembly for high power density converters. 2019 IEEE International Workshop on Integrated Power Packaging (IWIPP), Toulouse, France, IEEE, 2019, pp. 73–77.

11. Lou Z., Mamee T., Hata K., Takamiya M., Nishizawa S.I., Saito W. IGBT power module design for suppressing gate voltage spike at digital gate control. *IEEE Access*, 2023, vol. 11, pp. 6632–6640.

12. Yang Y., Dorn-Gomba L., Rodriguez R., Mak C., Emadi A. Automotive power module packaging: Current status and future trends. *IEEE Access*, 2020, vol. 8, pp. 160126–160144.

13. Smirnov V.I. Khodakov A.M. Gavrikov A.A. [Modeling of thermoelectrical processes in a power MOSFET modules]. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2023, Vol. 15, no. 2. pp. 117–124 (in Russ.).

14. Marchesini J.L., Jeannin P.O., Avenas Y., Delaine J., Buttay C., Riva R. Implementation and switching behavior of a PCB-DBC IGBT module based on the power chip-on-chip 3D-concept. *IEEE Transactions on Industry Applications*, 2016, vol. 53, no. 1, pp. 362–370.

15. Choi A.Y., Jeon J., Lim J.H., Yoon S.W. Embedded PCB- and DBC-integrated design of 650 V GaN power modules with enhanced thermal and electrical performance. *Microelectronics Reliability*, 2026, vol. 179, pp. 116053.

16. Tian X., Jia N., DeVoto D., Paret P., Bai H., Tolbert L.M., Cui H. PCB-on-DBC GaN power module design with high-density integration and double-sided cooling. *IEEE Transactions on Power Electronics*, 2023, vol. 39, no. 1, pp. 507–516.

17. Morganti T. PCB Layout Rules for Converters and Motor Drives. Efficient Power Conversion. Available at: https://epc-co.com/epc/Portals/0/epc/documents/presentations/pcb-layout-rules-04292025.pdf (accessed: 20 March 2026).

18. EPC2016C – eGaN FET DATASHEET. Efficient Power Conversion. Available at: https://epc-co.com/epc/portals/0/epc/documents/datasheets/EPC2016C_datasheet.pdf#page=1.00 (acces-sed: 20 March 2026).

19. Polyntsev E.S., Kodorova I.Y., Aksenov A.A., Kagadey V.A. Design and Optimization of 100 V GaN Multi-chip Power-Micromodule Based on AlN DBC Substrate. 2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM), Altai, Russian Federation, IEEE, 2024, pp. 1180–1184.

20. Kodorova I.Y., Polyntsev E.S., Kagadey V.A. Thermal Analysis of Packaging Solution for GaN Multi-chip Power Micromodule. 2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM), Altai, Russian Federation, IEEE, 2024, pp. 1230–1233.

21. Gan C.L., Huang C.Y. Advanced flip chip packaging. *Interconnect Reliability in Advanced Memory Device Packaging*. Cham: Springer International Publishing, 2023, pp. 67–94.

22. Double pulse testing: The how, what and why / D. Levett, Z. Zheng, T. Frank. – Infineon Technologies, 2020, pp. 30–35. Available at: https://ru.scribd.com/document/803783300/Infineon-Double-pulse-testing-Bodos-power-systems-Article-v01-00-EN (accessed: 20 March 2026).

Egor S. Polyntsev
Postgraduate student, Department of Physical Electronics (PE), Tomsk State University
of Control Systems and Radioelectronics (TUSUR)
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0003-4368-8258
Phone: +7 (382-2) 41-48-61
Email: egoleo@mail.ru

Irina Y. Kodorova
Electronics Engineer JSC «Radar MMS»
28/3, V. Vysotsky Str., Tomsk, Russia, 634040
ORCID: 0009-0009-8521-8343
Phone: +7 (382-2) 61-40-00
Email: irina_tusur@mail.ru

Aleksandr I. Bartenev
Postgraduate student,
Department of Industrial Electronics (IE), TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0009-0001-2082-9313
Phone: +7 (382-2) 41-48-61
Email: san4ubas@mail.ru

Andrey A. Aksenov
Postgraduate student, Institute of High Current Electronics,
Siberian Branch of the Russian Academy of Sciences
2/3, Akademichesky Ave., Tomsk, Russia, 634055
Phone: +7-905-990-57-14
Email: aksenov@mail.tsu.ru

Maxim S. Vorobyov
Doctor of Technical Sciences, Senior Researcher
National Research Tomsk Polytechnic University
30, Lenin Ave., Tomsk, Russia, 634050
ORCID: 0000-0001-5136-5905
Phone: +7 (913) 886-95-37
Email: vormas@tpu.ru

Valery A. Kagadey
Doctor of Physical and Mathematical Sciences, Professor,
Department of Physical Electronics (PE), TUSUR
40, Lenin Ave., Tomsk, Russia, 634050
Phone: +7 (382-2) 41-48-61
Email: vak@mail.tsu.ru

Received: 29.04.2026.
Accepted: 07.09.2026.

УДК 621.3.011.7
Шифр специальности ВАК: 2.4.1

Ю.А. Бурцев

Развитие методов расчета переходных процессов на основе матричной экспоненты для цепей с нелинейными резисторами

Актуальность. Разработка новых методов расчета переходных процессов в электрических цепях актуальна, несмотря на наличие множества известных методов и компьютерных программ. Причина в том, что практические задачи накладывают набор требований, которым существующие методы иногда не вполне удовлетворяют, поэтому дают большую погрешность или даже качественно неверные результаты либо требуют большого времени счета, оказываются очень сложными или совсем неприменимыми. Методы должны быть: пригодными для жестких, колебательных и жестко-колебательных цепей с большим разбросом параметров; высокоточными (порядок точности не менее 3); экономичными по вычислениям; *L*-устойчивыми (это устраняет ложные колебания и ограничения на шаг интегрирования); приспособленными для расчета нелинейных цепей и цепей с частыми коммутациями; простыми и удобными, не требовать настройки, предварительных расчетов и оценок. ***Цель исследования.*** Расширить ранее предложенное автором семейство методов расчета переходных процессов в линейных электрических цепях на цепи с нелинейными резисторами. ***Методы.*** В основе представленной работы лежит новое семейство высокоточных неявных численных методов расчета переходных процессов в линейных электрических цепях. Новые методы основаны на разложении аппроксимации Паде матричной экспоненты на простейшие дроби и удовлетворяют всем перечисленным выше требованиям. Они очень хорошо приспособлены к структуре систем уравнений электрических цепей, значительно проще других высокоточных методов. Для решения нелинейных уравнений используются метод простой итерации и интерполяция тока внутри интервала интегрирования. ***Научная новизна.*** Новые эффективные методы расчета переходных процессов в линейных цепях удалось приспособить к расчету цепей с нелинейными резисторами. Предложены оригинальная, повышенной точности линеаризация вольт-амперной характеристики резистора и способ расчета токов внутри интервала интегрирования. Все результаты получены лично автором. ***Результаты.*** Новые методы позволяют рассчитывать произвольные режимы цепей, содержащих нелинейные резисторы, с погрешностью на несколько порядков меньшей, чем у одного из наиболее часто применяемых для этого метода трапеций. ***Практическая значимость.*** Расчеты динамических процессов в цепях с нелинейными элементами имеют очень много практических приложений. В частности, они нужны для проектирования силовых полупроводниковых инверторов, цифровой техники, систем электроснабжения, электронных и радиотехнических устройств и компонентов.

Ключевые слова: электрические цепи, нелинейные резисторы, переходные процессы, *L*-устойчивые методы, высокоточные методы, матричная экспонента, аппроксимация Паде.
DOI: 10.21293/1818-0442-2026-29-2-189-198

Введение

Основные задачи расчета электрических цепей появились и были так или иначе решены десятки лет назад. Однако, во-первых, многие методы расчета и программы показывают большие различия расчетов с реальностью [1–3] и имеют другие перечисленные ниже недостатки. Во-вторых, существующие методы и программы не могут покрыть все задачи проектирования новой техники. В настоящей статье ранее разработанные автором методы для линейных цепей расширены на цепи с нелинейными резисторами. При этом используется оригинальная линеаризация вольт-амперной характеристики (ВАХ) резистора, отличающаяся повышенной точностью.

В исследованиях на тему, представленную в статье, в большинстве случаев используют термин «переходные процессы», что отражено и в названии статьи. Однако полезно иметь в виду, что на самом деле решается более широкая задача, так как если в цепи действуют периодические источники ЭДС и тока и к тому же заданы специальные начальные условия либо переходной процесс в классическом его понимании закончился, мы получим расчет периодического режима электрической цепи, как это далее сделано в простейших тестовых задачах.

Отметим, что для широкополосной по частотному спектру нелинейной цепи расчет периодического процесса часто бывает трудно получить частотными методами. Если в цепи действуют непериодические источники энергии или имеют место автоколебания, то, применяя методы, формально предназначенные для расчета переходных процессов, можно рассчитывать непериодические и автоколебательные процессы. Таким образом, по существу в статье идет речь о расчете произвольных режимов электрических цепей. Сделав такую оговорку и следуя традиции, далее все же будем применять термин «переходные процессы», понимая его в широком смысле без дальнейших уточнений.

В [4] предложено семейство новых высокоточных *L*-устойчивых [5] численных методов для расчета переходных процессов в линейных электрических цепях с кусочно-постоянными параметрами и с учетом вариации заданных функций источников внутри шагов интегрирования. Основа методов – формула Коши [6] для решения системы обыкновенных дифференциальных уравнений (СОДУ), в которой матричная экспонента [6] представлена в виде разложения аппроксимации Паде [7] на сумму простейших дробей [8]. *L*-устойчивые методы не имеют

ограничений на шаг интегрирования и подавляют паразитные колебания [5].

Метод ε-вложения [5] позволил преобразовать формулы для решения СОДУ к форме, пригодной для непосредственного решения дифференциально-алгебраических систем уравнений (ДАСУ) электрических цепей. Матрицы в этих формулах легко формируются прямо по схеме электрической цепи.

Это важное отличие новых методов от других высокоточных методов, для которых необходим непростой переход от исходной ДАСУ электрической цепи к СОДУ, либо нужны другие сложные действия. Второе отличие – экономичность вычислений. Новые методы требуют в несколько раз меньше арифметических операций, чем аналогичные по точности и устойчивости методы типа Рунге–Кутты [8]. Они годятся для обычных цепей и для цепей с жесткими и колебательными системами уравнений [4, 8].

Наиболее популярные методы – метод трапеций и метод на основе формул дифференцирования назад (ФДН) 2-го порядка точности учитывают поведение свободного члена только в узлах временной сетки. Метод трапеций имеет склонность к раскачке паразитных колебаний [2, 3], метод на основе ФДН дает существенное методическое затухание решения. У явных методов и методов на основе ФДН с порядком точности больше 2 возможны ограничения на шаг интегрирования. Методы на основе матричной экспоненты свободны от перечисленных недостатков. Они могут иметь порядок точности 3 и более, они одношаговые и одностадийные, учитывают вариации свободного члена внутри шага интегрирования.

В работе [2] предложен гибридный метод на основе неявного метода Эйлера и метода трапеций. Метод расчета на основе ФДН и метода Ньютона–Рафсона для динамических цепей с диодами показан в [9]. В [10] рассмотрен метод Эйлера с адаптацией шага интегрирования. Описанный в [11, 12] способ на основе матричной экспоненты годится лишь для простых линейных цепей. В работе [13] предложен метод на основе преобразования Лапласа и *z*-преобразования для расчета переходных процессов в длинных линиях. Существуют также различные высокоточные методы расчета электрических цепей, например, разработанные в [1, 14–16].

В [17] разработан численный метод решения нелинейных СОДУ в форме Коши. Подробный обзор современных методов для нелинейных цепей представлен в [18, 19]. Общий недостаток известных высокоточных методов – сложность теории и алгоритмов, ведущая к большим объемам вычислений при изменении параметров цепи. Это представляет проблему при расчетах нелинейных цепей и требует от инженеров глубоких знаний математики.

Переходные процессы играют важную роль в силовых импульсных устройствах, в цифровой технике, в электронике, радиотехнике и энергетике – фактически в любых электротехнических изделиях. Поэтому создание простых высокоточных методов для их расчета является актуальной задачей.

Цель представленного исследования состоит в том, чтобы приспособить семейство новых численных методов на основе аппроксимации Паде матричной экспоненты [4] к расчетам переходных процессов в цепях с нелинейными резисторами.

Линеаризация ВАХ нелинейного резистора

Наиболее популярным методом расчета нелинейных электрических цепей является метод Ньютона–Рафсона. Однако он требует дифференцирования ВАХ и имеет высокую чувствительность к начальному приближению, поэтому он часто дает сбои. В нем используется локальная линеаризация ВАХ вида $u(i) = u(i_0)i_0 + R_\mathrm{d}(i_0)(i - i_0)$, здесь R_d – дифференциальное сопротивление. Погрешность такого уравнения, как правило, растет от начала к концу шага интегрирования (рис. 1). Уменьшить ее можно, но это требует непростых вычислений [15].

Другой метод локальной линеаризации основан на замене нелинейного резистора статическим сопротивлением, его уравнения имеют вид $u(i) = R_\mathrm{st}(i_0)i$ или $u(i) = R_\mathrm{st}(i_0 + \Delta i))i$, здесь R_st – статическое сопротивление. Метод на основе статических сопротивлений обычно сходится медленнее, чем метод Ньютона, но реже дает отказы. Погрешность такой линеаризации также равна нулю только в одной точке шага интегрирования (см. рис. 1). Метод статических сопротивлений исследован в [20]. Идея предлагаемого метода линеаризации состоит в том, чтобы точно представить напряжение нелинейного резистора в пределах шага интегрирования в виде

$$u(i(t)) = R_\Delta i(t) + E_R(i(t)),$$
$$R_\Delta = (u(i_0 + \Delta i) - u(i_0))/\Delta i, \qquad (1)$$

где R_Δ – конечно-разностное приближение производной (дифференциального сопротивления), $E_R(i(t))$ – добавочная функция (рис. 2). Далее R_Δ включается в матрицу **H** представленной ниже формулы расчета переходных процессов, а функция $E_R(i(t))$ в виде $E_T(t) = u(i(t)) - R_\Delta i(t)$ – в свободный член **f** этой формулы. При этом методы на основе матричной экспоненты позволяют учитывать динамику свободного члена в пределах шага интегрирования. Такая идея в несколько ином виде используется в [15].

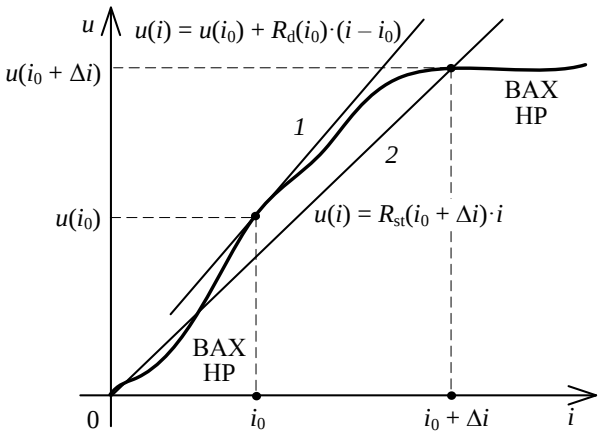


Рис. 1. Линеаризация ВАХ нелинейного резистора (ВАХ НР) с помощью дифференциального сопротивления (*1*) и с помощью статического сопротивления (*2*)

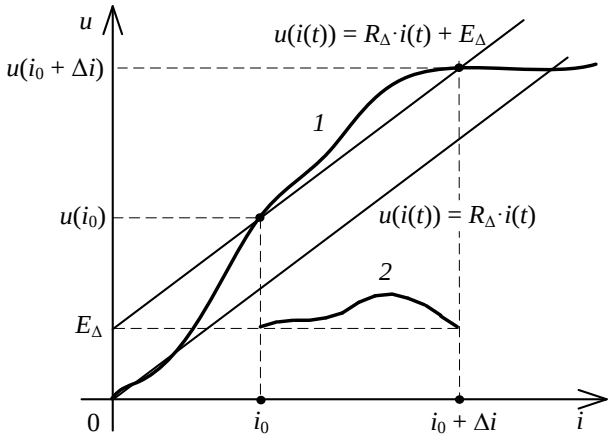


Рис. 2. Линеаризация ВАХ нелинейного резистора (ВАХ НР) с помощью конечно-разностного представления дифференциального сопротивления и добавочной функции: *1* – ВАХ НР, *2* – $E_R(i(t))$

В простейшем случае функции $E_R(i(t))$ и $E_T(t)$ могут быть приближенно представлены константой E_Δ (см. рис. 2). Такое представление годится в том числе для метода Эйлера, метода трапеций и других методов расчета переходных процессов, использующих значения решения и свободного члена системы уравнений только в узлах временной сетки. Однако это упрощение снижает точность расчетов.

Чтобы найти функцию $E_T(t)$, нужно знать зависимость $i(t)$ в пределах шага интегрирования. Положим $t = 0$ в начале шага и $t = \Delta t$ в конце шага. Начальное значение $i_0 = i(0)$ известно. Значение $i_0 + \Delta i = i(\Delta t)$ уточняем с помощью последовательных приближений.

Для первого приближения используем сопротивление R_Δ и функцию $E_T(t)$ с предыдущего шага интегрирования, рассчитываем состояние цепи в момент $t = \Delta t$, находим $i(\Delta t)$. Далее получаем новое значение R_Δ по (1) и вычисляем функцию $i(t)$ в пределах шага интегрирования линейной интерполяцией по значениям $i(0)$ и $i(\Delta t)$ либо параболической интерполяцией по значениям $i(-\Delta t)$, $i(0)$ и $i(\Delta t)$, либо кубической интерполяцией по значениям $i(-2\Delta t)$, $i(-\Delta t)$, $i(0)$ и $i(\Delta t)$. Полиномы более высокой степени не дают уменьшения погрешности, по-видимому, из-за раскачки значений между точками коллокации. На каждой итерации имеем очередное приближение значения R_Δ и функции $E_T(t)$, которые подставляем в формулу для расчета состояния цепи в момент $t = \Delta t$.

Расчет переходных процессов

Поясним суть метода для расчета переходных процессов в линейных электрических цепях [4], на основе которого возможен расчет процессов в цепях с нелинейными резисторами. Точнее, это семейство методов, из которого в этом примере мы используем вариант 3-го порядка точности на основе аппроксимации Паде матричной экспоненты R_{12}. Уравнения цепи запишем в виде

$$\mathbf{G}\frac{d\mathbf{x}(t)}{dt} = \mathbf{H}\mathbf{x}(t) + \mathbf{f}(t),$$

где **G** – матрица, содержащая индуктивности катушек и емкости конденсаторов, **H** – матрица с коэффициентами алгебраических уравнений цепи и правых частей дифференциальных уравнений, $\mathbf{x}(t)$ – вектор состояния цепи, содержащий токи ветвей и потенциалы узлов, $\mathbf{f}(t)$ – вектор-функция с источниками напряжения и тока. Покажем структуру матриц на простейшем примере цепи, изображенной на рис. 3.

Для этой цепи

$$\mathbf{G} = \begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & C \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -R & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix},$$

$$\mathbf{x}(t) = \begin{bmatrix} i_1(t) \\ i_2(t) \\ i_3(t) \\ \phi_1(t) \end{bmatrix}, \qquad \mathbf{f}(t) = \begin{bmatrix} E_1(t) \\ E_2(t) \\ -C\dfrac{dE_3(t)}{dt} \\ -J_1(t) \end{bmatrix}.$$

Состояние цепи $\mathbf{x}(\Delta t)$ в момент времени $t = \Delta t$ по ее состоянию $\mathbf{x}(0)$ в момент $t = 0$ и известным значениям функции $\mathbf{f}(0)$, $\mathbf{f}(\Delta t/3)$, $\mathbf{f}(2\Delta t/3)$, $\mathbf{f}(\Delta t)$ дает формула

$$\mathbf{x}(\Delta t) = \\ = 2\,\mathrm{Re}\left[(\Delta t\,\mathbf{H} - z_1\mathbf{G})^{-1}\left(y_1\mathbf{G}\mathbf{x}(0) + \Delta t\sum_{k=0}^{3} b_{1k}\mathbf{f}\left(\frac{k\Delta t}{3}\right)\right)\right], \quad (2)$$

здесь $z_1 = 2 - j\sqrt{2}$, $y_1 = 1 + j\,2{,}5\sqrt{2}$;

$b_{10} = 1{,}0606601717798216\,j$,

$b_{11} = -1{,}5909902576697323\,j$,

$b_{12} = 3{,}181980515339464\,j$,

$b_{13} = -0{,}5 - 1{,}2374368670764584\,j$.

Формула (2) может быть легко преобразована к виду, в котором главная матрица имеет такую же структуру, как у метода узловых потенциалов [4]. Это значительно сокращает размерность задачи, но для описания расширения метода на нелинейные цепи (2) будет более понятной. Отметим, что матрицы в (2) также легко получаются по схеме цепи.

Формула (2) требует задания значений $\mathbf{f}(t)$ в 4 точках шага интегрирования. Наибольшее число точек равно порядку аппроксимации Паде + 1, для аппроксимации 3-го порядка R_{12} – это 4. Однако численные эксперименты показывают, что для аппроксимаций более высоких порядков увеличение числа точек практически не повышает точность расчетов. Возможно, точность ограничена степенью полинома для расчета тока внутри шага интегрирования.

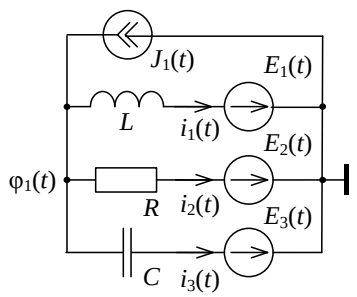


Рис. 3. Схема электрической цепи

Условия сходимости для резистивной цепи

Выделим в системе координат, в которой изображена ВАХ нелинейного резистора (ВАХ НР) (см. рис. 1 и 2), прямоугольник, в котором расположен участок ВАХ, имеющий отношение к расчету переходного процесса в цепи на данном шаге по времени и данной итерации. Это прямоугольник, ограниченный токами i_0, $i_0 + \Delta i$ и напряжениями $u(i_0)$, $u(i_0 + \Delta i)$ (рис. 4).

Введем в этом прямоугольнике локальные координаты $\tilde{i}$ и $\tilde{u}$. Диагональ прямоугольника имеет уравнение $\tilde{u}(\tilde{i}(t)) = R_\Delta \tilde{i}(t)$, а ВАХ НР представляется в виде $\tilde{u}(\tilde{i}(t)) = R_\Delta \tilde{i}(t) + \tilde{E}_R(\tilde{i}(t))$. Линеаризация ВАХ – это замена участка ВАХ на диагональ $\tilde{u}(\tilde{i}(t)) = R_\Delta \tilde{i}(t)$. В локальных координатах это ВАХ статического сопротивления $R_{st} = R_\Delta$. Тогда можно применить условие сходимости итерационного процесса расчета цепи с нелинейным резистором методом статических сопротивлений, полученное в [20]. В самом плохом для сходимости случае, когда нелинейный резистор подключен к идеальному источнику напряжения, это условие имеет вид

$$0 < R_d < 2R_{st}. \qquad (3)$$

Это условие выполняется для ВАХ вида $\tilde{u}(\tilde{i}) = c\tilde{i}^n$ при $0 < n < 2$, где $c > 0$ – некоторая константа. Поэтому локальное условие сходимости итераций для нелинейного резистора, включенного в любую резистивную цепь, можно сформулировать так: график ВАХ должен лежать внутри заштрихованной на рис. 4 области. Снизу эта область ограничена параболой $\left[u(i_0 + \Delta i) - u(i_0)\right]\left(\tilde{i}/\Delta i\right)^2$. Из (3) следует, что ВАХ должна быть монотонно возрастающей функцией.

В общем случае, если нелинейный резистор включен в резистивную цепь с эквивалентным сопротивлением R_0, условие сходимости для метода статических сопротивлений выглядит так [20]:

$$\left|R_d - R_{st}\right| < R_0 + R_{st}.$$

Отсюда можно получить условие сходимости для случая $R_d < R_{st}$ (для области над диагональю прямоугольника на рис. 4): $-R_0 < R_d$.

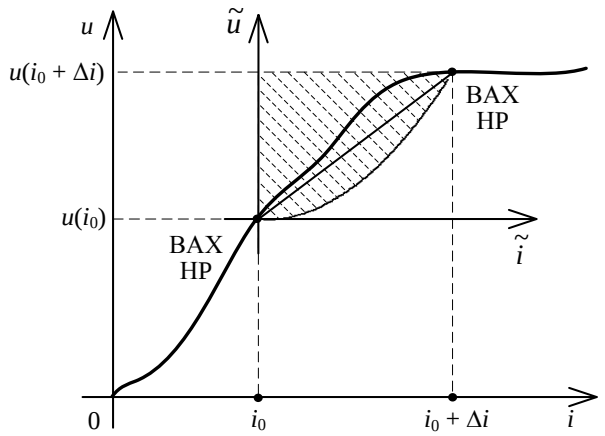


Рис. 4. Локальные координаты и область сходимости итерационного процесса (заштрихована)

Для случая $R_d > R_{st}$ (для области под диагональю прямоугольника) условие примет вид $\tilde{u}(\tilde{i}) > -R_0\tilde{i} + c\tilde{i}^2$. Область сходимости будет ограничена снизу параболой

$$\left[u(i_0 + \Delta i) - u(i_0) + R_0\Delta i\right]\left(\tilde{i}/\Delta i\right)^2 - R_0\tilde{i}.$$

Условие (3) будет расширено: $-R_0 < R_d < 2R_{st} + R_0$.

Чтобы не усложнять задачу теоретически возможными локальными колебаниями, мы рассматриваем только монотонные ВАХ.

Теорема Тевенена позволяет распространить полученные результаты на любые линейные и линеаризованные резистивные цепи, подключенные к нелинейному резистору. Итерационный процесс не раскачивается решением ДАСУ цепи благодаря L-устойчивости используемых методов ее решения, кроме метода трапеций, который не обладает свойством L-устойчивости.

Для ускорения сходимости полезно применение релаксации в виде

$$R_{\Delta\, k+1} = 0{,}5\left[R_{\Delta\, k} + u(i(\Delta t)) - u(i(0))/(i(\Delta t) - i(0))\right].$$

Решение тестовых задач

Для тестирования программы на основе предложенного метода использовались различные ВАХ нелинейных резисторов. Далее показаны результаты для симметричной кусочно-линейной ВАХ, у которой в узлах характеристики дифференциальное сопротивление растет скачком с ростом тока в 1,8 раза (таблица, рис. 5). Для других ВАХ, в том числе с непрерывной зависимостью $R_d(i)$, характер представленных результатов существенно не меняется.

Нелинейный резистор с идеальным источником напряжения включался последовательно поочередно с постоянным резистором (R = 10 Ом), катушкой (L = 0,2 Гн) и конденсатором (C = 200 мкФ).

ВАХ резистора с растущим сопротивлением R_d

i, А	0	0,25	0,5	0,75	1,0	1,25	1,5	1,75	2,0
u, В	0	1,6	4,4	9,5	18,8	35,3	65,1	118,8	215,3

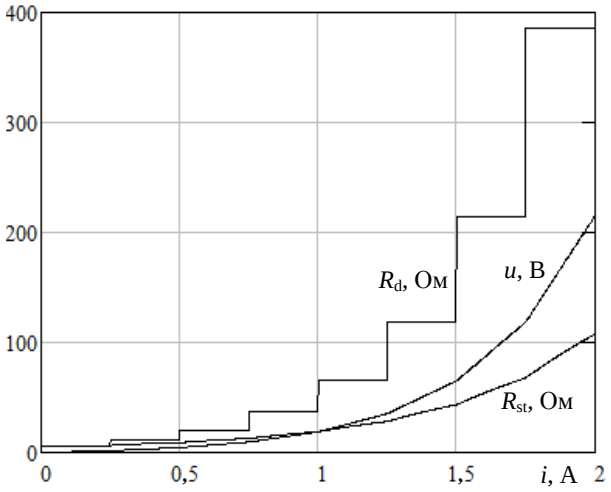


Рис. 5. ВАХ нелинейного резистора с растущим сопротивлением R_d

Зависимость ЭДС источника от времени предварительно определялась при синусоидальном токе амплитудой 2 А частотой 10 Гц. Затем эта ЭДС использовалась для численного расчета процесса в тестовой цепи. ВАХ нелинейного резистора масштабировалась так, чтобы его среднеквадратическое напряжение было равно напряжению постоянного пассивного элемента. Погрешность вычислялась сравнением численно рассчитанного тока нелинейного резистора с заданной идеальной синусоидой.

На рис. 6 для цепи с конденсатором и катушкой представлена зависимость относительной среднеквадратической погрешности тока δ от числа шагов интегрирования на период N.

Графики даны для метода Эйлера, метода трапеций и методов на основе аппроксимации Паде матричной экспоненты R_{12} 3-го порядка точности и R_{23} 5-го порядка точности. Отметки $R_{12}(2)$ и $R_{23}(2)$ соответствуют параболической интерполяции тока в пределах шага интегрирования, $R_{12}(3)$, $R_{23}(3)$ – кубической интерполяции.

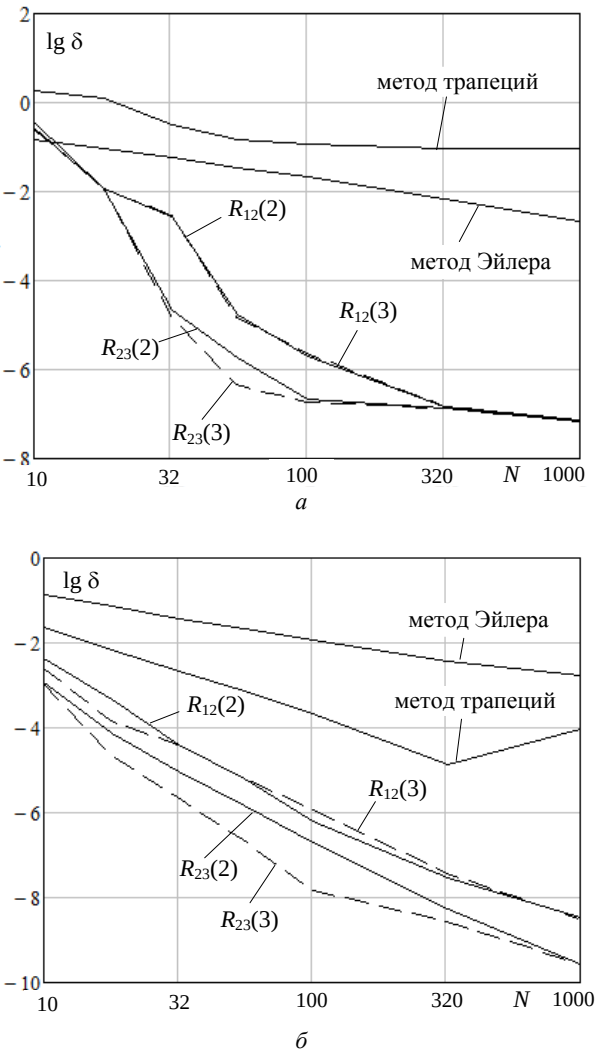


Рис. 6. Логарифм относительной среднеквадратической погрешности от числа шагов интегрирования на период. *а* – нелинейный резистор с конденсатором, *б* – нелинейный резистор с катушкой

На рис. 7 показано среднее число итераций для цепи с катушкой, которое потребовалось для достижения относительной разности между последовательными приближениями тока менее 10^{-6}.

Далее покажем пример расчета переходного процесса в более сложной цепи, схема которой показана на рис. 8.

Параметры цепи: R_1 = 4 Ом, R_2 = 20 кОм, L_1 = 10 мГн, L_2 = 1 мГн, C_1 = 20 мкФ, C_2 = 0,2 мкФ. ВАХ нелинейного резистора задана формулой $u_3(i_3) = 10i_3^5$. Начальные условия: $i_4(0)$ = 2А, $i_5(0) = 0$, $u_6(0)$ = –5 В, $u_7(0) = 0$.

Для сравнения численного решения с точным аналитическим решением применялся следующий прием. Вначале операторным методом рассчитывался ток $i_4(t)$ для переходного процесса в линейной цепи, показанной на рис. 9. При этом постоянный источник ЭДС E_1 имел напряжение 20 В. Далее напряжение источника ЭДС $E_{13}(t)$ в цепи с нелинейным резистором, показанной на рис. 8, полагалось равным $E_1 + u_3(i_4(t))$. Тем самым для линейной части цепи нелинейный резистор фактически исключался, так как его напряжение компенсировалось источником $E_{13}(t)$ – при условии верного расчета тока и напряжения этого резистора численным методом. В целом же численный расчет получался для реальной нелинейной цепи рис. 8 без каких-либо упрощений.

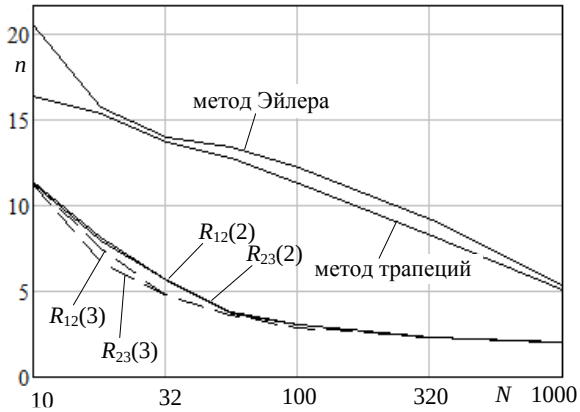


Рис. 7. Зависимость среднего числа итераций для решения нелинейной системы уравнений цепи на каждом шаге по времени от числа шагов интегрирования на период. Нелинейный резистор с катушкой

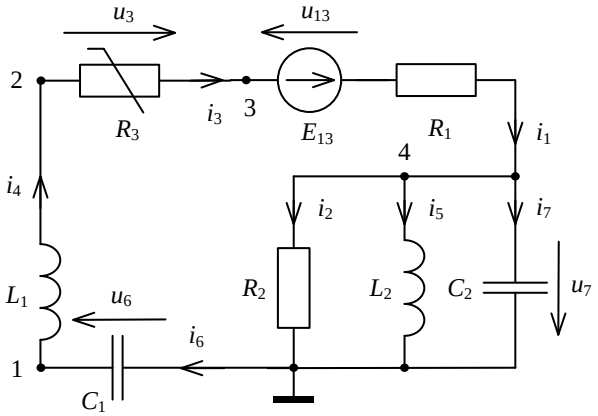


Рис. 8. Цепь с нелинейным резистором для численного расчета

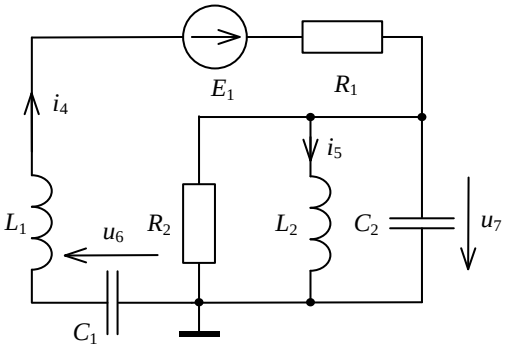


Рис. 9. Линейная цепь для расчета операторным методом

В исходном виде методы расчета переходных процессов непосредственно содержат уравнения ветвей, состоящих из одного пассивного двухполюсника с последовательно включенным источником ЭДС или без него. Для записи таких уравнений нужны простые узлы № 1–3. Можно использовать составные *RL*- и *RC*-ветви, что уменьшает число простых узлов, уравнений и неизвестных, но матрицы становятся чуть сложнее. Матрицы $\Delta t\mathbf{H} - z_k\mathbf{G}$ нетрудно привести к виду, подобному матрице метода узловых потенциалов. Это существенно сокращает размерность задачи [4]. Однако здесь для ясного представления методов покажем самый простой вариант формул. Этот вариант можно непосредственно использовать для расчетов, применяя методы решения систем уравнений для разреженных матриц.

Матрицы для цепи рис. 8 имеют вид

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & L_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & L_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -C_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & C_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{H} = \begin{bmatrix} -R_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & -R_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -R_\text{d} & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{x} = \begin{bmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 & i_7 & \phi_1 & \phi_2 & \phi_3 & \phi_4 \end{bmatrix}^\text{T},$$

$$\mathbf{f} = \begin{bmatrix} E_{13}(t) & 0 & -E_T(t) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\text{T}.$$

Формула для метода на основе аппроксимации Паде матричной экспоненты R_{23} выглядит так:

$$\mathbf{x}(\Delta t) = (\Delta t\,\mathbf{H} - z_1\mathbf{G})^{-1}\left(y_1\mathbf{G}\mathbf{x}(0) + \Delta t\sum_{k=0}^{3} b_{1k}\mathbf{f}\left(\frac{k\Delta t}{3}\right)\right) +$$

$$+2\,\mathrm{Re}\left[(\Delta t\,\mathbf{H} - z_2\mathbf{G})^{-1}\left(y_2\mathbf{G}\mathbf{x}(0) + \Delta t\sum_{k=0}^{3} b_{2k}\mathbf{f}\left(\frac{k\Delta t}{3}\right)\right)\right],$$

где

$z_1 = 3{,}6378342527444945$, $y_1 = -18{,}297498174845835$,
$z_2 = 2{,}6810828736277523 - 3{,}050430199247411\,j$,
$y_2 = 7{,}648749087422917 - 4{,}171640244747441\,j$,
$b_{10} = -1{,}4456897385025358$,
$b_{11} = -3{,}8031665043431016$,
$b_{12} = 1{,}001957896137494$, $b_{13} = -0{,}7828795111042717$,
$b_{20} = 0{,}722844869251269 - 0{,}10110612415667874\,j$,
$b_{21} = 1{,}9015832521715508 - 0{,}14117930024617908\,j$,
$b_{22} = -0{,}5009789480687474 + 1{,}3180723497684803\,j$,
$b_{23} = -0{,}10856024444786472 - 0{,}33927937032482625\,j$.

Формула для метода на основе аппроксимации Паде матричной экспоненты R_{34} имеет вид

$$\mathbf{x}(\Delta t) = 2\,\mathrm{Re}\left[(\Delta t\,\mathbf{H} - z_1\mathbf{G})^{-1}\left(y_1\mathbf{G}\mathbf{x}(0) + \Delta t\sum_{k=0}^{3} b_{1k}\mathbf{f}\left(\frac{k\Delta t}{3}\right)\right) +\right.$$

$$\left. + (\Delta t\,\mathbf{H} - z_3\mathbf{G})^{-1}\left(y_3\mathbf{G}\mathbf{x}(0) + \Delta t\sum_{k=0}^{3} b_{3k}\mathbf{f}\left(\frac{k\Delta t}{3}\right)\right)\right],$$

где

$z_1 = -0{,}6232430928037013 + 1{,}3600845532157446\,j$,
$y_1 = -11{,}301539995971506 + 12{,}4716758502502\,j$,
$z_3 = 4{,}787193103128463 + 1{,}5674764168952087\,j$,
$y_3 = 13{,}301539995971508 - 60{,}071732737047355\,j$,
$b_{10} = -0{,}6232430928037013 + 1{,}3600845532157446\,j$,
$b_{11} = 1{,}0359181466945429 + 2{,}663457546206927\,j$,
$b_{12} = 0{,}46944067839791304 - 1{,}41102265408329\,j$,
$b_{13} = -0{,}18073859691170266 + 0{,}2273466755528692\,j$,
$b_{30} = 0{,}623243092803702 - 5{,}110492939064246\,j$,
$b_{31} = -1{,}0359181466945424 - 7{,}715491508183748\,j$,
$b_{32} = -0{,}4694406783979126 + 0{,}9543069059443732\,j$,
$b_{33} = -0{,}31926140308829787 - 0{,}283378909525589\,j$.

Переходной процесс в цепи рис. 8 и рис. 9 имеет две составляющие с постоянными времени $\tau_1 = 5{,}501$ мс и $\tau_2 = 6{,}983$ мс, с периодами колебаний $T_1 = 2{,}958$ мс и $T_2 = 84{,}72$ мкс.

На рис. 10 показаны напряжение и ток нелинейного резистора, на рис. 11 – относительная среднеквадратическая погрешность тока нелинейного резистора в зависимости от числа точек интегрирования на период высокочастотной составляющей переходного процесса, на рис. 12 – число итераций k, нужное для выполнения условия

$$\left\|\mathbf{x}_k(\Delta t) - \mathbf{x}_{k-1}(\Delta t)\right\| / \left(\left\|\mathbf{x}_k(\Delta t)\right\| + \left\|\mathbf{x}_{k-1}(\Delta t)\right\|\right) < 10^{-6}.$$

Интервал интегрирования – один период низкочастотных колебаний $T_1 = 2{,}958$ мс. На рис. 10 и 11 обозначено: $R_{12}(2)$, $R_{23}(2)$, $R_{34}(2)$ – расчет методами на основе соответствующих аппроксимаций Паде мат-

ричной экспоненты с полиномом 2-й степени для вычисления тока нелинейного резистора внутри интервала интегрирования; $R_{12}(3)$, $R_{23}(3)$, $R_{34}(3)$ – с полиномом 3-й степени. Во всех представленных расчетах для функции **f** использовались 4 точки внутри шага интегрирования.

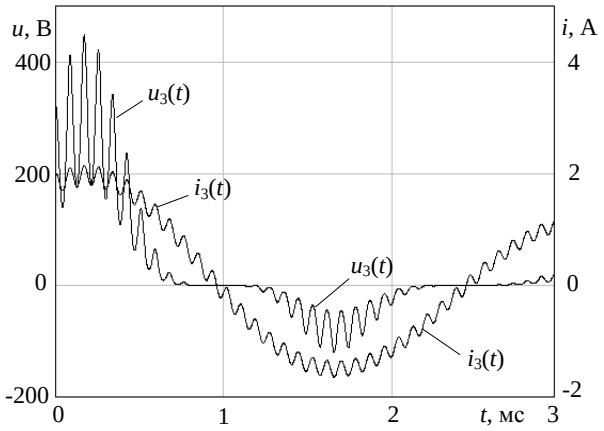


Рис. 10. Напряжение и ток нелинейного резистора

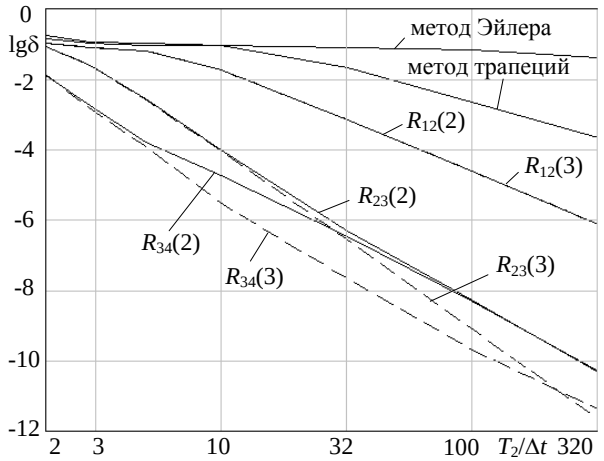


Рис. 11. Логарифм относительной среднеквадратической погрешности расчета тока нелинейного резистора в зависимости от числа точек интегрирования на период высокочастотных колебаний T_2

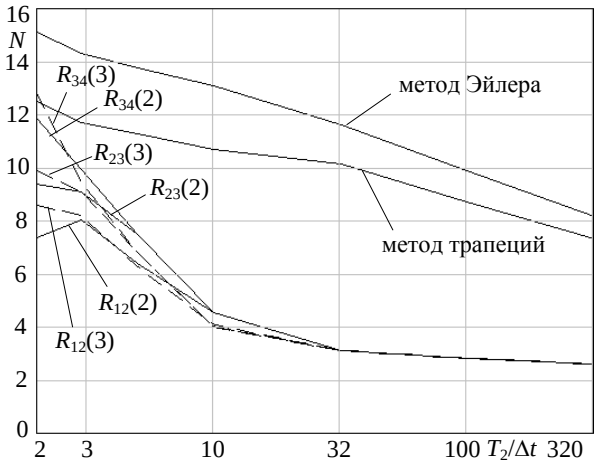


Рис. 12. Среднее число итераций, нужное для решения нелинейной системы уравнений разными методами

Анализ результатов

Для простейших цепей с катушкой и конденсатором, а также для сложной цепи новые методы дают погрешность на несколько порядков меньше, чем метод Эйлера и метод трапеций. Метод трапеций для цепи с конденсатором из-за раскачки колебаний дает неприемлемую относительную погрешность порядка 0,1 для любого числа шагов интегрирования на период. Расчет цепи с нелинейным резистором последовательно с постоянным резистором дает относительную погрешность порядка 10^{-6} для метода Эйлера и метода трапеций и 10^{-7} для методов на основе матричной экспоненты почти независимо от величины шага интегрирования.

Благодаря теореме Тевенена результаты решения представленных простейших тестовых задач можно распространить на любые линейные или линеаризованные цепи, у которых эквивалентное сопротивление представляет собой постоянный элемент, с которым были проделаны расчеты.

Несмотря на то, что на одну итерацию метод R_{12} требует примерно в 4 раза больше вычислений, чем метод Эйлера и метод трапеций, метод R_{23} в 5 раз, а R_{34} в 8 раз больше, число арифметических операций для достижения заданной точности у новых методов существенно меньше (с учетом соответствующего числа итераций и точек интегрирования на период колебаний решения). Достоинством новых методов является возможность высокоточных расчетов с даже с большим шагом интегрирования. Зависимость погрешности методов от величины шага интегрирования и от способа расчета токов нелинейных резисторов внутри шага интегрирования требует дальнейшего исследования.

Отметим, что рассмотренные выше условия сходимости являются, во-первых, локальными и, во-вторых, достаточными, но не необходимыми. То есть они обеспечивают сходимость, но при невыполнении этих условий сходимость не исключается. Выше рассмотрены условия сходимости итераций для нелинейного резистора, включенного последовательно к источнику ЭДС и линейному резистору. Переход к локальным координатам имеет строгое обоснование, которое не приводится из-за ограничений по объему статьи. В самом начале расчета значения R_Δ и E_Δ определяются с помощью численного или аналитического дифференцирования ВАХ нелинейного резистора при начальном токе. На самой первой итерации принимаем $E_T(t) \approx E_\Delta$.

Сходимость для цепей с катушкой и конденсатором требует дополнительных исследований. В статье рассмотрен нелинейный резистор, управляемый током. Аналогичный метод можно применить и для резистора, управляемого напряжением. При нарушении сходимости на каком-либо участке ВАХ проблема решается локальным уменьшением шага.

Если в пределах шага интегрирования функция $\mathbf{f}(t)$ задана в виде $\mathbf{f}(t) = \mathbf{f}_0 + \mathbf{f}_1 t + \mathbf{f}_2 t^2 + \mathbf{f}_3 t^3$, то $\mathbf{x}(\Delta t)$ для аппроксимации Паде матричной экспоненты R_{12} рассчитывается по формуле

$$\mathbf{x}(\Delta t) = 2\,\mathrm{Re}\left[(\Delta t\mathbf{H} - z_1\mathbf{G})^{-1}\left(y_1\mathbf{G}\mathbf{x}(0) + \Delta t \sum_{k=0}^{3} a_{1k}\mathbf{f}_k(\Delta t)^k \right)\right],$$

где $a_{10} = -0{,}5 + j\sqrt{2}$, $a_{11} = -0{,}5 + 0{,}25j\sqrt{2}$, $a_{12} = -0{,}5$, $a_{13} = -0{,}5 - 0{,}25j\sqrt{2}$.

Для остальных аппроксимаций Паде также возможно представление функции **f** в виде полинома. Его степень не превышает порядок точности метода, который равен сумме степеней числителя и знаменателя аппроксимации Паде. Для аппроксимации R_{12} он равен 3, для R_{23} равен 5, для R_{34} – 7. Неявный метод Эйлера 1-го порядка точности и метод трапеций 2-го порядка точности могут быть получены с помощью аппроксимаций Паде матричной экспоненты R_{01} и R_{11} соответственно.

Заключение

Применение новых высокоточных методов расчета переходных процессов в линейных электрических цепях расширено на цепи с нелинейными резисторами.

Предложен новый повышенной точности способ линеаризации ВАХ, сочетающий свойства линеаризации с помощью дифференциального и статического сопротивления, позволяющий перенести нелинейность в свободный член системы уравнений и учитывать ее в решении.

Предложен способ тестирования методов расчета нелинейных цепей с помощью аналитических расчетов линейных цепей.

Новые методы отличаются предельной простотой и универсальностью. Они хорошо приспособлены к расчетам нелинейных цепей, а также цепей с частыми коммутациями, для которых из-за изменений параметров цепи нужно решение задачи с новыми матрицами на каждом шаге интегрирования. Численные расчеты подтверждают работоспособность и надежность новых методов.

Далее предполагается развитие методов для их применения к цепям с нелинейными индуктивностями и емкостями, а также к параметрическим и нелинейно-параметрическим цепям.

Литература

1. Гридин В.Н. Численно-аналитическое моделирование радиоэлектронных схем / В.Н. Гридин, В.Б. Михайлов, Л.Б. Шустерман. – М.: Наука, 2008. – 342 с.

2. Пилипенко А.М. Применение гибридных методов численного решения обыкновенных дифференциальных уравнений для анализа автоколебательных цепей с различной динамикой // Известия ЮФУ. Технические науки. – 2026. – № 1 (249). – С. 112–122.

3. Study of a Numerical Integral Interpolation Method for Electromagnetic Transient Simulations / K. Sun, K. Chen, H. Cen, F. Tan, X. Ye // Energies (MDPI). – 2024. – Vol. 17, No. 15. – Art. 3837. DOI: 10.3390/en17153837.

4. Бурцев Ю.А. Расчет переходных процессов в линейных электрических цепях с использованием аппроксимации Паде матричной экспоненты // Электричество. – 2024. – № 8. – С. 26–35.

5. Хайрер Э. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи: пер. с англ. / Э. Хайрер, С. Нёрсетт, Г. Ваннер. – М.: Мир, 1999. – 685 с.

6. Гантмахер Ф.Р. Теория матриц. – Изд. 5-е. – М.: Физматлит, 2010. – 560 с.

7. Бейкер Дж. мл. Апрроксимации Паде: пер. с англ. / Дж. мл. Бейкер, П. Грейвс-Моррис. – М.: Мир, 1986. – 502 с.

8. Бурцев Ю.А. Решение задачи Коши высокоточными методами на основе аппроксимации Паде матричной экспоненты // Труды Института системного анализа РАН. – 2024. – Т. 74, № 1. – С. 3–12.

9. Орлов А.И. Моделирование динамических режимов электрических цепей с диодами / А.И. Орлов, С.В. Волков, В.Т. Сидорова // Вестник Чувашского ун-та. – 2025. – № 4. – С. 98–110. DOI: 10.47026/1810-1909-2025-4-98-110.

10. Hernandez H. A New Adaptive Step-size Euler ODE Solver // ForsChem Research Reports. – 2025. – Vol. 10 (part 1). – P. 233–264.

11. Корнев А.С. Алгоритм расчета переходных процессов электрических цепей методом переменных состояния / А.С. Корнев, С.И. Вольский, И.П. Викулов, Н.В. Кузнецов, Д.А. Сорокин // Интеллектуальная электротехника. – 2022. – № 1. – С. 99–112.

12. Корнев А.С. Алгоритм поиска аналитического решения для расчета переходных процессов методом матричной экспоненты / А.С. Корнев, Н.В. Кузнецов, С.И. Вольский, И.П. Викулов // Интеллектуальная электротехника. – 2023. – № 2. – С. 4–17.

13. Shilin A.N. Improvement of the Accuracy of Transient Process Analysis in Complex Power Systems Using the Z-Transform Method / A.N. Shilin, A.A. Shilin, P.V. Dikarev, D.T. Ngo // Problemele Energeticii Regionale. – 2025. – No. 4 (68). – P. 118–130. DOI:10.52254/1857-0070.2025.4-68.09.

14. Бычков Ю.А. Аналитически-численный метод расчета динамических систем / Ю.А. Бычков, С.В. Щербаков. – 2-е изд., доп. – СПб.: Энергоатомиздат, 2002. – 368 с.

15. Михайлов В.Б. Спектральный численно-аналитический метод Ньютона–Канторовича–Михайлова для решения нелинейных дифференциально-алгебраических систем уравнений и его применение для анализа нелинейных режимов радиоэлектронных схем / В.Б. Михайлов, В.В. Румянцев // Электричество. – 2012. – № 1. – С. 50–54.

16. Артым А.Д. Новый метод расчета процессов в электрических цепях / А.Д. Артым, В.А. Филин, К.Ж. Есполов. – СПб.: Элмор, 2001. – 192 с.

17. Бычков Ю.А. Численный расчет неавтономных динамических моделей с нелинейностями в виде нестепенных элементарных функций / Ю.А. Бычков, Е.Б. Соловьева, С.В. Щербаков // Известия СПбГЭТУ «ЛЭТИ». – 2022. – Т. 15, № 2. – С. 58–66.

18. Ланцов В.Н. Состояние в области методов моделирования нелинейных ВЧ-электронных устройств связи (обзор): ч. 1 // Проектирование и технология электронных средств. – 2012. – № 4. – С. 2–11.

19. Ланцов В.Н. Состояние в области методов моделирования нелинейных ВЧ-электронных устройств связи (обзор): ч. 2 // Проектирование и технология электронных средств. – 2013. – № 1. – С. 16–23.

20. Бурцев Ю.А. Расчет переходных режимов в полупроводниковых преобразователях с применением статических сопротивлений и неявного метода Эйлера / Ю.А. Бурцев, А.В. Павленко, И.В. Васюков // Электротехника. – 2020. – № 12. – С. 13–19.

Бурцев Юрий Алексеевич
Канд. техн. наук, доцент каф. электроснабжения
и электроприводов (ЭиЭ) Южно-Российского
государственного политехнического университета
им. М.И. Платова (ЮРГПУ НПИ)
ул. Просвещения, 132, г. Новочеркасск, Россия, 346428
ORCID: 0000-0001-6663-3239
Тел: +7-961-305-22-76
Эл. почта: proton36@yandex.ru

Поступила в редакцию: 05.05.2026.
Принята к публикации: 04.08.2026.

Burtsev Y.A.

Development of Matrix-Exponential-Based Methods for Transient Analysis of Circuits with Nonlinear Resistors

Relevance. The development of new methods for transient analysis of electrical circuits is relevant despite the availability of numerous well–known methods and computer programs. The reason is that practical tasks impose a set of requirements that existing methods do not always fully satisfy. As a result, they give significant inaccuracy or even qualitatively incorrect results, or require excessive computational time and become very complicate or completely inapplicable. The methods should be suitable for stiff, oscillatory, and stiff-oscillatory circuits with a wide range of parameters; highly accurate (order of accuracy of at least 3); computationally efficient; be *L*-stable (this eliminates false oscillations and restrictions on the integration step); adapted for calculating nonlinear circuits and circuits with frequent switching; simple and convenient, without requiring configuration, preliminary calculations and evaluations. ***The purpose of the study.*** To expand the set of methods previously proposed by the author for calculating transients in linear electrical circuits on circuits with nonlinear resistors. ***Methods.*** The presented work is based on a new set of high-precision implicit numerical methods for calculating transients in linear electrical circuits. The new methods are based on decomposition of the Padé approximation of the matrix exponential into elementary partial fractions and satisfy all the requirements listed above. They are very well adapted to the structure of electrical-circuit equation systems, and they are much simpler than other high-precision methods. Nonlinear equations are solved using the method of simple iterations and current interpolation within the integration interval. ***Scientific novelty.*** New effective methods for calculating transients in linear circuits have been adapted to the calculation of circuits with nonlinear resistors. An original, high-precision linearization of the resistor current-voltage characteristic is proposed, together with a method for calculating currents within the integration interval. All results presented in this study were obtained personally by the author. ***Results.*** The proposed new methods make it possible to calculate arbitrary modes of circuits containing nonlinear resistors with an error several orders of magnitude smaller than that of the trapezoidal rule integration, one of the most commonly used for this purpose. ***Practical significance.*** Calculations of dynamic processes in circuits with nonlinear elements have numerous practical applications. In particular, they are required for the design of power semiconductor inverters, digital equipment, power supply systems, electronic and radio engineering devices and their components.

Keywords: electrical circuits, nonlinear resistors, transient processes, *L*-stable methods, high precision methods, matrix exponential, Padé approximation.
DOI: 10.21293/1818-0442-2026-29-2-189-198

References

1. Gridin V.N., Mikhaylov V.B., Shusterman L.B. *Chislenno-analiticheskoe modelirovaniye radioielektronnyh shem* [Numerical-analytical modelling of radioelectronic schemes]. Moscow, Nauka Publ., 2008, 342 p. (in Russ.).

2. Pilipenko A.M. [Application of hybrid methods for numerical solving of ordinary Differential Equations for Analysis of self-oscillating circuits with various dynamics]. *Izvestiya SFedU. Engineering sciences*, 2026, no. 1 (249), pp. 112–122 (in Russ.).

3. Sun K., Chen K., Cen H., Tan F., Ye X. Study of a Numerical Integral Interpolation Method for Electromagnetic Transient Simulations. *Energies (MDPI)*, 2024, vol. 17, no. 15, art. 3837. DOI: 10.3390/en17153837

4. Burtsev Y.A. [Calculation of transients in linear electrical circuits using the Pade approximation of a matrix exponent]. *Elektrichestvo*, 2024, no. 8, pp. 26–35 (in Russ.).

5. Hairer E., Wanner J. Solving ordinary differential equations II. Stiff and differential-algebraic problems. Second revised edition. Berlin, Springer-Verlag, 1996, 599 p.

6. Ganmaher F.R. *Teoriya matritz* [Theory of matrixes – 5th edition]. Moscow, Fizmatlit Publ., 2010, 560 p. (in Russ.).

7. Baker J. Jr., Graves-Morris P. Padé approximants. Massachusetts, Addison-Wesley Publishing Company, 1981, 459 p.

8. Burtsev Y.A. [Solving of Cauchy Problem with High Precision Methods Based on Matrix Exponent]. *Trudy Instituta sistemnogo analiza rossiyskoy akademii nauk (Proceedings of the institute for systems analysis Russian academy of sciences)*, 2024, vol. 74, no. 1, pp. 3–12 (in Russ.).

9. Orlov A.I., Volkov S.V., Sidorova V.T. [Modeling of dynamic modes of electrical circuits with diodes]. *Vestnik Chuvashskogo universiteta*, 2025, no. 4, pp. 98-110 (in Russ.).

10. Hernandez H. A New Adaptive Step-size Euler ODE Solver. *ForsChem Research Reports*, 2025, vol. 10, part 1, pp. 233–264.

11. Kornev A.S., Volskiy S.I., Vikulov I.P., Kuznetsov N.V., Sorokin D.A. [Algorithm for calculating transients in electrical circuits by the state-variable method]. *Smart Electrical Engineerin*g, 2022, no. 1, pp. 99–112 (in Russ.).

12. Kornev A.S., Kuznetsov N.V., Volskiy S.I., Vikulov I.P. [Search algorithm for calculating transients by matrix exponential method]. *Smart Electrical Engineering*, 2023, no. 2, pp. 4–17 (in Russ.).

13. Shilin A.N., Shilin A.A., Dikarev P.V., Ngo D.T. Improvement of the Accuracy of Transient Process Analysis in Complex Power Systems Using the Z-Transform Method. *Problemele energeticii regionale*, 2025, no. 4 (68), pp. 118–130.

14. Bychkov Y.A., Scherbakov S.V. *Analiticheski-chislenniy metod rascheta dinamicheskih sistem* [Analytical-numerical method for calculation of dynamic systems]. Saint-Petersburg, Energoatomizdat Publ., 2002, 368 p. (in Russ.).

15. Mikhaylov V.B., Rumyantsev V.V. [Newton–Kantorovich–Mikhailov spectral numerical analytical method for solving nonlinear differential algebraic equation systems and its application to the analysis of nonlinear modes of electronic circuits]. *Elektrichestvo*, 2012, no. 1, pp. 50–54 (in Russ.).

16. Artym A.D., Filin V.A. *Noviy metod rascheta processov v elektricheskih tsepyah* [New method of calculation processes in electrical circuits]. Saint-Petersburg, Elmor Publ., 2001, 192 p. (in Russ.).

17. Bychkov Y.A., Solovieva E.B., Scherbakov S.V. Numerical Calculation of Non-Autonomous Dynamic Models with Nonlinearities in the Form of Non-Power Elemental Functions. *LETI Transactions on Electrical Engineering & Computer Science*, 2022, vol. 15, no. 2, pp. 58–66 (in Russ.).

18. Lantsov V.N. *Sostoyaniye v oblasti metodov modelirovaniya nelineynykh VCH elektronnykh ustroystv svyazi (obzor) chast' 1* [The state of the art in modeling of nonlinear HF electronic communication devices (review) Part 1]. *Design and technology of electronic means*, 2012, no. 4, pp. 2–11 (in Russ.).

19. Lantsov V.N. *Sostoyaniye v oblasti metodov modelirovaniya nelineynykh VCH elektronnykh ustroystv svyazi (obzor) chast' 2* [The state of the art in modeling of nonlinear HF electronic communication devices (review) Part 2]. *Design and technology of electronic means*, 2013, no. 1, pp. 16–23 (in Russ.).

20. Burtsev Y.A., Pavlenko A.V., Vasyukov I.V. [Calculation of transient modes in semiconductor converters using static resistances and the implicit Euler method]. Russian Electrical Engineering, 2020, vol. 91, no. 12, pp. 730–736 (in Russ.)

Yuri A. Burtsev
Candidate of Technical Sciences, Associate Professor,
Electrical Power Department,
Platov South-Russian Polytechnic University (NPI)
132, Prosvesheniya st., Novocherkassk, Russia, 346411
ORCID: 0000-0001-6663-3239
Phone: +7-961-305-22-76
Email: proton36@yandex.ru

Received: 05.05.2026.
Accepted: 04.08.2026.

Требования к подготовке статей,

представляемых в редакцию для публикации в журнале

«Доклады Томского государственного университета систем управления и радиоэлектроники»

(в требования и шаблон на сайте внесены дополнения и изменения от 01.04.2026)

Статья в электронном виде, поданная ответственным автором по эл. почте на адрес редакции, автоматически является соглашением автора на выполнение всех требований редакции журнала к статье и ответственностью за ее содержание. Плата за публикацию статей не взимается, гонорар не выплачивается.

1. Электронные варианты статьи должны быть предоставлены в виде файлов в форматах не ниже версий MS Word 2007 и сохранённых как «Документ Word» и .pdf, названных русской фамилией первого автора **по эл. почте на адрес редакции: journal@tusur.ru.**

2. Объем статьи должен быть не менее 5 стр. (но не более 20 стр.), включая рисунки и таблицы, список литературы должен включать не менее 15 источников. Объем обзорной статьи – до 30 стр., включая рисунки и таблицы, список литературы – от 20 источников.

3. Статьи, поступившие в редакцию и соответствующие тематическим направлениям журнала, проверяются в системе «Антиплагиат». Оригинальность текста должна составлять не менее 80%. При выявлении многочисленных заимствований редакция действует в соответствии с правилами COPE.

4. Статья должна иметь: УДК; шифр специальности ВАК, которому должно соответствовать содержание статьи, И.О. Фамилии авторов; заглавие; аннотацию от 50 до 300 слов; ключевые слова; основной текст статьи; список библиографий под заголовком «Литература» – не менее 15 п.; сведения об авторах в полном объеме, включая наименование и адрес места работы, ORCID (при наличии), номер тел., эл. адрес. Далее на англ. яз.: Ф.И.О., заглавие статьи, аннотация, ключевые слова, References, сведения об авторах.

5. Текст статьи должен быть размещен в две колонки без принудительных переносов через один интервал шрифтом Times New Roman 10-го кегля на страницах формата А4. Для облегчения форматирования прилагается **шаблон статьи**, размещенный на сайте: **journal.tusur.ru.**

6. Одни и те же символы в тексте, формулах, таблицах и рисунках должны быть единообразными по написанию. Русские буквы и греческие символы набираются прямым шрифтом, а переменные, обозначенные латинскими, – курсивом, кроме слов, их сокращений, имен функций, программ, фирм и химических формул.

7. Формулы должны быть набраны **в формульном редакторе (MathType)** программы Word **(не конструктором!**).

Набор простых формул в строке допускается в текстовой форме – $F(x_n, y_1)$. В формулах русские буквы, индексы-метки от слов ($f_{n\text{ОС}}$ – частота n-го опорного сигнала); греческие символы (θ, ω, β, α, …); математические знаки (+, –, ×, ∈, =, скобки, …) и цифры – всегда набираются прямым нежирным шрифтом.

Переменные, обозначенные латинскими буквами, – курсивом, кроме англ. слов, их сокращений, имен функций, названий программ, элементов, фирм и химических формул (const; summa; MatLab Simulink; Cisco Systems; Adobe Acrobat; элемент Keysight N5A3-90Hz; sign(…); $\log x(t_1)$; cos θ; H_2O; $LiNbO_3$ и т.д.); TV_1 (транзистор); D_2 (детектор); $C_{1\ldots n}$; $L_{1\ldots m}$, $R_{1\ldots k}$ – конденсаторы, катушки индуктивности, резисторы с переменными значениями емкости, индуктивности, сопротивления; $U_{\text{вых}}$, $\mathbf{G}_{\text{sum}}$, $f_{\text{пч}}$ (сокращения – вых, sum, пч (промежут. частота) – это метки от слов: прямо, не курсив); T_z (где z – переменный параметр, курсив); λ_2.

Векторные величины, буквенные обозначения матриц, множеств, графов – жирным, прямо, не курсив – $\mathbf{A}_1$, $\mathbf{M}(f_{2n})$, $\boldsymbol{\beta}_x$ (где x, n, f – переменные параметры; цифровые индексы и числа – всегда не курсив, не жирным!).

Шаблоны для набора формул в редакторе MathType с заданными размерами для всех элементов размещены на сайте в шаблоне статьи.

8. Все употребляемые обозначения и сокращения должны быть пояснены.

9. Единицы измерения физических величин должны соответствовать Международной системе единиц (СИ) и написаны по-русски через пробел (x, ГГц; 43 дБ; 6,5 км/ч; $9{,}7\times10^{-5}$ А/см^2; T, град; 7 °С); 5% – без пробела). Десятичные числа пишутся через запятую (не точку).

10. **Таблицы и рисунки** должны иметь тематические заголовки (не повторяющие фразы-ссылки на них в тексте). (Рис. 1. Название рисунка. Таблица 1. Название таблицы). Большие блоки расшифровки условных обозначений лучше приводить в тексте. Подписи и текстовые надписи на рис. – Times New Roman, 9 пт (после масштабирования), не жирным, не курсивом; *кривые, обозначенные цифрами на графиках, – курсивом*, не жирным; переменные и параметры – так же, как и в тексте. На все рисунки и таблицы должны быть ссылки в тексте (… на рис. 1, … в табл. 1).

Гиперссылки и вкладки в статьях не использовать!

11. Рисунки, иллюстрации и фотографии должны быть **черно-белыми**, четкими, контрастными, аккуратными, сгруппированными. Графики – не жирно, сетка – четко. Единицы измерения – на русском. Десятичная запятая (не точка). Рисунки могут быть выполнены в программах CorelDraw, Illustrator, Word, Visio и должны давать возможность внесения исправлений. Для пояснительных подписей на рисунках используйте шрифт TNR. Размер изображения после масштабирования – 8 или 16,7 см по ширине при условии **читаемости всех надписей**, выполненных на русском языке не жирным, как основной текст, шрифтом TNR, после масштабирования размер надписей должен быть ~9 пт (на скриншотах может быть язык оригинала).

12. На все источники, указанные в списке литературы, должны быть ссылки по тексту (нумерация в порядке упоминания, например, [1, 2], [5–7]). Описание источников должно соответствовать ГОСТам 7.1–2003, Р 7.0.5–2008 и содержать всю необходимую для идентификации источника информацию, а именно: *для непериодических изданий* – Ф.И.О. автора, полное название работы, место издания, название изд-ва, год издания, кол-во стр.; *для периодических изданий* – Ф.И.О. автора, полное название работы, название журнала, год выпуска, номера тома и страниц (см. примеры оформления библиографии на след. стр. и на сайте в шаблоне).

Рекомендуется чтобы **не менее 50% публикаций**, на которые есть ссылки в статье**, были за последний 5-летний период.**

Контактная информация

Адрес редакции: 634050, г. Томск, пр. Ленина, 40.
МК, каб. 310, тел. +7 (382-2) 70-15-82 (редакция),
+7 (382-2) 51-21-21 с 14:00 до 17:30 ч
(*отв. за прием статей – Анна Михайловна Богдан*)

Эл. почта: journal@tusur.ru

9 771818 044708 02007

Правила и примеры оформления публикаций в списках лит-ры по ГОСТам

Ссылка дается на языке оригинальной публикации или издания

Внимание: *перед названием публикации – И.О. – пишется после фамилии автора* ***(Ф… И.О.)****, а после косой черты – перед* ***(И.О. Ф…)*** (*инициалы ИО от фамилии запятой не отделять, И.О. между собой – без пробела:* Иванов И.И.)

Описание для периодики (журналы):

1) 1 авт., продолжающееся издание, т.е. журнал:

Ф. И.О. Назв. публ. // Название журнала. – Год. – Т. _, № _. – С. _–_.

1. Беляев Ю.К. Аналитические случайные процессы // Теория вероятностей и ее применение. – 2019. – Т. 4, № 4. – С. 437–459.

2) 2, 3 автора, журнал:

Первый автор (Ф. И.О.). Название статьи / авторы (И.О. Ф…), …, … // Названия ж-ла. – Год. – Т. __, № _. – С. _–_.

2. Андреев Ю.А. Комбинированная антенна с расширенной полосой пропускания / Ю.А. Андреев, Ю.И. Буянов, В.И. Кошелев // Радиотехника и электроника. – 2005. – Т. 50, № 5. – С. 585–594.

3) более 4 авторов, журнал:

Назв. статьи / авторы (И.О. Ф…), …, … // Название ж-ла. – Год. – Т. _, № _, ч.__. – С. _–_.

3. Ultra-Wideband Source and Antenna Research / W.D. Prather, C.E. Baum, J.M. Lehr, J.P. O'Loughlin, S. Tyo, J.S. Schoenberg, R.J. Torres, T.C. Tran, D.W. Scholfield, J. Gaudet, J.W. Burger // IEEE Transactions on Plasma Science. – 2000. – Vol. 28, No. 5. – P. 117–130.

4. Исследование способов увеличения направленности симметричных направленных ответвителей диапазона 2–20 ГГц / Г.Г. Гошин, А.В. Зорин, С.А. Подлиннов и др. // Доклады ТУСУР. – 2016. – Т. 19, № 3. – С. 5–7.

5. Инфекционная теория новообразований и нанобактерия (перспективы исследований) / Г.В. Смирнов, В.Т. Волков, С.М. Шихман, Д.А. Васин, В.А. Хон // Сибирский медицинский журнал (Иркутск). – 2003. – Т. 26. – С. 9–20.

Описание для книг:

1) 1 автор

Ф… И.О. (авт.) Название книги. – Город: Изд-во, год. – кол-во стр.

1. Беличенко В.П. Сверхширокополосные импульсные радиосистемы. – 2-е изд., перераб. и доп. – Новосибирск: Наука, 2019. – 118 с.

2. Balanis A.C. Antenna Theory: Analysis and Design. – 4th ed. – Hoboken, NJ, USA: Wiley, 2016. – 1104 p.

2) несколько авторов

1. Тихонов А.Н. Методы решения некорректных задач / А.Н. Тихонов, В.Я. Арсенин. – 2-е изд. – М.: Наука, 2022. – 286 с.

2. Беличенко В.П. Сверхширокополосные импульсные радиосистемы / В.П. Беличенко, Ю.И. Буянов, В.И. Котов. – Новосибирск: Наука, 2015. – 588 с.

или (более 3)

3. Сверхширокополосные импульсные радиосистемы / В.П. Беличенко, Ю.И. Буянов, А.Б. Петров и др. – М.: Техносфера, 2022. – 200 с.

Описание для статей в сборнике:

1) 1 автор или более 2, 3, 4 авторов

Первый автор (Ф... И.О.). Название статьи / авторы (И.О. Ф…) // Название сб. – Город: Изд-во, год. – С. __–__.

1. Кон А.В. Результаты применения материала Арфлон AR200 для изготовления СВЧ-устройств // Наука и практика: проекты – от идеи до внедрения: матер. VIII регион. конф.: в 2 ч. – Томск: ТУСУР, 2019. – Ч. 2. – С. 146–152.

2. Elmansouri M.A. TEM horn inspired wideband antennas for diverse applications / M.A. Elmansouri, D.S. Filipovic // 2019 IEEE-APS Topical conference on Antennas and Propagation in wireless communications (APWC), 2019. – P. 016–019.

3. Васильев Е.П. Современные САПР СВЧ и их особенности / Е.П. Васильев, И.А. Ермолаев, И.М. Сомов // Современные технологии в науке и образовании (СТНО–2019): сб. трудов II Междунар. науч.-техн. форума: в 6 т. / под общ. ред. О.В. Миловзорова. – Рязань: Рязан. гос. радиотехн. ун-т, 2019. – Т. 6. – С. 170–177.

4. Нечеткая логика и искусственные нейронные сети / В.В. Круг, М.Т. Долгов, Р.Ю. Голунов и др. // Радиотехника и искусственный интеллект: сборник науч. статей. – М.: Наука, 2024. – С. 85–94.

Патенты, авторские св-ва и заявки

1. Пат. 124059 РФ, МПК H 01 R 13/646. Диэлектрическая шайба / О.Ю. Мороз (РФ), Ф.А. Михов (РФ). – № 2 012 127 060 / 07; заявл. 27.06.2012; опубл. 10.01.2013, Бюл. № 1. – 9 с.

2. Пат. 4 867 703 США, МПК H 01 R 24/542. High temperature dielectric bead for coaxial connector / J.R. Flanagan (США), D.J. Critell (США). № 233 152; заявл. 17.08.1988; опубл. 19.09.1989. – 7 с.

3. А.с. 1765238 СССР, МКИ2 С 22 С 37/10. Износостойкий чугун / М.И. Карпенко (СССР). – № 4870354/02; заявл. 02.11.90; опубл. 30.09.92, Бюл. № 36. – 4 с.

4. Св-во о гос. регистрации программы для ЭВМ № 2012660373. TALGAT 2011 / Т.Р. Газизов, А.О. Мелкозеров, Т.Т. Газизов и др. – Заявка № 2012618426. Дата поступления 5 октября 2012 г. Зарегистрировано в Реестре программ для ЭВМ 16 ноября 2012 г.

Электронные издания и сведения в публикациях из Интернета

1 автор или более 2, 3, 4 авторов –

так же, как в книгах, статьях в сборниках

1. Fuks R. New dielectric bead for millimeter-wave coaxial components // Microwave journal. – 2001 [Электронный ресурс]. – URL: https://www.microwavejournal.com/articles/3205-new-dielectric-bead-for-millimeter-wave-coaxial-components (дата обращения: 12.03.2025).

2. Clupper T. Cut-off frequency prediction for MMW coaxial interconnects / T. Clupper, Ch. Blair, J. Broomall // EDI CON online – 2019 [Электронный ресурс]. – URL: https://edicononline.com/wpcontent/uploads/sites/6/2019/10/7_Cut-off-frequency-prediction-for-MMW-coaxial-interconnects-2.pdf (дата обращения: 12.03.2025).

3. Integrated ferrite film inductor for power system-on-chip (PowerSoC) smart phone applications / J. Lee, S. Bae, Y.K. Hong, J. Jalli, J. Park, G.S. Abo, G.B.C. Choi [Электронный ресурс]. – URL: https://static.chipdip.ru/lib/663/DOC01166 3283.pdf, свободный (дата обращения: 11.12.2024).

4. Lithium Battery 72Ah. LT-spice model // Electronics Forum (Circuits, Projects and Microcontrollers) [Электронный ресурс]. – URL: https://www.electro-tech-online.com/threads/lithium-battery-72ah-ltspice-model.159067, свободный (дата обращения: 01.10.2024).

5. Vikramkumar. Bayes and naive bayes classifier / Vikramkumar, B. Vijaykumar et al.// arXiv e-prints. – 2014. – Art. no. arXiv:1404.0933. DOI:10.48550/arXiv.1404. 0933.



9 771818 044708 02007