

ТРУДЫ УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Темы номера:

- ✓ Повышение помехоустойчивости систем дальнемагистральной связи ДКМВ-диапазона
- ✓ Точность временной синхронизации оборудования при DRM-приеме
- ✓ Расположение воздушных базовых станций на основе методов глубокого обучения с подкреплением

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Научный журнал

**ТРУДЫ
УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ**

Том 11. № 2

Proceedings of Telecommunication Universities

Vol. 11. Iss. 2

Санкт-Петербург

2025

Описание журнала

Научный журнал. Включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (распоряжение Минобрнауки России № 21-р от 12.02.2019), по специальностям (распоряжение № 33-р от 01.02.2022):

1.2.2. Математическое моделирование, численные методы и комплексы программ

2.2.6. Оптические и оптико-электронные приборы и комплексы

2.2.13. Радиотехника, в том числе системы и устройства телевидения

2.2.14. Антенны, СВЧ-устройства и их технологии

2.2.15. Системы, сети и устройства телекоммуникаций

2.2.16. Радиолокация и радионавигация

2.3.1. Системный анализ, управление и обработка информации, статистика

2.3.6. Методы и системы защиты информации, информационная безопасность

Журнал позиционирует себя как научный, в связи с этим его целями являются ознакомление научной общественности (научного сообщества) с результатами оригинальных исследований, выполненных ведущими учеными и специалистами и их коллективами, а также апробация научных результатов, полученных при подготовке кандидатских и докторских диссертаций для повышения качества (уровня) проводимых исследований. Издание ставит перед собой задачу расширения инфокоммуникативного пространства взаимодействия российских и зарубежных ученых. Целевой аудиторией журнала являются ученые и специалисты-практики в области связи и телекоммуникаций и смежных направлениях науки и техники, а также профессорско-преподавательский состав и студенты, обучающиеся по программам аспирантуры, магистратуры, специалитета и бакалавриата профильных вузов и кафедр.

Выпускается с 1960 года. Выходит 6 раз в год. Издается на русском и английском языках.

Редакционный совет

Киричек Р.В. <i>Главный редактор</i>	д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия
Владыко А.Г. <i>Зам. Главного редактора</i>	к.т.н., доцент, Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия
Буйневич М.В.	д.т.н., проф., Санкт-Петербургский университет государственной противопожарной службы МЧС России, г. Санкт-Петербург, Россия
Зеневич А.О.	д.т.н., проф., Белорусская государственная академия связи, г. Минск, Республика Беларусь
Розанов Н.Н.	д.ф.-м.н., проф., чл.-корр. РАН, АО «Государственный оптический институт им. С.И. Вавилова» (ГОИ), г. Санкт-Петербург, Россия
Дукельский К.В.	д.т.н., доцент, АО «Государственный оптический институт им. С.И. Вавилова» (ГОИ), г. Санкт-Петербург, Россия
Кучерявый Е.	PhD, Технологический университет Тампере, г. Тампере, Финляндия
Каримов Б.Т.	к.т.н., доцент, Институт электроники и телекоммуникаций, Кыргызский государственный технический университет И. Раззакова (КГТУ), г. Бишкек, Кыргызстан
Тиамийу О.А.	PhD, Университет Илорина, г. Илорин, Нигерия
Козин И.Д.	д.ф.-м.н., проф., Алматинский университет энергетики и связи, г. Алма-Аты, Казахстан
Самуйлов К.Е.	д.т.н., проф., Российский университет дружбы народов (РУДН), г. Москва, Россия
Степанов С.Н.	д.т.н., проф., Московский технический университет связи и информатики (МТУСИ), г. Москва, Россия
Росляков А.В.	д.т.н., проф., Поволжский государственный университет телекоммуникаций и информатики (ПГУТИ), г. Самара, Россия
Кучерявый А.Е.	д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия
Канаев А.К.	д.т.н., проф., Петербургский университет путей сообщения имени Александра I (ПГУПС), г. Санкт-Петербург, Россия
Новиков С.Н.	д.т.н., проф., Сибирский государственный университет телекоммуникаций и информатики (СибГУТИ), г. Новосибирск, Россия
Дворников С.В.	д.т.н., проф., Военная академия связи им. Маршала Советского Союза С.М. Буденного (ВАС), г. Санкт-Петербург, Россия
Коржик В.И.	д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия
Ковалгин Ю.А.	д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия

Description

Scientific journal. The journal is included in the List of reviewed scientific publications, in which the main scientific results of dissertations for the degree of candidate of science and for the degree of doctor of science should be published (order of the Ministry of Education and Science of Russia No 21-r of 12 February 2019) in the field of (order of the Ministry of Education and Science of Russia No 33-r of 01 February 2022):

1.2.2. Mathematical modeling, numerical methods and complexes of programs

2.2.6. Optical and optoelectronic devices and complexes

2.2.13. Radio engineering, including television systems and devices

2.2.14. Antennas, microwave devices and its technologies

2.2.15. Systems, networks and telecommunication devices

2.2.16. Radiolocation and radio navigation

2.3.1. System analysis, management and information processing, statistics

2.3.6. Methods and systems of information security, cybersecurity

The journal positions itself as a scientific one, in this regard, its goals are to familiarize the scientific community (scientific community) with the results of original research carried out by leading scientists and specialists and their teams, as well as approbation of scientific results obtained in the preparation of candidate and doctoral dissertations to improve the quality (level) of ongoing research. The publication sets itself the task of expanding the infocommunicative space of interaction between Russian and foreign scientists. The target audience of the journal are scientists and practitioners in the field of communications & telecommunications and related fields of science & technology, as well as faculty and students enrolled in postgraduate, master's, specialisation and bachelor's programs of profiled universities and departments.

Since 1960. Published 6 times per year. Published in Russian and English.

Editorial Board

R.V. Kirichek DSc, prof., The Bonch-Bruевич Saint-Petersburg State University
Editor-in-chief of Telecommunications (SPbSUT), Saint-Petersburg, Russia

A.G. Vladiko PhD, associate prof., The Bonch-Bruевич Saint-Petersburg State University
Deputy editor-in-chief of Telecommunications (SPbSUT), Saint-Petersburg, Russia

M.V. Buinevich DSc, prof., Saint-Petersburg University of State Fire Service of EMERCOM of Russia,
Saint-Petersburg, Russia

A.O. Zenevich DSc, prof., Belarusian State Academy of Communications, Minsk, Republic of Belarus

N.N. Rozanov DSc, prof., member-corr. RAS, Open Joint Stock Company «S.I. Vavilov State Optical Institute» (SOI),
Saint-Petersburg, Russia

K.V. Dukel'skii DSc, associate prof., Open Joint Stock Company «S.I. Vavilov State Optical Institute» (SOI),
Saint-Petersburg, Russia

Y. Koucheryavy PhD, Tampere University of Technology, Tampere, Finland

B.T. Karimov PhD, Institute of Electronics and Telecommunications, Kyrgyz State Technical University
named after I. Razzakov, Bishkek, Kyrgyzstan

O.A. Tiarniyu PhD, University of Ilorin, Ilorin, Nigeria

I.D. Kozin DSc, prof., Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan

K.E. Samuilov DSc, prof., Peoples' Friendship University (RUDN), Moscow, Russia

S.N. Stepanov DSc, prof., Moscow Technical University of Communication and Informatics (MTUCI),
Moscow, Russia

A.V. Roslyakov DSc, prof., Povolzhskiy State University of Telecommunications and Informatics (PSUTI), Samara,
Russia

A.E. Koucheryavy DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT),
Saint-Petersburg, Russia

A.K. Kanaev DSc, prof., Emperor Alexander I-st Petersburg State Transport University (PSTU),
Saint-Petersburg, Russia

S.N. Novikov DSc, prof., Siberian State University of Telecommunications and Information Sciences (SibSUTIS),
Novosibirsk, Russia

S.V. Dvornikov DSc, prof., Military Academy of Telecommunications named after Marshal Union S.M. Budyonny,
Saint-Petersburg, Russia

V.I. Korzhik DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT),
Saint-Petersburg, Russia

Yu.A. Kovalgin DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT),
Saint-Petersburg, Russia

РЕГИСТРАЦИОННАЯ ИНФОРМАЦИЯ / REGISTRATION INFORMATION

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций: ПИ № 77-77501 от 17.01.2020 г. (пред. рег. № 77-17986 от 07.04.2004 г.)

Размещение в РИНЦ (elibrary.ru) по договору: № 59-02/2013R от 20.02.2013

Registered by Federal Service for Supervision of Communications, Information Technology and Mass Media on 17.01.2020: PI No. 77-77501 (prev. reg. on 04.07.2004: No. 77-17986)

Accommodation in RINC (elibrary.ru) by agreement on 20.02.2013: No. 59-02/2013R



Товарный знак № 929373.

Правообладатель:

Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

191186, Санкт-Петербург, наб. реки Мойки, 61, литера А

Trademark No. 929373.

Copyright holder:

Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruevich Saint-Petersburg State University of Telecommunications» (SPbSUT)

191186, St. Petersburg, emb. Moika River, 61, letter A

КОНТАКТНАЯ ИНФОРМАЦИЯ / CONTACT INFORMATION

Учредитель и издатель: Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Адрес учредителя: 191186, Санкт-Петербург, набережная реки Мойки, д. 61, литера А

Адрес редакции: 193232, Санкт-Петербург,

пр. Большевиков, 22/1, к. 334/2

Тел.: +7 (812) 326-31-63, м. т. 2022

E-mail: tuzs@sut.ru

Web: <http://tuzs.sut.ru>

ВК: <http://vk.com/spbtuzs>

Ответственный редактор **Татарникова И.М.**

Выпускающий редактор **Яшугин Д.Н.**

Дизайн: **Коровин В.М.**, изображение на обложке сгенерировано ИИ freepik <https://ru.freepik.com/ai>

Publisher: Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruevich Saint-Petersburg State University of Telecommunications» (SPbSUT)

Publisher address: 191186, Saint Petersburg, Moika river embankment, 61-A

Post address: 193232, Saint Petersburg, Prospekt Bolshevikov, 22/1

Phone: +7 (812) 326-31-63, local 2022

E-mail: tuzs@sut.ru

Web: <http://tuzs.sut.ru>

Executive Editor **Tatarnikova I.M.**

Commissioning Editor **Yashugin D.N.**

Design: **Korovin V.M.**, cover image generated by AI freepik <https://ru.freepik.com/ai>

ВЫХОДНЫЕ ДАННЫЕ / IMPRINT

Дата выхода в свет: 06.05.2025

Тираж: 1000 экз. Цена свободная.

Отпечатано в типографии

Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича»

Release date: 06.05.2025

Circulation: 1000 copies. Free price.

Printed in the printing office

Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruevich Saint-Petersburg State University of Telecommunications»



СОДЕРЖАНИЕ

CONTENTS

ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ

Грицкевич И.Ю. Гибридный метод локального контрастирования изображений с нейросетевой регулировкой параметров	7	Gritskevich I.Yu. A hybrid approach to local contrast enhancement using adaptive neural network parameter control
Гуминский О.А., Ковалгин Ю.А. О требуемой точности временной синхронизации оборудования при DRM-приеме	20	Guminsky O.A., Kovalgin Yu.A. On the Required Accuracy of Equipment Time-Synchronization at DRM Reception
Чан Т.З., Кучерявый А.Е. Оптимизация распределения пользователей, мощности и расположения воздушной базовой станции на основе методов глубокого обучения с подкреплением	32	Tran T.D., Koucheryavy A.E. User association, power allocation, and aerial base stations placement in hybrid networks: a deep reinforcement learning approach
Диязитдинова А.А. Интерактивная обработка изображений для робастного распознавания геометрических примитивов	41	Diyazitdinova A.A. Interactive image processing for robust geometric primitives recognition
Маммадов А.М. Оптимальный выбор диаметра лазерного луча в атмосферных сетях связи распределенного и мобильного типа	49	Mammadov A.M. Optimal choice of laser beam diameter in atmospheric communication networks of distributed and mobile type
Рыбаков А.И., Фокин Г.А. Разработка расчетной модели радиотрассы систем дальнемагистральной связи декаметрового диапазона для анализа параметров ионосферного канала	56	Rybakov A.I., Fokin G.A. Development of a radio wave propagation model for long-haul decameter-band communication systems in the decameter range for the analysis of ionospheric channel parameters
Блан Ф.С., Елагин В.С. Перспективная архитектура сетей, определяемых знаниями (KDN)	67	Blan F.S., Elagin V.S. Future architecture of knowledge-defined networking (KDN)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

Александров И.В., Вивчарь Р.М., Киричек Р.В. Модель представления решающих функций для метрических методов распознавания образов	84	Aleksandrov I.V., Vivchar R.M., Kirichek R.V. Decision functions model for metric methods of pattern recognition
Сабри Н.Х. Реализация электронной подписи ЕСС в ограниченных средах	101	Sabbry N.H. Implementation of an ECC digital signature technique in constrained environments
Ульянов И.В. Метод вычисления полевой свертки на основе разложения многозначного расширенного поля Галуа	109	Ulyanov I.V. Method for calculating field convolution based on the decomposition of a multi-valued extended Galois field

ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ

**2.2.6 – Оптические
и оптико–электронные приборы
и комплексы**

**2.2.13 – Радиотехника, в том числе системы
и устройства телевидения**

**2.2.14 – Антенны, СВЧ–устройства
и их технологии**

**2.2.15 – Системы, сети и устройства
телекоммуникаций**

2.2.16 – Радиолокация и радионавигация

Научная статья

УДК 621.397

<https://doi.org/10.31854/1813-324X-2025-11-2-7-19>

EDN:TKAPTM



Гибридный метод локального контрастирования изображений с нейросетевой регулировкой параметров

Иван Юрьевич Грицкевич, gritskevich.iu@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. Современные методы обработки изображений направлены на повышение их визуального качества, в частности, на адаптивное локальное контрастирование. Для достижения высокой эффективности контрастирования ранее применялись классические алгоритмы, однако они не учитывали глобальный контекст сцены и могли приводить к усилению шумовых искажений. В связи с этим в данной работе предложен гибридный метод адаптивного локального контрастирования изображений с использованием нейросетевой регулировки параметров.

Целью статьи является разработка алгоритма, обеспечивающего оптимальное усиление контраста при минимизации шумовых артефактов и искажений, повышение контрастности и точности обнаружения объектов в режиме реального времени.

Сущность решения: адаптивная настройка начальных параметров локального контрастирования с помощью сверточной нейронной сети, учитывающей яркостные и текстурные особенности. Сверточная нейронная сеть динамически подбирает параметры обработки и размеры локальных областей для объектов и фона, улучшая видимость деталей и подавляя артефакты обработки (ореолы, блочность). Метод реализован в виде программно-аппаратного комплекса для компьютерного зрения, обработки аэрофото-снимков, видеонаблюдения и поиска пострадавших при катаклизмах.

Научная новизна работы заключается в разработке алгоритма, позволяющего автоматически регулировать параметры контрастирования на основе анализа глобального и локального контекста сцены с использованием искусственного интеллекта.

Теоретическая значимость работы состоит в гибридном подходе к адаптивной обработке изображений, основанном на применении сверточной нейронной сети для управления параметрами локального контрастирования. Управление параметрами осуществляется на основе анализа текстурных и частотных характеристик изображения, автоматическую адаптацию под которые производит нейронная сеть. Методика обеспечивает адаптацию к нестационарным условиям наблюдения и, как следствие, повышает устойчивость алгоритма к сложным условиям.

Практическая значимость разработанного алгоритма определяется реализацией повышения контраста объектов изображений, полученных в видимом и инфракрасном диапазонах спектра и достоверностью их распознавания с использованием искусственного интеллекта.

Ключевые слова: контраст, искажения, оценка качества изображений, обработка изображений, нейронные сети, частотные характеристики, локальная оценка, адаптация, локальные особенности

Ссылка для цитирования: Грицкевич И.Ю. Гибридный метод локального контрастирования изображений с нейросетевой регулировкой параметров // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 7–19. DOI:10.31854/1813-324X-2025-11-2-7-19. EDN:TKAPTM

Original research

<https://doi.org/10.31854/1813-324X-2025-11-2-7-19>

EDN:TKAPTM

A Hybrid Approach to Local Contrast Enhancement Using Adaptive Neural Network Parameter Control

 Ivan Yu. Gritskevich, gritskevich.iu@sut.ru

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance. Modern image processing techniques are focused on enhancing visual quality, particularly through adaptive local contrast enhancement. Previously, classical algorithms were employed to achieve high contrast efficiency; however, these approaches failed to account for the global scene context and often led to noise amplification. This paper proposes a hybrid method for adaptive local image contrast enhancement utilizing neural network-based parameter adjustment.

The **aim** of this research is to develop an algorithm that provides optimal contrast enhancement while minimizing noise artifacts and distortions, thereby improving contrast and real-time object detection accuracy.

The **essence** of the proposed solution lies in employing a convolutional neural network for automatic configuration of local contrast parameters based on statistical brightness characteristics and textural image features. The proposed method incorporates image segmentation into local regions, analysis of their properties, and adaptive adjustment of processing parameters. This results in improved discernibility of low-contrast objects under various imaging conditions. The algorithm's operating principle is **based** on dynamically selecting local region dimensions and contrast parameters depending on background and target scene objects. The integration of a neural network module enables precise adjustment of processing parameters while minimizing undesirable artifacts such as halos and blockiness. The methodology has been **implemented** as software and hardware for an optoelectronic system designed for computer vision applications, aerial image processing, video surveillance systems, and locating victims in various disaster scenarios.

The **scientific novelty** of this work lies in the development of an algorithm that automatically regulates contrast parameters based on analysis of both global and local scene context using artificial intelligence.

The **theoretical significance** of the work consists in the development of a contrast enhancement algorithm and image quality assessment method that accounts for contrast perception characteristics by both humans and AI systems under challenging observational conditions, such as fog, smoke, low illumination, etc.

The **practical significance** of the developed algorithm is determined by its implementation of contrast enhancement for objects in images acquired in both visible and infrared spectral ranges, and by the reliability of their recognition using artificial intelligence.

Keywords: contrast, distortion, image quality assessment, image processing, neural networks, frequency characteristics, local evaluation, adaptation, local features

For citation: Gritskevich I.Yu. A Hybrid Approach to Local Contrast Enhancement Using Adaptive Neural Network Parameter Control. *Proceedings of Telecommunication Universities*. 2025;11(2):7–19. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-7-19. EDN:TKAPTM

Введение

Современные системы компьютерного зрения, медицинской диагностики, аэрофотосъемки и автономного транспорта требуют качественной обработки изображений для повышения разборчивости объектов и точности алгоритмов анализа. Однако изображения, полученные в сложных усло-

виях, часто страдают от низкой контрастности. Это может быть вызвано различными факторами, такими как неравномерное освещение сцены, когда тени или блики скрывают важные детали [1], ограниченный динамический диапазон сенсора, не позволяющий одновременно захватить яркие и темные области [2], или же свойствами самих материалов, например, низкой отражательной способностью [3].

Низкий контраст приводит к тому, что важные детали становятся неразличимыми, затрудняя обнаружение объектов [4–6], сегментацию изображений [7–9] и 3D-реконструкцию [10–12, 13]. В результате, два близко расположенных объекта могут быть восприняты как один [14].

Эффективное улучшение контраста имеет первостепенное значение для широкого спектра приложений, включая криминалистику, мониторинг дорожного движения, инспекцию промышленной безопасности [15], медицинскую визуализацию для диагностики [3], анализ спутниковых снимков для мониторинга окружающей среды и систем наблюдения, работающие в сложных условиях. В связи с этим задача адаптивного контрастирования становится критически важной для улучшения восприятия и анализа визуальной информации.

Различные методы повышения контрастности уже широко применяются в обработке изображений, однако их эффективность ограничивается рядом недостатков. Хотя глобальные методы просты в реализации [14], они часто не могут устранять локальные вариации контраста внутри изображения.

Применение однородного преобразования ко всему изображению может привести к усилению шума в одних областях и подавлению важных деталей в других [15, 16]. Это связано с тем, что локальное контрастирование, по сути, является полосовым фильтром изображения [17]. Глобальные методы, такие как гистограммная эквализация [7], не учитывают эти локальные частотные характеристики, что и приводит к нежелательным эффектам, к потере деталей и усилению шумов [18].

Локальные методы, включая CLAHE (аббр. от *англ.* Contrast Limited Adaptive Histogram Equalization) [19] и Retinex [8], адаптируются к локальному контексту, но могут вызывать артефакты, такие как блочность и ореолы [1, 3]. К распространенным недостаткам существующих алгоритмов локального контрастирования относятся усиление шума, особенно в однородных областях, появление ореолов вокруг резких краев, искажения цвета и потенциальная потеря мелких деталей, особенно в очень ярких или очень темных областях изображения [15].

Глубокие нейросетевые модели обучаются напрямую преобразовывать пиксели входного изображения в пиксели выходного, минимизируя ошибку по сравнению с эталонным изображением, и при этом сами «выясняют», как лучше всего выполнить это преобразование [12, 20, 21], но они требуют значительных вычислительных ресурсов [2]. Основной проблемой улучшения контраста на основе сверточных нейронных сетей (CNN, аббр. от *англ.* Convolutional Neural Network) является их сильная зависимость от больших объемов изображений с низким контрастом и соответствующих изображений с высоким контрастом [15].

В инфракрасном (ИК) диапазоне значительный вклад в шум вносит фоновое излучение. Это компромисс между пространственным и температурным (энергетическим) разрешением. Выбор спектрального диапазона влияет на температурный коэффициент чувствительности, поскольку интенсивность излучения черного тела и контраст между объектами и фоном зависят от длины волны.

Фундаментальной проблемой улучшения контраста является достижение значительного увеличения качества изображения без одновременного внесения нежелательных артефактов или потери ценных деталей изображения. Эти ограничения создают необходимость в разработке гибридного метода, объединяющего преимущества традиционных и нейросетевых алгоритмов. Интеграция традиционных методов обработки изображений с современными возможностями обучения моделей искусственного интеллекта предлагает перспективный путь к разработке более адаптивных, и, в конечном итоге – более эффективных методов улучшения изображений [9, 22].

Искусственный интеллект, особенно модели глубокого обучения, такие как CNN, может использоваться для анализа локальных характеристик изображения и последующей адаптивной настройки параметров в традиционных алгоритмах улучшения контраста [21, 23]. CNN фактически выполняет многомасштабный анализ, извлекая признаки на разных уровнях абстракции. Это обеспечивает более контекстно-зависимое применение этих алгоритмов. И наоборот, традиционные методы, которые часто основаны на хорошо изученных математических принципах [17, 24, 25], могут обеспечить уровень надежности и интерпретируемости, которого может не хватать чисто сквозным подходам, управляемым искусственным интеллектом.

Сравнение традиционных методов и CNN приведено в таблице 1, где используются следующие сокращения:

- MSR (аббр. от *англ.* Multi-Scale Retinex) – метод улучшения изображений (особенно контраста и цветопередачи), основанный на теории Retinex и применяющий ее на нескольких пространственных масштабах (с разными размерами окон);
- MSRCR (аббр. от *англ.* Multi-Scale Retinex with Color Restoration) – расширение MSR, добавляющее шаг восстановления цвета для коррекции искажений, которые может вносить базовый MSR;
- SSR (аббр. от *англ.* Single-Scale Retinex) – базовый вариант Retinex, применяемый только на одном пространственном масштабе (с одним размером окна) для улучшения контраста;
- HE (аббр. от *англ.* Histogram Equalization) – метод глобального контрастирования изображения путем выравнивания его гистограммы распределения яркостей.

ТАБЛИЦА 1. Сравнение методов контрастирования

TABLE 1. Comparison of Contrast Methods

Метод	Принцип	Преимущества	Ограничения	Типичные приложения	Вычислительная сложность
Контрастно-ограниченное адаптивное выравнивание гистограммы (CLAHE)	Локальное выравнивание гистограммы с ограничением контраста	Улучшение локального контраста, ограничение усиления шума, уменьшение артефактов на границах	Более сложная реализация по сравнению с HE, требует настройки параметров (размер плитки, предел отсечения)	Медицинская визуализация, подводная фотография, низкая видимость	Средняя
Многомасштабный Retinex (MSR)	Взвешенная сумма результатов SSR в разных масштабах	Лучший баланс между сжатием динамического диапазона и цветопередачей по сравнению с SSR	Может приводить к цветовым искажениям, требует выбора нескольких масштабов	Улучшение контраста и цветопередачи в различных условиях освещения	Высокая
Многомасштабный Retinex с восстановлением цвета (MSRCR)	MSR с дополнительным шагом восстановления цвета	Улучшение контраста и цветопередачи с уменьшением цветовых искажений	Может инвертировать цвета в насыщенных изображениях, требует настройки дополнительных параметров	Улучшение контраста и цветопередачи изображений, не соответствующих предположению о «сером мире»	Высокая
Сверточные нейронные сети (CNN)	Обучение сложным отображениям между изображениями с низким и высоким контрастом	Высокая эффективность, способность к изучению сложных признаков, адаптивность к различным задачам	Требуют больших объемов данных для обучения, значительных вычислительных ресурсов, интерпретируемость может быть затруднена	Улучшение изображений с низкой освещенностью (подводных, медицинских)	Очень высокая (обучение), высокая (вывод)

Одним из ключевых принципов современной обработки изображений является многомасштабный анализ, позволяющий учитывать как локальные особенности, так и глобальный контекст сцены. Этот принцип находит отражение в различных подходах, включая фрактальный анализ, вейвлет-преобразования и нейросетевые модели, извлекающие признаки на разных уровнях абстракции [6, 25]. Объединение этих сильных сторон классических подходов и современных нейросетевых алгоритмов приводит к более надежным, адаптивным и потенциально интерпретируемым решениям для локальной регулировки контраста.

Чтобы увеличить контраст гибридным методом, требуется решение следующих задач:

- разработка алгоритма с адаптивным ограничением контраста и динамическим выбором размера окна обработки;
- применение CNN для автоматической настройки параметров контрастирования;
- интеграция хроматического контрастного веса, минимизирующего цветовые искажения;
- оптимизация метода для реализации в системах реального времени (включая программируемую логическую интегральную схему (ПЛИС));
- экспериментальное тестирование на различных наборах данных (ИК изображения, аэрофото-съемка и т. д.);
- сравнение с существующими методами.

Особое внимание в статье уделено обработке ИК изображений, поскольку в этом спектральном диапазоне контрастность сцены зависит от темпера-

турных характеристик объектов. Это приводит к дополнительным проблемам, таким как температурная зависимость, фиксированные шумы сенсоров и атмосферные искажения [10, 11]. Применение в ИК съемке наиболее важно для задач поиска и спасения людей, поскольку позволяет выделить объекты с разной температурой. В ИК диапазоне контраст зависит от теплового излучения, что влияет на восприятие объектов, однако стандартные методы, например CLAHE, усиливают шум, что делает анализ менее достоверным. Помимо внешних факторов (шум сенсора и атмосферные искажения) на качество восприятия изображений влияет внутренний шум зрительной системы [6]. Этот шум обусловлен различными процессами, включая спонтанную активность и тепловые флуктуации.

Разработка адаптивного гибридного метода позволила объединить преимущества локального контрастирования и нейросетевых технологий, обеспечивая устойчивую работу в различных условиях без ручной настройки параметров, минимизацию артефактов (блочности, ореолов, усиления шума) и оптимальную вычислительную эффективность для работы в реальном времени. Таким образом, разработанный алгоритм предоставляет перспективное решение для повышения качества изображений в различных областях применения.

Предложенный метод включает несколько последовательных этапов.

Этап 1. Выбор начальных параметров скользящего окна, исходя из обстановки либо исходя из предыдущего кадра.

Этап 2. Анализ статистических характеристик внутри скользящего окна и его изменение во время обработки.

Этап 3. Адаптивное усиление контраста.

Этап 4. Оценка всего кадра с помощью собственной безэталонной метрики для регулировки последующего кадра.

Этап 5. Регулировка параметров обработки с помощью нейросетевой модели.

Разработанный алгоритм локального контрастирования имеет несколько настраиваемых параметров:

- размер окна (динамический, ограниченный диапазоном $[r_{\min}, r_{\max}]$);
- порог отсечения гистограммы T_c ;
- коэффициенты в критериях, используемых для адаптации размера окна и порога отсечения.

В ходе исследования проведен анализ, как выбор этих параметров влияет на качество выходного изображения.

Динамическое скользящее окно

При максимальном размере окна (рисунок 1) ($r_{\max} \rightarrow$ размер изображения) локальное контрастирование вырождается в глобальное. Таким образом, выходная яркость $I'(x, y)$ всех пикселей определяется глобальной функцией преобразования, рассчитанной по гистограмме всего изображения. Это максимизирует общий контраст, но теряет локальные детали. При минимальном окне ($r_{\min} = 0$) локальная гистограмма вырождается, и $I'(x, y) \approx I(x, y)$.

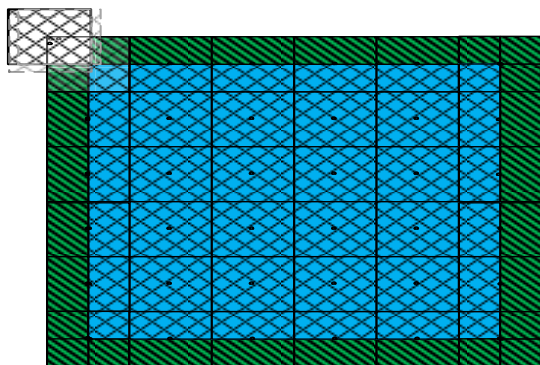


Рис. 1. Скользящее окно по кадру

Fig. 1. Sliding Window for Local Neighborhood Analysis

Если размер окна $r(x, y)$ кусочно-постоянен, то внутри области с $r(x, y) = r_0$ локальное преобразование $T_{\Omega r_0}$ эквивалентно свертке с пространственно-инвариантным оператором. $I'(x, y)$ зависит только от пикселей в окрестности $\Omega r_0(x, y)$ и не зависит от абсолютного положения (x, y) внутри области. Это преобразование непрерывно и монотонно, не создавая разрывов градиента.

На границе областей с разными размерами окон r_1 и r_2 соседние пиксели (x, y) и $(x + 1, y)$ обрабаты-

ваются в окрестностях $\Omega r_1(x, y)$ и $\Omega r_2(x + 1, y)$. Разность их выходных значений $|I'r_1(x, y) - I'r_2(x + 1, y)|$ обусловлена различием гистограмм в этих окрестностях. При значительном перекрытии окон разность мала.

Для минимизации пограничного эффекта используется взвешенное объединение:

$$I'(x, y) = \alpha(x, y) \cdot I'r_1(x, y) + (1 - \alpha(x, y)) \cdot I'r_2(x, y), \quad (1)$$

где $\alpha(x, y)$ ($0 \leq \alpha \leq 1$) – весовой коэффициент, плавно изменяющийся вблизи границы.

Динамическое окно обеспечивает локально-адаптивное контрастирование. Плавное изменение $r(x, y)$ и взвешенное объединение минимизируют артефакты на границах. Существует непрерывный переход от I' при $r_{\min} = 0$ (нет обработки) к глобальному контрастированию при $r_{\max}$.

Адаптивный размер окна

Выбран размер окна $r(x, y)$ в каждой точке изображения (рисунок 2).

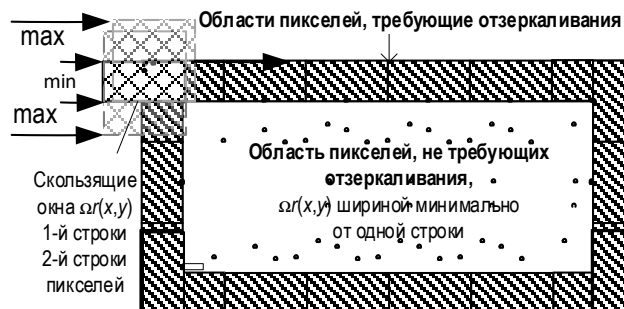


Рис. 2. Обработка изображения «скользящим окном» $\Omega r(x, y)$

Fig. 2. Processing Using the $\Omega r(x, y)$ Sliding Window

При промежуточных размерах окна возникает компромисс: меньший радиус лучше сохраняет мелкие текстуры (усиливает их контраст), но повышает шум. Предлагается использовать локальную энтропию, частотные характеристики и результаты безэталонной оценки качества как количественные показатели, определяющие размер окна. Обработка изображений производится в соответствии с алгоритмом.

Локальная энтропия

Оценка неоднородности (энтропия показывает, насколько разнообразны уровни яркости в локальной области):

$$H\Omega r(x, y) = \sum_{i=0}^N p_i \log(p_i), \quad (2)$$

где p_i – относительная частота яркости i в окрестности $\Omega r(x, y)$; N – количество пикселей в локальной области; i – уровни яркости.

Из теории информации известно: мера Хартли (энтропия источника = максимальная энтропия) – частный случай меры Шеннона (энтропия сообщения) при равновероятных символах в сообщении. Локальная энтропия используется для управления параметрами не только окна, но и контрастера. Энтропийный критерий используется для оценки неоднородности яркости в локальной области изображения, что позволяет алгоритму адаптивно реагировать на изменения текстурных характеристик сцены и динамически выбирать параметры контрастирования.

Частотный анализ

На этапе частотного анализа к фрагменту изображения $I(x + i, y + j)$ размером $(2r + 1) \times (2r + 1)$ применяется двумерное дискретное косинусное преобразование (ДКП) для оценки его спектральных характеристик.

Первый коэффициент ДКП представляет собой среднюю яркость в окне (ДК-компонента), а коэффициенты F_{uv} (коэффициенты в разложении) при больших значениях u и v соответствуют высокочастотным компонентам изображения (мелким деталям, резким перепадам яркости, шуму).

Введем $RHF(x, y)$ как функцию оценки частот в локальной области. Доля высокочастотной информации $RHF(x, y)$ оценивается как отношение энергии высоких частот к полной энергии в окне:

$$RHF(x, y) = \frac{(\sum_{u,v=1}^r (u, v) \in H |F(u, v)|^2)}{(\sum_{u,v=1}^r (u, v) \in \Omega |F(u, v)|^2)}, \quad (3)$$

где u и v – индексы частот по горизонтали и вертикали, соответственно, и $-r \leq u, v \leq r$; Ω – множество всех пар индексов (u, v) в пределах окна ДКП, т. е. H – множество пар индексов (u, v) , соответствующих высоким частотам и определяемое пороговым значением D , которое определяет границу между низкими и высокими частотами:

$$H = \{(u, v) \mid u^2 + v^2 > D^2\};$$

$|F(u, v)|^2$ – квадрат модуля коэффициента ДКП, представляющий собой энергию соответствующей частотной компоненты.

Выбор D зависит от размера максимального окна $r_{\max}$ обработки и максимальной частоты фона. Выбрано $D = r/2$, что соответствует половине максимальной частоты в обрабатываемом окне.

Безэталонная оценка качества изображения E вычисляется как взвешенное среднее величин искажений по всем пикселям:

$$E = \frac{(\sum_{k=1}^H K_{\text{иск}}) + C_1}{(N + C_1)}, \quad (4)$$

где E – результат оценки качества изображений; $K_{\text{иск}}$ – коэффициент искажения (ни в одном из вы-

шеперечисленных случаев не превышает значения 1); N – число активных локальных областей в изображении, на которые было произведено разбиение; C_1 – положительная константа (в рассматриваемом случае $C_1 = 1$), которая введена для предотвращения неопределенности при знаменателе переходящем в 0.

Итоговый результат оценки качества изображений определяется значениями: $0,1T < K_{\text{иск}} < T$ [4].

Итоговый результат регулировки окна:

$$r_{\text{target}}(x, y) = (1 - \beta) \times \frac{[r_{\min} + (r_{\max} - r_{\min}) \times (\eta(x, y) + v(x, y))]}{2} + \beta * r(x, y) \times \frac{E(x, y)}{100}, \quad (5)$$

где $\eta = \frac{(H_{\max} - H)}{H_{\max}}$ – мера энтропии; $v = 1 - RHF$ частотный коэффициент; β – весовой коэффициент для общей оценки качества и количества артефактов в предыдущем кадре.

Для устранения проблем с обработкой ИК изображений, а также для сюжета с широким динамическим диапазоном, в алгоритм добавлено уменьшение размера окна. У границ изображения и в областях с разной яркостью или сюжетом $\Omega r(x, y)$ – окно уменьшается:

$$\Omega r(x, y) \leftarrow \Omega r(x, y) \times \frac{(1 - \alpha * ||\nabla I(x, y)||)}{\max_{\text{gradient}} - ||\nabla I(x, y)||} > \text{gradient}_{\text{threshold}}, \quad \text{при } ||\nabla I(x, y)|| > \text{gradient}_{\text{threshold}}, \quad (6)$$

Ограничение контраста (задается порогом T_c) снижает коэффициент усиления до величины, пропорциональной $\frac{(L-1)T_c}{|\Omega r|}$, где $|\Omega r|$ – количество пикселей в окне. Чем меньше размер окна (и, следовательно, меньше ΔI), тем больше потенциальный коэффициент усиления шума. Коэффициент T_c выбирается в зависимости от оценки уровня шума в текущей области обработки, регулируя баланс между усилением контраста и подавлением шума:

$$|T(i) - T(j)| \leq K|i - j|, \quad (7)$$

где $K < \infty$ и определяется значением $\frac{T_c}{|\Omega r|}$.

Алгоритм обеспечивает адаптивный выбор размера окна, стремясь удержать локальную энтропию и долю высоких частот на заданном уровне, а также учитывая качество обработки на предыдущем кадре и предотвращая появление артефактов у границ.

Внедрение ИИ и использование нейросетевой модели

Ключевой особенностью разработанного метода является применение нейронных сетей для автоматической настройки параметров адаптивного

локального контрастирования (АЛК). Согласно теореме Цыбенко (универсальной аппроксимации) [11], любая непрерывная функция на компактном подмножестве конечномерного пространства может быть сколь угодно точно приближена нейронной сетью, с одним скрытым слоем, с достаточно большим числом нейронов к параметрам адаптивного локального контрастирования. Это теоретически обосновывает возможность использования нейросетей для аппроксимации нелинейных преобразований яркости и регрессии оптимальных параметров алгоритма контрастирования, что особенно важно при обработке изображений с низкой контрастностью.

Архитектура нейросетевой модели

Для реализации предложенного метода используется гибридный подход с применением CNN, состоящий из двух основных компонентов:

1) CNN для сегментации фона, основанная на архитектуре U-Net, которая благодаря своей U-образной форме с пропускными соединениями между энкодером и декодером, эффективно сочетает локальную информацию (для точной локализации границ) и глобальный контекст (для распознавания областей фона);

2) CNN для регрессии параметров АЛК.

Общая структура гибридной модели

Модель принимает на входе исходное изображение и выдает два результата:

- маску фона (вероятность принадлежности пикселей к фоновой области);
- вектор параметров АЛК (размер окна обработки и порог ограничения контраста).

Архитектура состоит из модулей (рисунок 3).

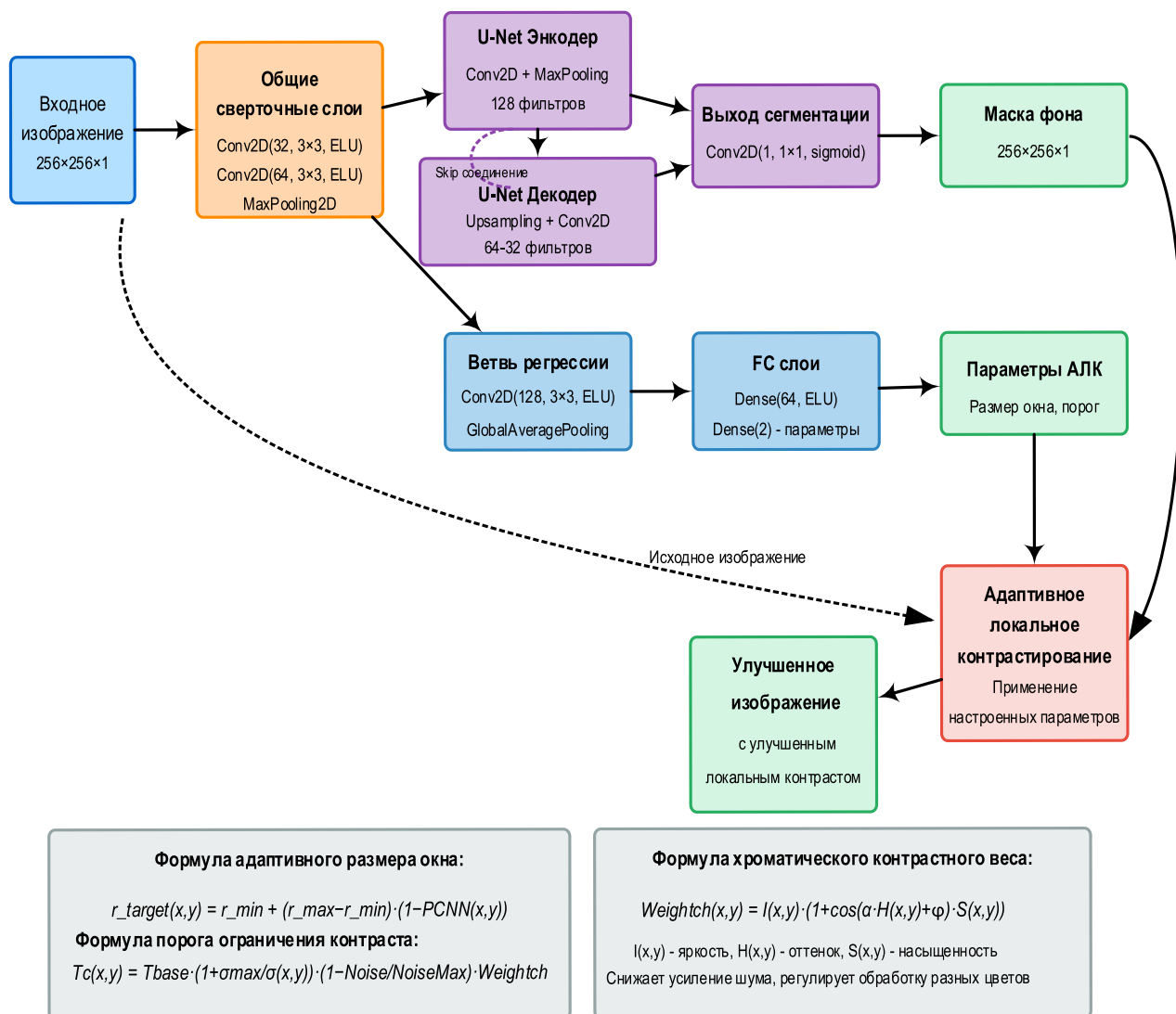


Рис. 3. Архитектура гибридной CNN для регулировки адаптивного локального контрастирования

Fig. 3. Architecture of the Hybrid Convolutional Neural Network (CNN) for Adaptive Local Contrast Enhancement Adjustment

1) Входной слой: принимает изображение в градациях серого или компоненты цветного изображения.

2) Общие сверточные слои: 3 сверточных слоев с фильтрами (количество которых составляет 32) и ядром 3×3, с функцией активации ELU для извлечения общих признаков из изображения.

3) Ветвь сегментации фона (U-Net) [13]:

– энкодер: сверточные слои и слои субдискретизации;

– декодер: сверточные слои и слои апсэмплинга;

– пропускные соединения между энкодером и декодером;

– выходной слой с сигмоидной активацией для получения маски фона.

4) Ветвь регрессии параметров:

– сверточные слои с глобальным усреднением;

– полносвязные слои;

– линейный выходной слой для предсказания параметров АЛК.

5) Skip connections соединяют соответствующие слои энкодера и декодера. На выходном слое используется сверточный слой для предсказания 4 параметров.

Обучение для видимого спектра: 150 эпох.

1) Размер батча: 99 изображений (равномерная выборка по 33 изображения из каждой подкатегории видимого спектра).

2) Оптимизатор: Стохастический градиентный спуск (SGD).

3) Начальная скорость обучения: 0,001 (скорость изменения коэффициентов слоев градиентного спуска).

4) Стратегия Learning Rate: Cosine annealing. Скорость обучения циклически изменяется по косинусоидальной функции, способствуя стабильности сходимости и предотвращению локальных минимумов. Параметры cosine annealing (период цикла, минимальная и максимальная скорость обучения) были настроены эмпирически для достижения оптимальной сходимости.

5) Функция потерь: среднее значение безэталонной оценки качества по батчу. Цель – минимизация функции потерь, т. е. улучшение качества изображения по собственной метрике безэталонной оценки качества.

CNN используется для оценки локальной контрастности изображения, определения параметров АЛК и сегментации фона. В качестве функции активации в сверточных слоях используется экспоненциальная линейная функция (ELU), которая, в отличие от ReLU, позволяет получать и отрицательные выходные значения, что полезно для представления информации о контрасте. На выходе сети формируется вектор значений полученных нейронной сетью.

Адаптивное ограничение контраста

Для предотвращения чрезмерного усиления шума применяется адаптивное ограничение контраста с порогом T_c , вычисляемым по формуле:

$$T_c(x, y) = T_{base} \cdot \left(1 + \frac{\sigma_{max}}{\sigma(x, y)}\right) \times \left(1 - \frac{Noisemax}{NoiseLevel(x, y)} \cdot Weightch(x, y)\right), \quad (8)$$

где T_{base} – базовый порог; $\sigma(x, y)$ – локальная дисперсия; σ_{max} – максимальная дисперсия; $Noisemax$ – максимальный уровень шума; $NoiseLevel(x, y)$ – оценка уровня шума; $Weightch(x, y)$ – хроматический контрастный вес.

Обучение нейросетевой модели

1) Прямой проход: Изображение -> Нейросеть -> 4 параметра -> АЛК -> Обработанное изображение.

2) Оценка качества: Обработанное изображение -> безэталонная оценка качества изображения -> Значение оценки.

3) Расчет функции потерь: Функция потерь = Значение безэталонной оценки качества изображения (среднее по батчу).

4) Обратное распространение: Вычисление градиентов функции потерь по весам CNN через backpropagation. Градиенты определяют направление изменения весов для минимизации функции потерь.

5) Обновление весов: Оптимизатор градиентный спуск обновляет веса CNN на основе градиентов для улучшения предсказания параметров АЛК, оптимизирующих значение безэталонной оценки.

Обучающая выборка

Для обучения гибридной сети использовался датасет из более чем 200 000 изображений, охватывающий видимый и ИК спектры, а также различные условия съемки. Данные разделены по спектральному диапазону.

Данные видимого Спектра (3 500 пар):

– черно-белые (1920×1200): 1000 пар (солнечно / пасмурно); центральная часть изображений без сжатия;

– черно-белые (1920×1200): 1500 пар (с / без поляризационного фильтра); центральная часть изображений без сжатия;

– цветные (1280×1024 HSV используется V-канал): 1000 пар для каждого фильтра KC10 ($\lambda > 600$ нм) и KC13 ($\lambda > 650$ нм) (красные фильтры).

Данные ИК спектра (11 500 изображений): 11 500 изображений с NUC (коррекция неоднородностей матриц с двумя кривыми второго порядка); для изображений 1280×024 использовалась центральная область 640×512 (16-битные изображения).

Для ИК и видимого диапазона были получены две отличающиеся модели нейронной сети для

настройки АЛК. Обучение модели проводилось на двух видеокартах NVIDIA Tesla K40X и заняло около 200 000 итераций.

Детализированные параметры обучения разработанной нейронной сети представлены в таблице 2.

ТАБЛИЦА 2. Параметры обучения

TABLE 2. CNN Training Parameters

Параметр	Значение	Обоснование
Эпохи	100	Обеспечивает сходимость без переобучения
Размер батча	99	Сбалансированное использование GPU
Learning rate	1e-4 → 1e-6	Используется scheduler ReduceLROnPlateau
Оптимизатор	Градиентный спуск	Снижение переобучения, стабильность
Функция потерь	Комбинированная безэталоная	Универсальность, не требует разметки
Инициализация	Xavier Uniform	Оптимальна для ELU
Активация	ELU ($\alpha = 0,1$)	Предотвращает затухание градиента
Регуляризация	Dropout 0,3 (в полносвязных слоях)	Устойчивость к переобучению
Аугментации	Различие выборок входных данных	Повышение обобщающей способности
Набор данных, без меток	> 3 500 изображений видимого спектра > 11 500 ИК изображений	Аэрофотосъемка 1920×1200 / 1280×1024, ИК кадры
Фреймворк	PyTorch + Albumentations	Быстрая отладка, качественные трансформации

Экспериментальная часть

Эффективность предложенного подхода подтверждается экспериментально с использованием объективных метрик. В качестве объективных метрик применялись:

– структурный индекс подобия (SSIM, аббр. от англ. Structural Similarity Index Measure – метрика для оценки схожести двух изображений, учитывающая структурную информацию, яркость и контраст; часто считается более близкой к человеческому восприятию, чем пиковое отношение сигнала к шуму или среднеквадратичная ошибка (MSE, аббр. от англ. Mean Squared Error): измеряет структурное сходство между двумя изображениями, принимая значения от 0 до 1, где 1 означает идеальное соответствие;

– средняя точность (mAP, аббр. от англ. mean Average Precision – средняя точность (Average Precision), усредненная по всем классам объектов или по разным порогами; стандартная метрика для оценки качества моделей обнаружения объектов (object detection): оценивает точность обнаружения объектов на изображении.

Для mAP использовались размеченные изображения:

– CityScapes (видимый спектр): 20 000 изображений 1920×1080 с людьми; оценка детекции YOLOv5 по mAP;

– FLIR (ИК спектр): 10 000 изображений 640×512 с людьми; оценка детекции YOLOv5 по mAP.

При mAP = 0,4 ошибка ϵ составляет не более 0,05 – это соответствует выборке из 20 000 изображений (в 20 раз больше 1/0,05). Ошибка ϵ , тестовая выборка $\geq 1/\epsilon$.

Для экспертной оценки качества использовался метод парного сравнения, в котором экспертам предлагалось выбрать изображение с лучшим контрастом из пары, включающей результат работы предложенного алгоритма и результат одного из следующих методов:

– CLANE: с параметрами, указанными в оригинальной работе [19];

– Multi-Scale Retinex: реализация, описанная в работе [1].

Для оценки статистической значимости различий между предложенным методом и сравниваемыми методами использовался t -критерий Стьюдента с уровнем значимости $p < 0,05$ [5].

Результат обработки

На рисунках 4a и 4b представлено изображение туманной улицы из набора CityScapes, а на рисунках 4c и 4d – туманной улицы автострადы вечером в тумане. Обработка увеличивает объективные показатели качества для малококонтрастных изображений в условиях ограниченной видимости.

При повышении точности обнаружения mAP получено 8-кратное увеличение количества обнаружений людей, а на ИК изображении – до 40 %; точность обнаружения в оригинальном ИК изображении 5 % (таблица 3).

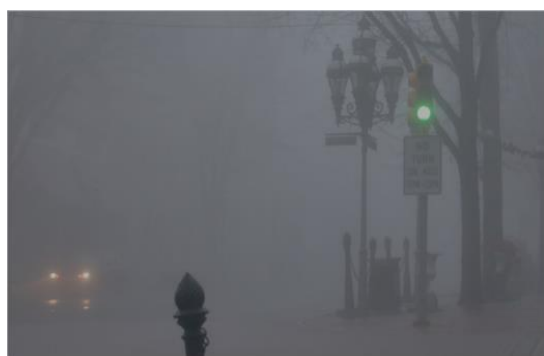
ТАБЛИЦА 3. Количественные оценки методов обработки

TABLE 3. Quantitative Metrics for Image Processing Methods

Метод (средняя оценка 20 000 CityScapes изображений)	SSIM	mAP	E (без эталона)	Экспертная оценка (по рекомендации MSE)
Оригинальные	–	0,05	0,1	0,05
CLANE	0,4	0,09	0,8	0,2
MSR	0,35	0,078	0,3	0,38
Разработанный гибридный метод	0,8	0,4	0,1	0,78

Результаты обнаружения людей на ИК изображениях с обработкой изображений методом CLANE показали излишнее усиление шума из-за чего получают ложные обнаружения (рисунок 5).

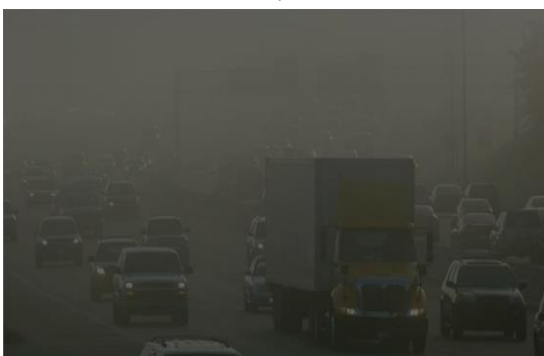
На рисунке 6 практической реализации разработанного алгоритма АЛК на ПЛИС, включая реализацию нейросетевой настройки.



a)



b)



c)



d)

Рис. 4. Результат обработки изображения: оригинальное (a, c) и обработанное гибридным алгоритмом (b, d)

Fig. 4. Processing Result: Original Image (a, c) and Image Processed with the Hybrid Algorithm (b, d)



a)



b)



c)



d)

Рис. 5. Результат обработки изображения: оригинальное ИК (a), обработанное методом Clahe (c) и гибридным алгоритмом (b, d)

Fig. 5. Processing Result: Original Infrared Image (a), Image Processed with CLAHE Method (c) and Images Processed with the Hybrid Algorithm (b, d)

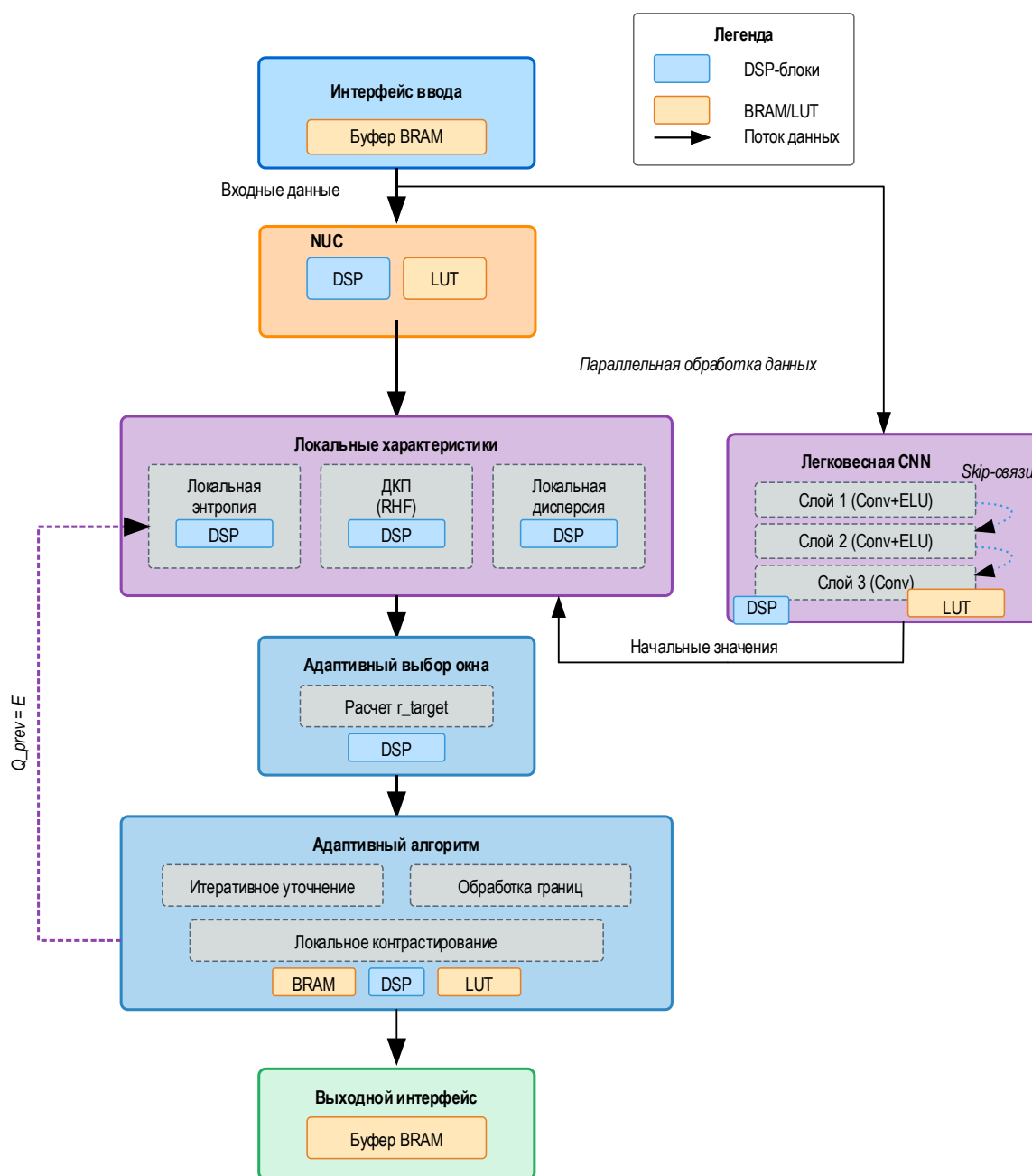


Рис. 6. Структурная схема реализации алгоритма на ПЛИС

Fig. 6. Block Diagram of the Algorithm Implementation on Field-Programmable Gate Array

Вывод

Разработанный алгоритм адаптивного локального контрастирования, основанный на динамическом определении размера окна и параметров ограничения контрастности, показал свою эффективность. Комплексный подход (АЛК + CNN) успешно протестирован на различных наборах данных, включая выборку ИК изображений (более 40 ТБ) полигонных испытаний и стандартный набор CityScapes.

Создана эффективная архитектура на ПЛИС, для реализации алгоритма с легковесной нейросетью, обеспечивающая обработку в реальном времени (65 кадров/с при разрешении 1920×1080) при потреблении менее 70 % ресурсов ПЛИС и энергопотреблении менее 5 Вт.

Применение разработанного алгоритма в качестве обработки для оценки обнаружений позволило существенно увеличить (на 40 %) обнаружение людей. Обработка в системах с использованием элементов искусственного интеллекта была успешно проверена (см. рисунки 4 и 5).

Список источников

1. Jobson D.J., Rahman Z., Woodell G.A. Properties and performance of a center/surround retinex // *IEEE Transactions on Image Processing*. 1997. Vol. 6. Iss. 3. PP. 451–462. DOI:10.1109/83.557356
2. Chen Y.S., Wang Y.C., Kao M.H., Chuang Y.Y. Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Salt Lake City, USA, 18–23 June 2018)*. IEEE, 2018. DOI:10.1109/CVPR.2018.00660
3. Paris S., Hasinoff S.W., Kautz J. Local Laplacian filtering: edge-aware image processing with a Laplacian pyramid // *Proceedings of the Conference on Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH '11, Vancouver, Canada, 7–11 August 2011)*. New York: Association for Computing Machinery, 2011. URL: https://people.csail.mit.edu/sparis/publi/2011/siggraph/Paris_11_Local_Laplacian_Filters_lowres.pdf (Accessed 25.04.2025)
4. Грицкевич И.Ю., Гоголь А.А. Алгоритм безэталонной оценки качества изображений // *Труды учебных заведений связи*. 2024. Т. 10. № 2. С. 16–23. DOI:10.31854/1813-324X-2024-10-2-16-23. EDN:TTPABW
5. Rec. ITU-R BT.500-11. Methodology for subjective assessment of the quality of television pictures. ITU-T. 2002.
6. Шелепин Ю.Е. Введение в нейроинформатику. СПб.: Троицкий мост, 2017. 352 с. EDN:YNTJRJ
7. Kim Y.T. Contrast enhancement using brightness preserving bi-histogram equalization // *IEEE Transactions on Consumer Electronics*. 1997. Vol. 43. Iss. 1. DOI:10.1109/30.580378
8. Rahman Z., Jobson D.J., Woodell G.A. Multi-scale retinex for color image enhancement // *Proceedings of the 3rd International Conference on Image Processing (Lausanne, Switzerland, 19 September 1996)*. IEEE, 1996. DOI:10.1109/ICIP.1996.560995
9. Ying Z., Li G., Gao W. A Bio-Inspired Multi-Exposure Fusion Framework for Low-light Image Enhancement // *arXiv pre-print arXiv:1711.00591*. 2017. DOI:10.48550/arXiv.1711.00591
10. Vala H.J., Baxi A. A review on Otsu image segmentation algorithm // *International Journal of Advanced Research in Computer Engineering & Technology*. 2013. Vol. 2. Iss. 2. PP. 387–389.
11. Cybenko G. Approximation by superpositions of a sigmoidal function // *Mathematics of Control, Signals and Systems*. 1989. Vol. 2. PP. 303–314. DOI:10.1007/bf02551274. EDN:OKSIPR
12. Wang R., Zhang Q., Fu C.W., Shen X., Zheng W.S., Jia J. Underexposed Photo Enhancement Using Deep Illumination Estimation // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR, Long Beach, USA, 15–20 June 2019)*. IEEE, 2019. DOI:10.1109/CVPR.2019.00701 (13)
13. Han Y., Ye J.C. Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT // *IEEE Transactions on Medical Imaging*. 2018. Vol. 37. Iss. 6. PP. 1418–1429. DOI:10.1109/TMI.2018.2823768
14. What is Histogram Equalization and how it works? // *Great Learning*. 2024. URL: <https://www.mygreatlearning.com/blog/histogram-equalization-explained> (Accessed 25.04.2025)
15. Liu J., Li D., Yuan C., Luo B., Wu G. A low-light image enhancement method with brightness balance and detail preservation // *PLoS One*. 2022. Vol. 17. Iss. 5. P. e0262478. DOI:10.1371/journal.pone.0262478. EDN:DFDSOY
16. Pizer S.M., Amburn E.P., Austin J.D., Cromartie R., Geselowitz A., Greer T., et al. Adaptive histogram equalization and its variations // *Computer Vision, Graphics, and Image Processing*. 1987. Vol. 39. Iss. 3. PP. 355–368. DOI:10.1016/S0734-189X(87)80186-X
17. Kaur M., Kaur J., Kaur J. Survey of Contrast Enhancement Techniques based on Histogram Equalization // *International Journal of Advanced Computer Science and Applications*. 2011. Vol. 2. Iss. 7. DOI:10.14569/IJACSA.2011.020721
18. Zuiderveld K. VIII.5. – Contrast Limited Adaptive Histogram Equalization // In: Heckbert P.S. (ed.) *Graphics Gems IV*. Academic Press, 1994. PP. 474–485. DOI:10.1016/B978-0-12-336156-1.50061-6
19. Lore K.G., Akintayo A., Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement // *Pattern Recognition*. 2017. Vol. 61. PP. 650–662. DOI:10.1016/j.patcog.2016.06.008
20. Wei C., Wang W., Yang W., Liu J. Deep Retinex Decomposition for Low-Light Enhancement. 2018. URL: http://39.96.165.147/Pub%20Files/2018/chen_bmvc18.pdf (Accessed 25.04.2025)
21. Liu X., Ma Y., Shi Z., Chen J. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing // *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV, Seoul, South Korea, 27 October – 02 November 2019)*. IEEE, 2019. DOI:10.1109/ICCV.2019.00741
22. Guo C., Li C., Guo J., Loy C.C., Hou J., Kwong S., Cong R. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR, Seattle, USA, 13–19 June 2020)*. IEEE, 2020. DOI:10.1109/CVPR42600.2020.00185
23. Hossain F., Alsharif M.R. Image Enhancement Based on Logarithmic Transform Coefficient and Adaptive Histogram Equalization // *Proceedings of the International Conference on Convergence Information Technology (ICCIT 2007, Gwangju, South Korea, 21–23 November 2007)*. IEEE, 2007. DOI:10.1109/ICCIT.2007.4420457
24. Stark J.A. Adaptive image contrast enhancement using generalizations of histogram equalization // *IEEE Transactions on Image Processing*. 2000. Vol. 9. Iss. 5. PP. 889–896. DOI:10.1109/83.841534
25. Потапова А.А. Новейшие методы обработки изображений. М.: ФИЗМАТЛИТ, 2008. 496 с.

References

1. Jobson D.J., Rahman Z., Woodell G.A. Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing*. 1997;6(3):451–462. DOI:10.1109/83.557356
2. Chen Y.S., Wang Y.C., Kao M.H., Chuang Y.Y. Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 18–23 June 2018, Salt Lake City, USA*. IEEE; 2018. DOI:10.1109/CVPR.2018.00660

3. Paris S., Hasinoff S.W., Kautz J. Local Laplacian filtering: edge-aware image processing with a Laplacian pyramid. *Proceedings of the Conference on Special Interest Group on Computer Graphics and Interactive Techniques, SIGGRAPH '11, 7–11 August 2011, Vancouver, Canada*. New York: Association for Computing Machinery; 2011. URL: https://people.csail.mit.edu/sparis/publi/2011/siggraph/Paris_11_Local_Laplacian.Filters_lowres.pdf [Accessed 25.04.2025]
4. Gritskevich I., Gogol A. No-Reference Image Quality Assessment Algorithm. *Proceedings of Telecommunication Universities*. 2024;10(2):16–23. (in Russ.) DOI:10.31854/1813-324X-2024-10-2-16-23. EDN:TTPABW
5. Rec. ITU-R BT.500-11. Methodology for subjective assessment of the quality of television pictures. ITU-T. 2002.
6. Shelepin Yu.E. *Introduction to Neuroiconics*. St. Petersburg: Troickij most Publ.; 2017. 352 p. (in Russ.) EDN:YNTJRI
7. Kim Y.T. Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Transactions on Consumer Electronics*. 1997;43:1. DOI:10.1109/30.580378
8. Rahman Z., Jobson D.J., Woodell G.A. Multi-scale retinex for color image enhancement. *Proceedings of 3rd IEEE International Conference on Image Processing, 19 September 1996, Lausanne, Switzerland*. IEEE; 1996. DOI:10.1109/ICIP.1996.560995
9. Ying Z., Li G., Gao W. A Bio-Inspired Multi-Exposure Fusion Framework for Low-light Image Enhancement. *arXiv preprint arXiv:1711.00591*. 2017. DOI:10.48550/arXiv.1711.00591
10. Vala H.J., Baxi A. A review on Otsu image segmentation algorithm. *International Journal of Advanced Research in Computer Engineering & Technology*. 2013;2(2):387–389.
11. Cybenko G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*. 1989;2:303–314. DOI:10.1007/bf02551274. EDN:OKSIPR
12. Wang R., Zhang Q., Fu C.W., Shen X., Zheng W.S., Jia J. Underexposed Photo Enhancement Using Deep Illumination Estimation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 15–20 June 2019, Long Beach, USA*. IEEE; 2019. DOI:10.1109/CVPR.2019.00701
13. Han Y., Ye J.C. Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT. *IEEE Transactions on Medical Imaging*. 2018;37(6):1418–1429. DOI:10.1109/TMI.2018.2823768
14. *Great Learning*. What is Histogram Equalization and how it works? 2024. URL: <https://www.mygreatlearning.com/blog/histogram-equalization-explained> [Accessed 25.04.2025]
15. Liu J., Li D., Yuan C., Luo B., Wu G. A low-light image enhancement method with brightness balance and detail preservation. *PLoS One*. 2022;17(5):e0262478. DOI:10.1371/journal.pone.0262478. EDN:DFDSOY
16. Pizer S.M., Amburn E.P., Austin J.D., Cromartie R., Geselowitz A., Greer T., et al. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*. 1987;39(3):355–368 DOI:10.1016/S0734-189X(87)80186-X
17. Kaur M., Kaur J., Kaur J. Survey of Contrast Enhancement Techniques based on Histogram Equalization. *International Journal of Advanced Computer Science and Applications*. 2011;2(7). DOI:10.14569/IJACSA.2011.020721
18. Zuiderveld K. VIII.5. – Contrast Limited Adaptive Histogram Equalization. In: Heckbert P.S. (ed.) *Graphics Gems IV*. Academic Press; 1994. p.474–485. DOI:10.1016/B978-0-12-336156-1.50061-6
19. Lore K.G., Akintayo A., Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement. *Pattern Recognition*. 2017;61:650–662. DOI:10.1016/j.patcog.2016.06.008
20. Wei C., Wang W., Yang W., Liu J. *Deep Retinex Decomposition for Low-Light Enhancement*. 2018. URL: http://39.96.165.147/Pub%20Files/2018/chen_bmvc18.pdf [Accessed 25.04.2025]
21. Liu X., Ma Y., Shi Z., Chen J. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. *Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV, 27 October – 02 November 2019, Seoul, South Korea*. IEEE; 2019. DOI:10.1109/ICCV.2019.00741
22. Guo C., Li C., Guo J., Loy C.C., Hou J., Kwong S., Cong R. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 13–19 June 2020, Seattle, USA*. IEEE; 2020. DOI:10.1109/CVPR42600.2020.00185
23. Hossain F., Alsharif M.R. Image Enhancement Based on Logarithmic Transform Coefficient and Adaptive Histogram Equalization. *Proceedings of the 2007 International Conference on Convergence Information Technology, ICCIT 2007, 21–23 November 2007, Gwangju, South Korea*. IEEE; 2007. DOI:10.1109/ICCIT.2007.4420457
24. Stark J.A. Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Transactions on Image Processing*. 2000;9(5):889–896. DOI:10.1109/83.841534
25. Potapova A.A. *The Latest Methods of Image Processing*. Moscow: FIZMATLIT Publ.; 2008. 496 p. (in Russ.)

Статья поступила в редакцию 18.03.2025; одобрена после рецензирования 14.04.2025; принята к публикации 21.04.2025.

The article was submitted 18.03.2025; approved after reviewing 14.04.2025; accepted for publication 21.04.2025.

Информация об авторе:

**ГРИЦКЕВИЧ
Иван Юрьевич**

аспирант кафедры телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича

 <https://orcid.org/0009-0007-4916-1816>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.

Научная статья

УДК 621.396.712

<https://doi.org/10.31854/1813-324X-2025-11-2-20-31>

EDN:LQOXKV



О требуемой точности временной синхронизации оборудования при DRM-приеме

Олег Андреевич Гуминский ✉, guminskii.oa@sut.ruЮрий Алексеевич Ковалгин, kowalgin@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. В настоящее время в России и в мире идет активный переход от аналоговых к цифровым технологиям. Этот процесс в полной мере относится и к радиовещанию. В России разрешено экспериментальное цифровое радиовещание в диапазоне ОВЧ в трех форматах: DRM, DAB, PAVIS. Однако решения Правительства о выборе одного из них в качестве национального стандарта пока нет, что объясняется в первую очередь недостаточной полнотой сведений, необходимых для производства и эксплуатации оборудования, относящего в первую очередь к системе DRM, единственной системе цифрового радиовещания, рекомендованной ITU-R для применения во всех полосах частот, выделенных для наземного радиовещания. Данная работа восполняет этот пробел в части, относящейся к разработке приемного оборудования.

Цель работы: оценка требуемой точности момента старта прямого дискретного преобразования Фурье относительно начала полезной части OFDM-символов, при котором сигнал ошибки для системы DRM при работе в режиме устойчивости E не превышает значений, требуемых стандартом и рекомендацией ITU-R BS.1660-8 (06/2019).

Методы. В качестве основы для проведения данных исследований выбрана имитационная модель приемопередающего тракта системы DRM для режима устойчивости E. Она дополнена блоками, позволяющими для разных видов модуляции и уровней защиты изменять в приемнике временной сдвиг между моментом старта прямого дискретного преобразования Фурье и началом полезной части OFDM-символа, а также при каждом значении временного сдвига оценивать качество принимаемого DRM-приемником сигнала.

Результаты. Критериями оценки качества принимаемого OFDM-сигнала выбраны коэффициент модуляционных ошибок и вероятность появления битовых ошибок. Показано, что для сохранения условий комфортного приема максимально допустимое временное рассогласование между моментами старта прямого дискретного преобразования Фурье и началом полезной части OFDM-символа при модуляции поднесущих частот QAM-4 не должно превышать значений 1,8–2,3 мкс; при модуляции QAM-16 – эта величина должна быть не более 0,8–1,3 мкс. Разброс полученных значений определяется уровнем защиты PL. Основное влияние на требуемую точность временной синхронизации DRM-приемника при работе в режиме устойчивости E оказывает порядок модуляции QAM и существенно меньшее влияние – уровень защиты PL в области, где вероятность появления битовых ошибок не превышает значение 10^{-4} . При $BER \leq 10^{-4}$ влиянием выбранного уровня защиты PL можно пренебречь. **Полученные результаты являются новыми** для системы DRM при работе в режиме устойчивости E (диапазон ОВЧ).

Теоретическая / практическая значимость. Предложен метод оценки допустимого временного рассогласования OFDM-сигнала при DRM-приеме. Учет полученных результатов необходим для разработки блока временной синхронизации приемников цифрового радиовещания стандарта DRM.

Ключевые слова: цифровое радиовещание, DRM, временная синхронизация

Ссылка для цитирования: Гуминский О.А., Ковалгин Ю.А. О требуемой точности временной синхронизации оборудования при DRM-приеме // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 20–31. DOI:10.31854/1813-324X-2025-11-2-20-31. EDN:LQOXKV

Original research

<https://doi.org/10.31854/1813-324X-2025-11-2-20-31>

EDN:LQOXXV

On the Required Accuracy of Equipment Time-Synchronization at DRM Reception

 **Oleg A. Guminsky** ✉, guminskii.oa@sut.ru

 **Yuri A. Kovalgin**, kowalgin@sut.ru

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Abstract

Relevance. Currently in Russia and in the world there is an active transition from analog to digital technologies. This process is also fully applicable to radio broadcasting. In Russia, experimental digital radio broadcasting (DRB) is authorized in the VHF band in three formats: DRM, DAB, and RAVIS. However, the Government has not yet decided to choose one of these formats as the national standard, which is explained primarily by the insufficient completeness of the information required for the production and operation of equipment related primarily to the DRM system, the only DRB system recommended by ITU-R for being used in all frequency bands allocated for terrestrial radio broadcasting. This paper fills this gap as it relates to the development of receiving equipment.

Purpose of this work: estimation of the required accuracy of the start moment of the direct discrete Fourier transform with respect to the beginning of the useful part of the OFDM symbols, where the error signal for the DRM system, while operating in the E stability mode, does not exceed the values required by the ITU-R standard and recommendation BS.1660-8 (06/2019).

Methods. As a basis to execute this research, the simulation model of the DRM system transceiver path for the stability mode E, was chosen. It is supplemented with modules that allow, for different types of modulation and protection levels, to change the time shift in the receiver between the start of the direct discrete Fourier transform and the beginning of the useful part of the OFDM-symbol, as well as to evaluate, at each value of the time shift, quality of the signal received by the DRM-receiver.

Results. The modulation error rate (MER) and bit error rate (BER) were chosen as the criteria for evaluating the quality of the received OFDM signal. It is shown that, in order to maintain conditions of a comfortable reception, the maximum allowable time mismatch between the start of the direct discrete Fourier transform and the beginning of the useful part of the OFDM-symbol, at modulation of subcarrier frequencies QAM-4, should not exceed 1,8 - 2,3 μ s; at modulation QAM-16 this value should not exceed 0,8-1,3 μ s. The range of values obtained is determined by the protection level PL. The main impact on the required accuracy of time synchronization of a DRM-receiver, while operating in the stability mode E, has the order of modulation QAM; significantly less influence has the level of protection PL if the probability of bit errors does not exceed 10^{-4} . For $BER \leq 10^{-4}$, the influence of the chosen level of protection PL can be neglected. **The obtained results are new** for the DRM system while operating in the E stability mode (VHF band).

Theoretical / Practical significance. The method of estimation of the acceptable time mismatch of OFDM-signal at DRM-reception is proposed. The results are necessary for the development of the time synchronization unit of digital radio broadcasting DRM-receivers.

Keywords: digital broadcasting, DRM, time synchronization

For citation: Guminsky O.A., Kovalgin Yu.A. On the Required Accuracy of Equipment Time-Synchronization at DRM Reception. *Proceedings of Telecommunication Universities*. 2025;11(2):20–31. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-20-31. EDN:LQOXXV

Введение

Технология OFDM широко используется в радио- и мобильной связи, включая и системы цифрового радиовещания, такие как DAB, DRM, IBOC HD-Radio,

РАВИС, CDR. При передаче OFDM-сигнал претерпевает целый ряд изменений, что может приводить к появлению межсимвольной интерференции и интерференции между поднесущими частотами.

Межсимвольная интерференция возникает в ситуациях, когда:

- на приемник, помимо основного сигнала, приходят его отраженные копии на интервале времени, превышающем длительность защитного интервала OFDM-символа;
- имеет место смещение OFDM-символа по времени (STO, аббр. от англ. Symbol Time Offset); появляется при наличии временного рассогласования переданного и принимаемого OFDM-сигнала в случае, когда эта величина меньше длительности импульсного отклика радиоканала.

Интерференция между поднесущими частотами возникает, когда:

- приемник из-за наличия временного рассогласования с передатчиком не точно оценивает начало полезной части OFDM-символа при старте прямого дискретного преобразования Фурье (ДПФ), т. е. достоверно не знает, с какого дискретизированного отсчета начинается полезная часть OFDM-символа;
- имеет место смещение несущей частоты радиоканала (CFO, аббр. от англ. Carrier Frequency Offset); возникает при рассогласовании частот тактовых генераторов передатчика и приемника.

Окно прямого ДПФ может иметь несколько вариантов расположения внутри OFDM-символа (рисунок 1).

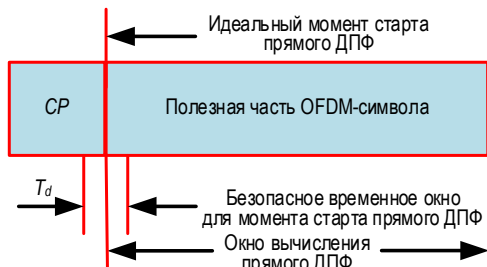


Рис. 1. Возможные варианты расположения окна прямого ДПФ относительно OFDM-символа, CP – защитный интервал OFDM-символа

Fig. 1. Possible Variants of the Direct DFT Window Location Relative to the OFDM-Symbol, CP – Guard Interval of the OFDM-Symbol

Безопасная точка старта прямого ДПФ должна располагаться после прихода импульсного отклика радиоканала. Это необходимо для отсутствия межсимвольной интерференции. Допустимы некоторые неточности T_d в моменте старта прямого ДПФ [1], при которых принятый OFDM-сигнал будет еще свободен от межсимвольной интерференции. Это позволяет говорить о разработке алгоритма оценки величины временного рассогласования с приемлемой точностью.

Известно большое число работ, посвященных исследованию проблемы временной синхронизации OFDM-сигналов, среди которых, прежде всего, следует назвать такие работы, как [1–7]. Однако анализ их содержания показывает, что они носят об-

щий характер, не учитывают особенностей структуры и характеристик OFDM-сигналов конкретных систем цифрового радиовещания, не предлагают для каждой из них наиболее эффективные (с учетом особенностей конкретного стандарта) решения.

Цель работы. Данная работа посвящена исследованию и оценке требуемой точности момента старта прямого ДПФ относительно начала полезной части OFDM-символов, при которой вероятность появления битовой ошибки (BER, аббр. от англ. Bit Error Ratio) для системы DRM не превышает значений, требуемых стандартом [8] и рекомендацией ITU-R BS.1660-8 (06/2019) [9].

Основные результаты

В качестве основы для проведения данных исследований была выбрана разработанная ранее в среде Matlab / Simulink имитационная модель приемопередающего тракта системы DRM для режима устойчивости E , выполненная в полном соответствии с требованиями стандарта [8]. Ее укрупненная структура приведена на рисунке 2а, а в развернутом подробном виде с пояснениями в [10]. Она была дополнена блоками, позволяющими для разных видов модуляции и уровней защиты изменять в обе стороны в приемнике временной сдвиг τ между моментом старта прямого ДПФ и началом полезной части OFDM-символа (рисунок 2б), а также при каждом значении временного сдвига τ оценивать качество принимаемого DRM-приемником сигнала (рисунок 2с). На выходе канального кодера (см. рисунок 2б) стоят блоки формирования OFDM-символа, после них стоит блок Sample Shifter, отвечающий за сдвиг принимаемого сигнала относительно старта прямого ДПФ, затем – блок AWGN Channel для ввода в канал аддитивного белого гауссовского шума (АБГШ) и OFDM-демодулятор.

Естественной величиной для оценки искажений OFDM-сигналов является мера отклонения точек сигнального созвездия поднесущих частот от своих истинных положений при каждом виде модуляции и уровне защиты [11–14]. С учетом этого критерием для оценки качества OFDM-сигнала изначально был выбран коэффициент модуляционных ошибок (MER, аббр. от англ. Modulation Error Ratio).

В документах Международного института телекоммуникационных стандартов (ETSI, аббр. от англ. European Telecommunications Standards Institute) величина MER [12] была впервые введена для оценки качества OFDM-сигнала, переданного передатчиком. В более поздней редакции [13] для сигналов DRM-передатчиков, работающих в диапазоне ОВЧ (30–3000 МГц), было установлено, что ее значение должно быть не менее 21 дБ.

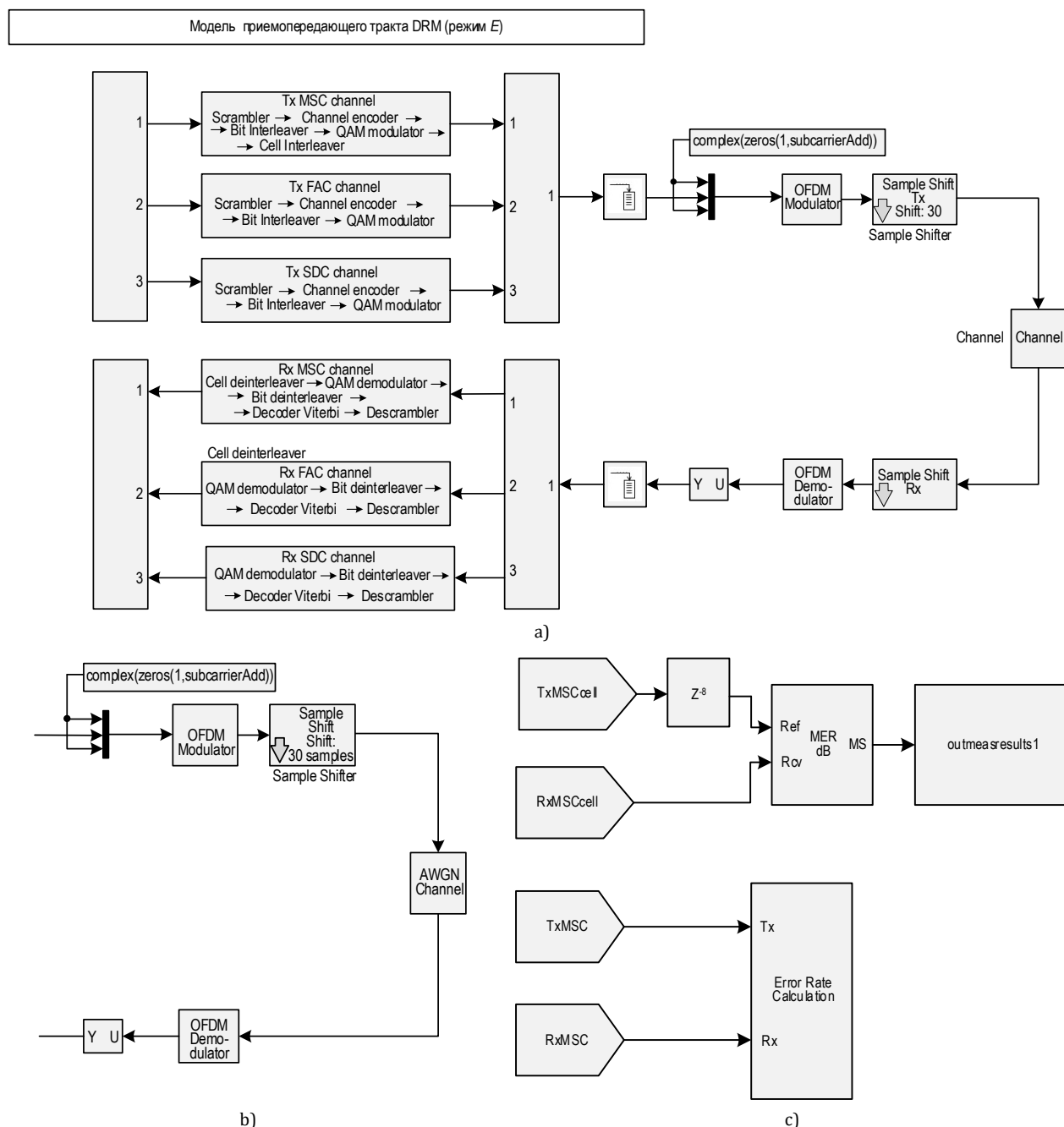


Рис. 2. Укрупненная схема имитационной модели для исследования требуемой точности временной синхронизации:
 а) укрупненный приемопередающий тракт системы DRM при работе в режиме устойчивости E; б) дополнительные блоки, необходимые для введения временного сдвига между моментом старта прямого ДПФ и началом полезной части OFDM-символа (Sample Shifter) и аддитивного белого гауссовского шума (AWGN Channel); в) блоки для оценки коэффициента модуляционных ошибок (блок оценки MER (dB), принимающий на вход символы модуляции с блоков Tx/Rx MCS Cell) и вероятности появления битовых ошибок BER (блок Error Rate Calculation, принимающий поток бит с блоков Tx/Rx MSC)

Fig. 2. Enlarged Scheme of the Simulation Model to Study the Required Accuracy of Temporal Synchronization: a) Enlarged Transceiver Path of the DRM System When Operating in the E Stability Mode; б) Additional Blocks Required to Introduce a Time Shift between the Start of the Direct Discrete Fourier Transform (DFT) and the Beginning of the Useful Part of the OFDM Symbol (Sample Shifter) and Additive White Gaussian Noise (AWGN Channel); в) Blocks for Estimation of Modulation Error Rate, MER (dB), Estimation Block, Receiving Modulation Symbols from Tx/Rx MCS Cell Blocks as Input) and Bit Error Rate BER (Error Rate Calculation Block, Receiving Bit Stream from Tx/Rx MSC Blocks)

В отличие от изложенного выше измерения, связанные с оценкой качества OFDM-сигнала, выполнялись в имитационной модели не на выходе ее передающей части, а в DRM-приемнике перед декоде-

ром. В нашем случае, где речь идет о величине коэффициента модуляционных ошибок, измеренных в приемнике, используется для его обозначения аббревиатура RMER (от слова «receive»). Кроме того,

для чистоты эксперимента было принято, что при DRM-приеме имеет место только временной сдвиг момента старта прямого ДПФ относительно начала полезной (информационной) части OFDM-символов, т. е. отсутствуют искажения, вызванные сдвигом несущей частоты, потерями в радиоканале, многолучевостью, а также рассогласованием периодов дискретизации цифроаналогового преобразования при преобразовании цифрового сигнала в аналоговый в передающем тракте и аналого-цифрового преобразования – в DRM-приемнике.

Исследования проводились в два этапа:

- 1) АБГШ в радиоканале отсутствовал (идеальные условия приема) при разных значениях τ ;
- 2) оценивалось влияние АБГШ в радиоканале на изменение условий приема при наличии временного сдвига τ между моментами старта прямого ДПФ и началом полезной части OFDM-символа (см. рисунок 1).

Напомним также, что RMER – это отношение средней мощности полезного (эталонного) сигнала P_c к средней мощности сигнала ошибки $P_{\text{ош}}$ за интервал наблюдения T , выраженное в дБ [11–14]. Можно сказать, что это аналог параметра *отношение сигнал / шум* (SNR, аббр. от англ. Signal-to-Noise Ratio) для OFDM-сигналов. При этом, если положения точек переданных сигнальных созвездий поднесущих частот оцениваются на DRM-передатчике, то для его обозначения используется аббревиатура MER, а при его оценке после их восстановления в DRM-приемнике использована аббревиатура RMER. Значения MER и RMER совпадают, когда имеют место идеальные условия для передающего и приемного трактов системы DRM.

В [15] предложено аппроксимирующее выражение, позволяющее по известному значению RMER вычислить соответствующее ему по мощности отношение SNR:

$$\text{SNR} = a + a_1(\text{RMER}) + a_2(\text{RMER})^2 + a_3(\text{RMER})^3 + \dots, \text{ (дБ)}. \quad (1)$$

Значения коэффициентов аппроксимирующего полинома зависят от вида модуляции и равны значениям, указанным в таблице 1, где: σ – среднеквадратическое отклонение; D – диапазон возможных значений SNR, в пределах которого расчетные значения, найденные по выражению 1, соответствуют экспериментальным данным.

ТАБЛИЦА 1. Значения коэффициентов аппроксимирующего полинома [15]

TABLE 1. Values of the Approximating Polynomial Coefficients [15]

Модуляция	a_0	a_1	a_2	a_3	σ , дБ	D , дБ
QAM-4	-2,672	1,518	-0,023	0	0,02	0–7
QAM-16	5,559	-2,959	0,455	-0,014	0,1	0–15

Заметим (рисунок 3), что выражение 1 дает достоверные результаты для модуляции QAM-4 при значениях SNR, лежащих в интервале от 0 до 7 дБ с величиной среднеквадратического отклонения, не превышающей 0,02 дБ; при модуляции QAM-16 эти значения – в интервале 0–15 дБ, при $\sigma \leq 0,1$ дБ и при отсутствии помехоустойчивого кодирования.

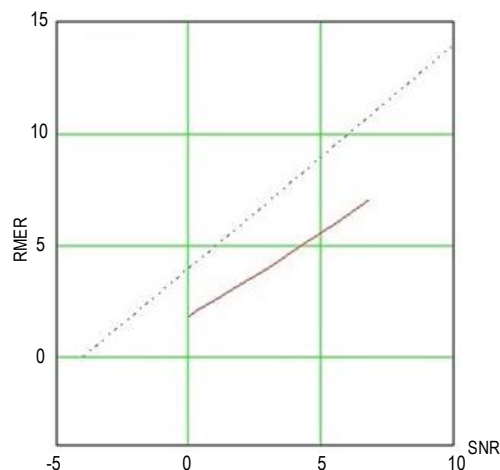


Рис. 3. Зависимости коэффициента RMER, дБ, от SNR при $\tau = 0$ мкс

Fig. 3. Dependences of the RMER, dB, on the SNR, at $\tau = 0$ μ s

Красная прямая линия (см. рисунок 3) – значения RMER для модуляции QAM-4 без учета помехоустойчивого кодирования (результат расчета по (1)). Верхняя прямая линия – значения RMER для модуляции QAM-4 при уровне защиты PL1 (скорость кода $R = 0,333$) для системы DRM при работе в режиме устойчивости E с помехоустойчивым (канальным) кодированием (экспериментальные данные, полученные в работе). Канальное кодирование, примененное в системе DRM, обеспечивает существенно более низкое значение SNR, необходимое для комфортного прослушивания. Например, при модуляции QAM-4 и уровне защиты PL1 величина RMER составляет 1,84 и 3,91 дБ при SNR = 0 дБ и, соответственно, 6,9 и 10,75 дБ при SNR = 6,8 дБ, т. е. увеличивается за счет канального кодирования на 2,07 дБ при SNR = 0 дБ и на 3,85 дБ при SNR = 6,8 дБ и, как следствие, при канальном кодировании требуется меньшее значение SNR для комфортного прослушивания. Иначе говоря, если при модуляции QAM-4 и уровне защиты PL1 минимальное значение RMER составляет 6 дБ, то канальное кодирование обеспечивает получение данного значения уже при SNR = 2,13 дБ. Из рисунка 3 следует, что, если величина SNR = 2,71 дБ, то требуемое для комфортного прослушивания значение RMER должно составлять не менее 6,6 дБ. Это хорошо согласуется с экспериментальными данными (рисунок 4, верхняя прямая линия), где требуемое минимальное значение RMER составляет $\sim 6,5$ дБ. Здесь модуляция поднесущих частот QAM-4, скорость кода $R = 0,333$.

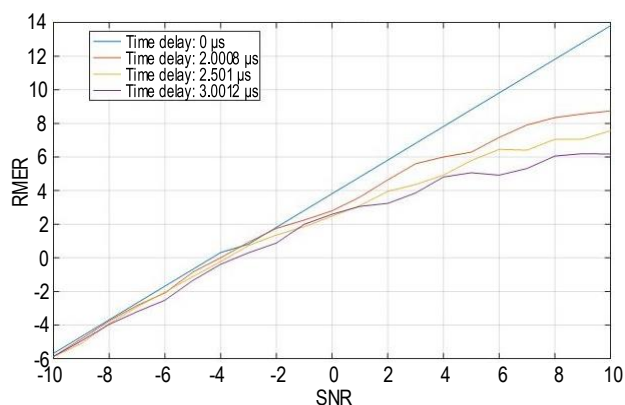


Рис. 4. Зависимости RMER, дБ, при наличии в радиоканале АБГШ от величины SNR, дБ, для разных значений временного сдвига τ окна прямого ДПФ относительно начала полезной части OFDM-символа

Fig. 4. Dependences of RMER, dB, in the Presence of AWGN in the Radio Channel on the Value of SNR, dB, for Different Values of Time Shift τ of the Direct Discrete Fourier Transform (DFT) Window Relative to the Beginning of the Useful Part of the OFDM-Symbol

Конечно, даже при выполнении изложенных выше условий на величину RMER могут влиять как способ передачи (OFDM-передача или передача на одной несущей частоте), виды модуляции и помехоустойчивого кодирования, возможно, другие факторы. Данное исследование не имеет своей целью оценку влияния каждого из них на величину RMER. Это предмет отдельного исследования.

Результаты оценки RMER, полученные при выполнении первого этапа исследований, приведены на рисунке 5.

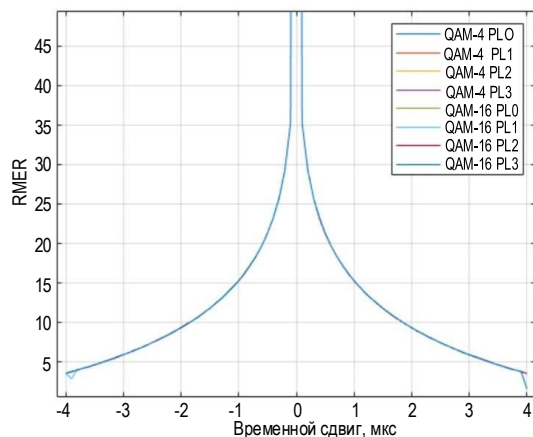


Рис. 5. Зависимости RMER, дБ, от временного сдвига τ между моментом старта прямого ДПФ и началом полезной части OFDM-символа при модуляции поднесущих частот QAM-4 и QAM-16, разных уровнях помехозащитности PL0–PL3 и отсутствии в радиоканале АБГШ, помех и многолучевости

Fig. 5. Dependence of RMER, dB, on the Time Shift τ between the Start of the Direct DFT and the Beginning of the Useful Part of the OFDM-Symbol at QAM-4 and QAM-16 Modulation of Subcarriers Frequencies and Different Protection Levels PL0-PL3 in the Absence of AWGN, Interference and Multipath in the Radio Channel

На рисунке 5 оценено влияние временного сдвига τ между моментами старта прямого ДПФ и

началом полезной части OFDM-символа на величину RMER для всех видов модуляции и уровней защиты при отсутствии в радиоканале АБГШ, помех и многолучевости. Величина RMER не зависит от порядка модуляции QAM поднесущих частот и уровня защиты OFDM-сигнала при отсутствии гауссовского шума, помех и многолучевости в радиоканале. Результат очевиден, ибо это идеальные условия приема, когда наличие канального кодирования не влияет на получаемые результаты. С появлением временного сдвига τ мощность сигнала ошибки растет и тем значительно, чем больше τ , что приводит к резкому уменьшению величины RMER. Отметим, что для величины RMER в стандарте на систему DRM не задано пороговое значение, при котором нарушается условие комфортного приема. С этой точки зрения полезен такой критерий, как вероятность появления BER. Для условий комфортного DRM-приема – $BER \leq 10^{-4}$ [9].

Заметим, что погрешность при измерении значений RMER и BER существенно зависит от числа обработанных OFDM-символов [15]: например, для подсчета BER с вероятностью не хуже 10^{-3} требуется обработка не менее 10^4 битов испытательного сигнала, для измерения BER с вероятностью не хуже 10^{-6} требуется обработка уже не менее 10^7 битов. В рассматриваемом случае при моделировании испытательный сигнал содержал 10^6 битов.

На рисунке 6а приведена полученная экспериментальным путем зависимость RMER от величины временного сдвига τ для модуляции QAM-4 и уровня защиты PL1 (скорость кода $R = 0,333$). На рисунке 6б дана экспериментальная зависимость вероятности появления BER также для QAM-4 и $R = 0,333$ от величины временного сдвига τ . На рисунке 6в показан вид сигнальных созвездий для трех значений τ (1, 2 и 3 мкс). Все эти данные получены при отсутствии в радиоканале АБГШ, помех и многолучевости. На рисунке 6 (справа) приведены аналогичные данные, но для модуляции QAM-16 и уровне защиты PL2 (скорость кода $R = 0,5$).

На рисунке 7 даны аналогичные сведения для разных видов модуляции и уровней защиты, которые возможны при работе системы DRM в режиме устойчивости E.

Параметрами каждой кривой на рисунках 6, 7 являются порядок модуляции QAM поднесущих частот OFDM-сигнала и уровень защиты PL.

Полученные для идеальных условий приема результаты позволяют сделать следующие выводы при работе системы DRM в режиме устойчивости E.

Во-первых, ошибка в оценке момента начала полезной части OFDM-символа при выполнении прямого ДПФ приводит для каждого выбранного вида модуляции (QAM-4, QAM-16) и уровня защиты (PL0–PL3) к росту BER (см. рисунки 6с, 6д и 7), ибо

сопровождается увеличением сигнала ошибки и, как следствие, ухудшением условий приема и сокращением радиуса зоны покрытия из-за порогового эффекта, свойственного системе DRM. Это явление

проявляется тем значительнее, чем больше величина временного сдвига τ окна прямого ДПФ от его идеального положения, когда $\tau = 0$ (см. рисунок 1).

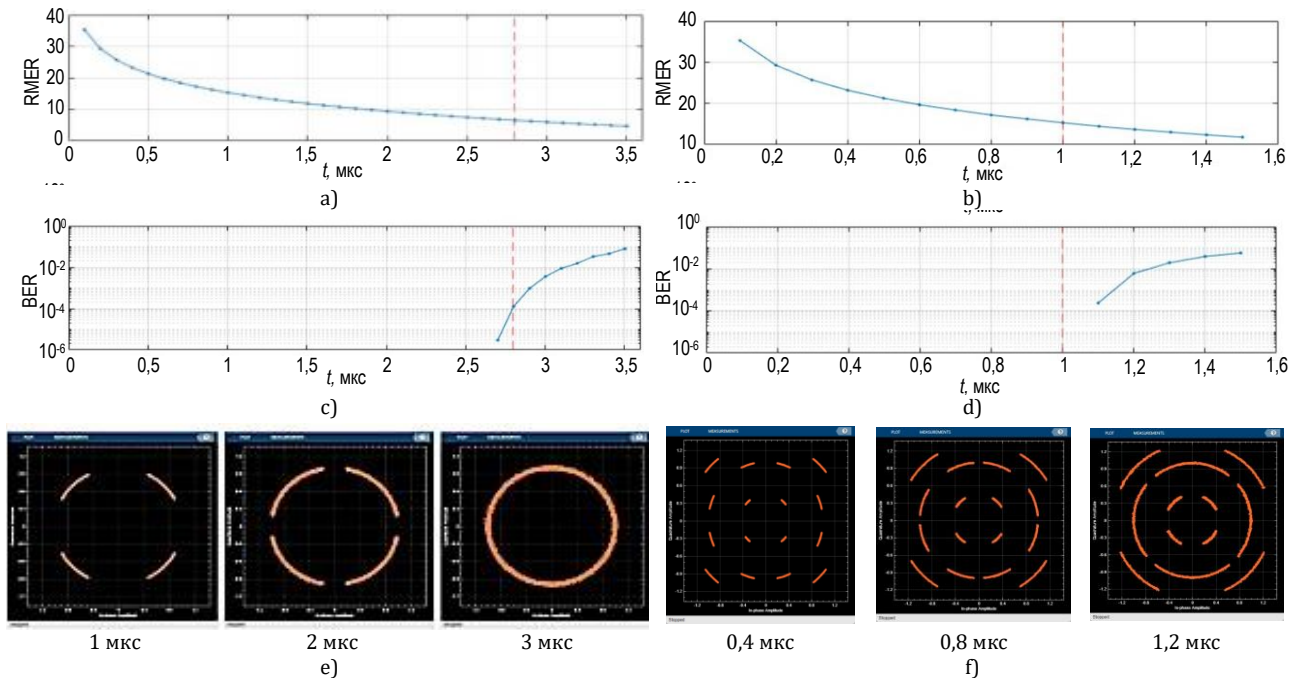


Рис. 6. Влияние временного сдвига τ между началом полезной части OFDM-символа и моментом старта прямого ДПФ на значение RMER, дБ, (a, b), на вероятность появления BER (c, d) и на вид сигнала созвездия (e, f) [19]

Fig. 6. Effect of the Time Shift τ between the Beginning of the Useful Part of the OFDM-Symbol and the Start of the Direct DFT on the Value of RMER, dB, (a, b), on the Probability of BER (c, d), and on the Type of Signal Constellation (e, f) [19]

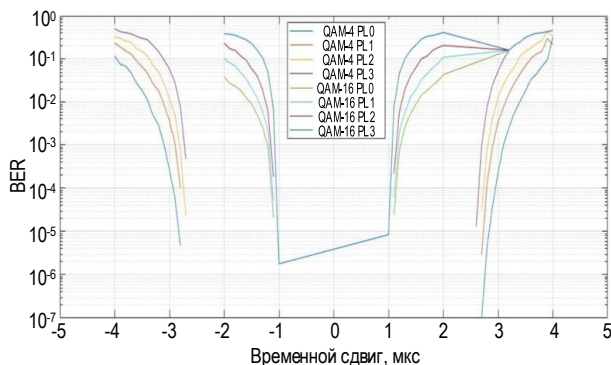


Рис. 7. Зависимости вероятности появления BER от временного сдвига τ между началом полезной части OFDM-символа и моментом старта прямого ДПФ

Fig. 7. Dependences of the Probability of BER on the Time Shift τ between the Beginning of the Useful Part of the OFDM-Symbol and the Moment of the Start of the Direct DFT

Во-вторых, наличие временного сдвига τ , вызванное неточностью временной синхронизации приемника, приводит к искажению (повороту) сигнальных созвездий поднесущих частот (см. рисунки 6e и 6f) и, следовательно, к дополнительному увеличению уровня шума при DRM-приеме. Это явление проявляется тем значительнее, чем больше число позиций вектора поднесущих частот при модуляции QAM.

В-третьих, величина BER (см. рисунок 7), в отличие от RMER, является не только функцией временного сдвига τ , но зависит также от вида модуляции и в меньшей степени – от уровня защиты. С уменьшением BER зависимость от уровня защиты уменьшается и при величине $BER \leq 10^{-4}$ становится незначительной, при этом существенным влияющим на эту величину фактором остается порядок модуляции QAM.

В-четвертых, для сохранения условий комфортного приема $BER \leq 10^{-4}$ допустимое значение временного сдвига τ должно быть не более 2,7 мкс (модуляция QAM-4) и не более 1 мкс (QAM-16), вместе с тем величина RMER не должна быть менее 6 дБ (см. рисунок 6a) при модуляции QAM-4 и 15,3 дБ (см. рисунок 6b) при модуляции QAM-16 для всех уровней защиты. Это пороговые значения RMER, необходимые для комфортного приема при отсутствии в радиоканале АБГШ, помех и многолучевости.

На втором этапе оценивалось влияние АБГШ в радиоканале на вероятность появления BER при наличии временного сдвига τ окна прямого ДПФ относительно начала полезной части OFDM-символа (рисунки 8 и 9).

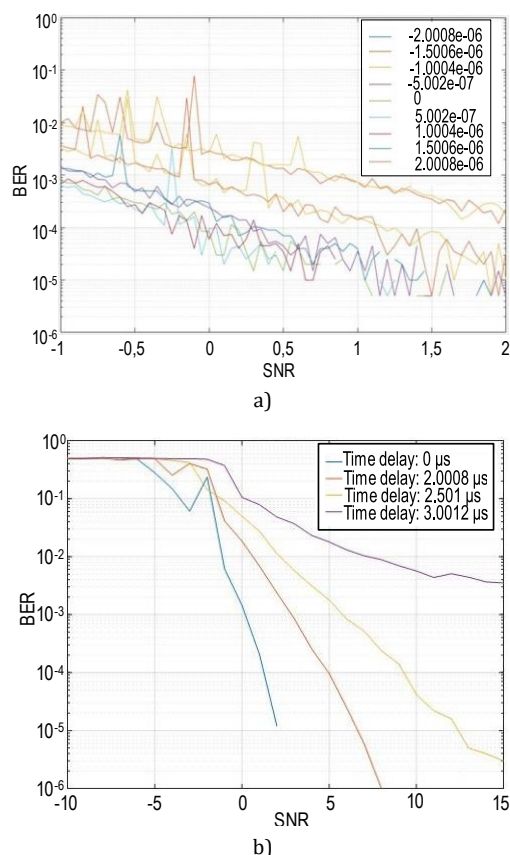


Рис. 8. Влияние SNR, дБ, при наличии в радиоканале АБГШ на вероятность появления BER для разных значений временного сдвига τ , мкс (модуляция поднесущих частот QAM-4): а) уровень защиты PL0; б) уровень защиты PL1, [19]

Fig. 8. The Effect of SNR, dB, in the Presence of AWGN in The Radio Channel, on the Probability of BER Occurrence for Different Values of the Time Shift τ , μ s (QAM-4 Subcarrier Frequency Modulation): а) Protection Level PL0; б) Protection Level PL1, [19]

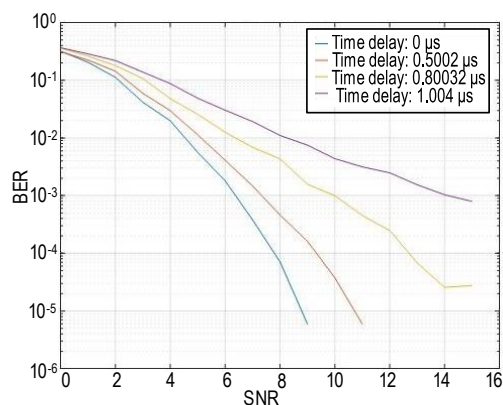


Рис. 9. Влияние SNR, дБ, при наличии в радиоканале АБГШ на вероятность появления BER для разных значений временного сдвига τ , мкс, окна прямого ДПФ относительно начала полезной части OFDM-символа для модуляции QAM-16, уровень защиты PL2 [19]

Fig. 9. Effect of SNR, dB, in the Presence of AWGN in the Radio Channel, on the BER for Different Values of the Time Shift τ , μ s, of the Direct DFT Window Relative to the Beginning of the Useful Part of the OFDM Symbol for QAM-16 Modulation, Protection Level PL2 [19]

Видно, что с увеличением временного сдвига τ при одном и том же значении SNR вероятность появления BER возрастает (см. рисунок 8а). Однако, несмотря на наличие значительных флуктуаций, аппроксимирующие зависимости для каждого значения τ имеют практически вид отрезков прямых линий с одинаковым наклоном. Зависимости, представленные на рисунках 8б и 9, подтверждают, что наличие временного рассогласования τ приводит к росту вероятности появления BER и к уменьшению RMER (см. рисунок 4) за счет вызванного этим явлением дополнительного ухудшения SNR при приеме.

Основное влияние на требуемую точность временной синхронизации DRM-приемника оказывает порядок модуляции QAM и существенно меньшее влияние – уровень защиты PL в области, где вероятность появления BER не превышает значение 10^{-4} . При $BER \leq 10^{-4}$ влиянием выбранного уровня защиты PL можно пренебречь (см. рисунок 7).

Покажем, что форма экспериментальных кривых, представленных на рисунке 9, подтверждается теорией. Величина RMER по определению равна:

$$RMER = 10 \lg \left(\frac{P_c}{P_{ош}} \right),$$

где P_c – средняя мощность полезного сигнала за интервал измерения T ; $P_{ош}$ – средняя мощность сигнала ошибки также за интервал T при наличии в радиоканале АБГШ и временного сдвига τ между моментами старта прямого ДПФ и началом полезной части OFDM-символа:

$$P_{ош} = \frac{P_c}{10^{RMER/10}},$$

где:

$$RMER_1 = 10 \lg \left(\frac{P_c}{P_{ош1}} \right), P_{ош1} = \frac{P_c}{10^{RMER_1/10}},$$

и

$$RMER_2 = 10 \lg \left(\frac{P_c}{P_{ош2}} \right), P_{ош2} = \frac{P_c}{10^{RMER_2/10}},$$

где $RMER_1$ и $P_{ош1}$ – RMER и средняя мощность сигнала ошибки при наличии в радиоканале АБГШ и отсутствии временного сдвига между моментами старта прямого ДПФ и началом полезной части OFDM-сигнала ($\tau = 0$); $RMER_2$ и $P_{ош2}$ – RMER и средняя мощность сигнала ошибки при отсутствии в радиоканале АБГШ и наличии временного сдвига ($\tau \neq 0$) между моментами старта прямого ДПФ и началом полезной части OFDM-символа.

Источники шумов $P_{ош1}$ и $P_{ош2}$ имеют разную природу появления, по этой причине их можно считать статистически независимыми случайными величинами, если это так, то справедливо выражение:

$$P_{ош} = P_{ош1} + P_{ош2},$$

и RMER при наличии АБГШ в радиоканале при $\tau \neq 0$ может быть найден по выражению:

$$RMER = 10 \lg \left(\frac{1}{\frac{1}{10^{\frac{RMER_1}{10}}} + \frac{1}{10^{\frac{RMER_2}{10}}}} \right). \quad (2)$$

Результаты измерений, приведенные на рисунках 4 (синяя линия) и 5, позволяют оценить влияние АБГШ в радиоканале на величину RMER при разных значениях временного сдвига $\tau \neq 0$. Полученные расчетным путем данные представлены на рисунке 9 для модуляции QAM-4 и уровне защиты PL1. Параметром каждой кривой здесь является значение временного сдвига τ (верхняя кривая для $\tau = 0$ мкс; вторая кривая сверху – $\tau = 1$ мкс; третья кривая сверху – $\tau = 2$ мкс; четвертая кривая сверху – $\tau = 2,5$ мкс; пятая кривая сверху – $\tau = 3$ мкс).

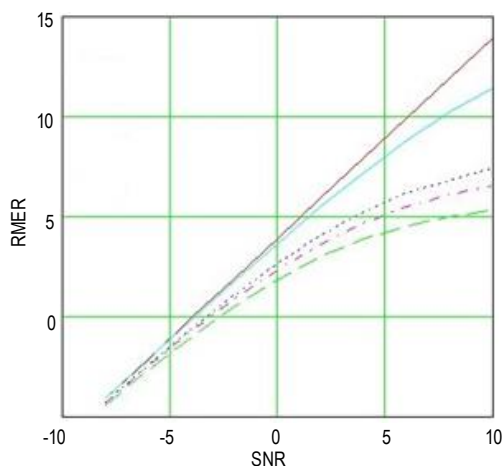


Рис. 10. Зависимости RMER, дБ, от SNR, дБ, при наличии в радиоканале АБГШ для разных значений временного сдвига τ между моментами старта прямого ДПФ и началом полезной части OFDM-символа

Fig. 10. Dependences of RMER, dB, on SNR, dB, in the Presence of AWGN in the Radio Channel for Different Values of the Time Shift τ between the Moments of the Start of the Direct DFT and the Beginning of the Useful Part of the OFDM-Symbol

Заметим, что при идеальном временном согласовании ($\tau = 0$) и отсутствии АБГШ в радиоканале RMER равен бесконечности (см. рисунок 4), если не учитывать шумы оборудования. При наличии временного сдвига ($\tau \neq 0$) или (и) АБГШ в радиоканале средняя мощность сигнала ошибки уже не будет равна нулю. Увеличение τ сопровождается ростом сигнала ошибки и, соответственно, уменьшением величины RMER (см. рисунок 5). Этот вывод подтверждается данными измерений (см. рисунок 4) и расчетов (см. рисунок 10).

Предположим, что в радиоканале присутствует АБГШ и отношение сигнал / шум равно $SNR_{\min}$ при выбранных значениях порядка модуляции QAM и скорости кода, а уменьшение этого отношения до величины порогового значения (SNR_{Π}), при кото-

ром прием сигнала становится невозможным, происходит за счет появляющегося временного сдвига τ , тогда выражение (2) преобразуется к следующему виду:

$$SNR_{\Pi} = 10 \lg \left(\frac{1}{\frac{1}{10^{\frac{(SNR_{\min})}{10}}} + \frac{1}{10^{\frac{RMER_2}{10}}}} \right) \quad (3)$$

и $SNR_{\Pi} = SNR_{\min} - 1$.

Выражение (3) позволяет рассчитать значение $RMER_2$ и затем, используя зависимость на рисунке 5, найти соответствующее ему значение максимально допустимого временного сдвига $\tau_{\text{доп}}$.

В системе DRM уменьшение SNR в радиоканале относительно его минимального значения ($SNR_{\min}$) на 1 дБ приводит к тому, что прием DRM-сигнала оказывается невозможным [16–18]. Ранее для системы DRM при работе в режиме устойчивости E в условиях стационарного приема Fx были измерены минимальные значения $SNR_{\min}$ (таблица 2, второй столбец), при которых еще сохраняется комфортный прием, т. е. вероятность появления BER, как того требует стандарт, не превышает 10^{-4} [8, 9].

ТАБЛИЦА 2. Максимально допустимое значение временного сдвига ($\tau_{\text{доп}}$) между моментами старта прямого ДПФ и началом полезной части OFDM-символов для разных значений порядка модуляции QAM и уровня помехозащитности PL требуемое для условий комфортного приема при наличии в радиоканале АБГШ

TABLE 2. Maximum Permitted Value of the Time Shift ($\tau_{\text{доп}}$) between the Start of the Direct DFT and the Beginning of the Useful Part of OFDM-Symbols for Different Values of the Modulation Order QAM and Protection Level PL, Required for Conditions of Comfortable Reception in the Presence of AWGN in the Radio Channel

Модуляция / Скорость кода / Уровень защиты	$SNR_{\min}$, дБ	SNR_{Π} , дБ	$RMER_2$, дБ	$\tau_{\text{доп}}$, мкс
QAM-4/0,250/PL0	2,27	1,27	8,2	2,3
QAM-4/0,333/PL1	2,71	1,71	8,6	2,2
QAM-4/0,4/PL2	2,93	1,93	8,8	2,1
QAM-4/0,5/PL3	4,09	3,09	10	1,8
QAM-16/0,330/PL0	7,64	6,64	13,5	1,3
QAM-16/0,411/PL1	9,22	8,22	15,1	1,0
QAM-16/0,500/PL2	10,24	9,24	16,1	0,9
QAM-16/0,411/PL1	11,40	10,40	17,3	0,8

Значения $SNR_{\min}$ и зависимости RMER (см. рисунки 4 (синяя кривая) и 5), полученные экспериментальным путем, позволяют найти максимально допустимое значение временного сдвига $\tau_{\text{доп}}$, при котором вероятность появления BER не превысит требуемое стандартом значение 10^{-4} . Результаты этих вычислений приведены в таблице 2 и на рисунке 11, где по оси ординат отложены значения $\tau_{\text{доп}}$, мкс, а по оси абсцисс – требуемое стандартом значение $SNR_{\min}$ для системы DRM.

И наконец, используя пилотные поднесущие частоты и метод билинейной интерполяции, были скомпенсированы вызванные временным сдвигом изменения амплитуды и фазы для информацион-

ных поднесущих частот при приеме. Зависимости RMER от временного сдвига для данного случая при отсутствии в радиоканале АБГШ, помех и многолучевости представлены на рисунке 12.

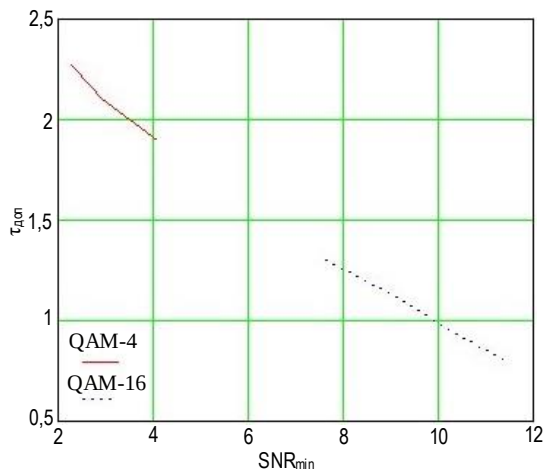


Рис. 11. Максимально допустимое значение временного сдвига между моментами старта прямого ДПФ и началом полезной части OFDM-символов при наличии в радиоканале АБГШ: левая верхняя линия получена для модуляции QAM-4; правая нижняя линия – для QAM-16

Fig. 11. Maximum Permitted Value of Time Shift between the Moments of the Start of the Direct DFT and the Beginning of the Useful Part of OFDM-Symbols in the Presence of AWGN in the Radio Channel: the Left Upper Line is Obtained for QAM-4 Modulation, the Right Lower Line is Obtained for QAM-16 Modulation

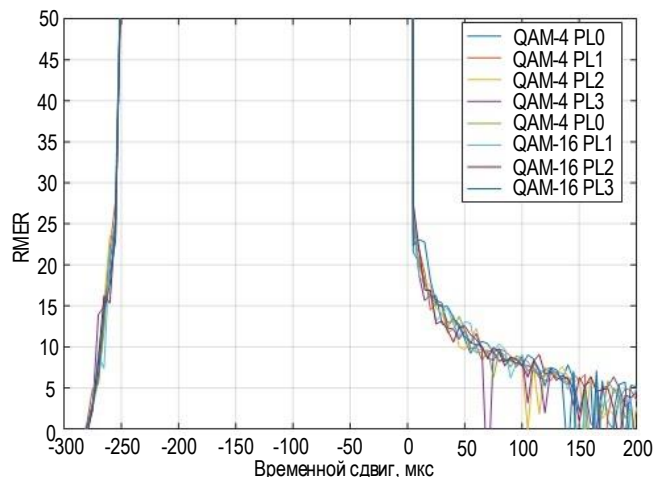


Рис. 12. Зависимости RMER, дБ, от величины временного сдвига при компенсации искажений для информационных поднесущих частот при разных видах модуляции QAM и уровнях защиты PL при отсутствии в радиоканале шума, помех и многолучевости

Fig. 12. Dependences of RMER, dB, on the Magnitude of the Time Shift during Distortion Compensation for Information Subcarriers with Different Types of QAM Modulation and PL Protection Levels in the Absence of Noise, Interference and Multipath in the Radio Channel

Компенсация искажений (см. рисунок 12) по амплитуде и фазе для информационных поднесущих частот позволяет сохранить высокое значение RMER при значительно большем значении τ . Видно, что когда старт прямого ДПФ происходит раньше начала полезной (информационной) части OFDM-символа, компенсация оказывается более эффективной. В худшем случае, когда момент старта прямого ДПФ берется позже начала полезной части OFDM-символа при модуляции QAM-4 и уровне защиты PL0, допустимое значение временного сдвига должно быть не более 190 мкс, при модуляции QAM-16 – не более 25 мкс. Эти результаты также подтверждают необходимость точной временной синхронизации при приеме OFDM-сигнала.

Заключение

Во-первых, сохранение условий комфортного приема сигнала стандарта DRM невозможно без точной временной синхронизации приемного оборудования. При модуляции поднесущих частот QAM-4 временной сдвиг τ между моментами старта прямого ДПФ и началом полезной части OFDM-символа не должен превышать 1,8–2,3 мкс. При мо-

дуляции QAM-16 – эта величина должна быть не более 0,8–1,3 мкс. Разброс полученных значений определяется выбранным уровнем защиты PL.

При компенсации искажений информационных поднесущих частот эти требования становятся менее жесткими: не более 190 мкс при модуляции QAM-4 и не более 25 мкс при модуляции QAM-16 при условии, что значения SNR соответствуют требованиям стандарта на систему DRM.

Во-вторых, основное влияние на требуемую точность временной синхронизации DRM-приемника при работе в режиме устойчивости E оказывает порядок модуляции QAM и существенно меньшее влияние – уровень защиты PL в области, где вероятность появления BER не превышает 10^{-4} . В случае, если $BER \leq 10^{-4}$, влиянием выбранного уровня защиты PL можно пренебречь.

В-третьих, появление τ приводит к уменьшению радиуса зоны покрытия, тем значительнее, чем больше эта величина.

В-четвертых, результаты вычислений подтверждены данными измерений и должны быть учтены при разработке блока временной синхронизации DRM-приемника.

Список источников

1. Бакулин М.Г., Крейнделин В.Б., Шлома А.М., Шумов А.П. Технология OFDM: учебное пособие для вузов. М.: Горячая линия – Телеком, 2016. 352 с.
2. Гуминский О.А., Мышьянов С.В. Разработка алгоритма синхронизации сигнала для радиоприемного устройства стандарта DRM mode // XI Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (АПИНО 2022, Санкт-Петербург, Российская Федерация, 15–16 февраля 2022 г.). СПб.: СПбГУТ, 2022. С. 178–183. EDN:USFELI
3. Keller T., Hanzo L. Orthogonal frequency division multiplex synchronization techniques for wireless local area networks // Proceedings of the 7th International Symposium on Personal, Indoor, and Mobile Communications (PIMRC '96, Taipei, Taiwan, 18 October 1996). IEEE, 1996. PP. 963–967. DOI:10.1109/PIMRC.1996.568424
4. Sandell M., Van De Beek J.J., Borjesson P.O. Timing and frequency synchronization in OFDM systems using the cyclic prefix // Proceedings of the International Symposium on Synchronization (Essen, Germany, 14–15 December 1995). 1995. PP. 16–19.
5. Chevillat P.R., Maiwald D., Ungerboeck G. Rapid Training of a Voiceband Data-Modem Receiver Employing an Equalizer with Fractional-T Spaced Coefficients // IEEE Transactions on Communications. 1987. Vol. 35. Iss. 9. PP. 869–876. DOI:10.1109/TCOM.1987.1096887
6. Van de Beek J.J., Sandell M., Isaksson M., Borjesson P.O. Low-complex frame synchronization in OFDM systems // Proceedings of the 4th International Conference on Universal Personal Communications (ICUPS 95, Tokyo, Japan, 06–10 November 1995). IEEE, 1995. DOI:10.1109/ICUPC.1995.497156
7. Moose P.H. A Technique for orthogonal frequency division multiplexing frequency offset correction // IEEE Transactions on Communications. 1994. Vol. 42. Iss. 10. PP. 2908–2914. IEEE, 1994. DOI:10.1109/26.328961
8. ETSI ES 201 980 V2.2.1 (2005-10) V4.2.1 (2021-01). Digital Radio Mondiale (DRM); System Specification.
9. Рекомендация МСЭ-R BS.1660-8 (06/2019). Техническая основа для планирования наземного цифрового звукового радиовещания в полосе ОВЧ. С. 33–49.
10. Гуминский О.А. Модель приемо-передающего тракта системы цифрового радиовещания DRM+. URL: <https://cloud.mail.ru/public/AxE9/9nJmYeYHv> (дата обращения 24.03.2025)
11. Дубов М.А., Полянин Ю.В., Стоянов Д.Д., Брюханов Ю.А. Оценка вероятности битовой ошибки приема сигналов с квадратурной модуляцией неэталонными методами // DSPA: Вопросы применения цифровой обработки сигналов. 2012. Т. 2. № 2. С. 173–177. EDN:ZHJHRT
12. ETSI EN 302 245-1 V1.1.1 (2005-01). Electromagnetic compatibility and Radio spectrum Matters (ERM); Transmitting equipment for the Digital Radio Mondiale (DRM) broadcasting service; Part 1: Technical characteristics and test methods.
13. ETSI EN 302 245 V2.2.1 (2022-05). Transmitting equipment for the Digital Radio Mondiale (DRM) service; Harmonised Standard for access to radio spectrum.
14. Gu Q. RF System Design of Transceivers for Wireless Communications. Springer, 2005. 467 p.
15. Дубов М.А., Приоров А.Л. Методика неэталонной оценки отношения сигнал/шум и вероятности битовой ошибки для сигналов с квадратурной модуляцией // Цифровая обработка сигналов. 2012. № 4. С. 37–43. EDN:PXXBYD
16. Ковалгин Ю.А. Цифровое радиовещание: системы и технологии. М.: Горячая линия – Телеком, 2021. 580 с.
17. Соколов С.А., Мышьянов С.В., Ковалгин Ю.А. Исследование аналого-цифровой DRM/ ЧМ-зоны радиовещания в полосе частот 87,5–108 МГц // Электросвязь. 2021. № 4. С. 23–29. DOI:10.34832/ELSV.2021.17.4.003. EDN:OBGGGA
18. Ковалгин Ю.А., Соколов С.А., Мышьянов С.В. и др. Построение опытной аналого-цифровой зоны радиовещания DRM+/УКВ ЧМ в диапазоне 87,5–108 МГц: отчет о НИР. Дигитон Системс, 2019. 223 с.
19. Гуминский О.А., Ковалгин Ю.А. Исследование и оценка требуемой точности временной синхронизации приемного оборудования в системе цифрового радиовещания DRM mode E // XIII Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (АПИНО 2022, Санкт-Петербург, Российская Федерация, 27–28 февраля 2024 г.). СПб.: СПбГУТ, 2024. С. 201–205. EDN:WDRCWV

References

1. Bakulin M.G., Kreindelin V.B., Shloma A.M., Shumov A.P. *OFDM Technology*. Moscow: Goriachaia liniia Telekom Publ.; 2016. 352 p. (in Russ.)
2. Guminskiy O., Myshyanov S. Development of a Signal Synchronization Algorithm for a DRM Mode E Radio Receiver. *Proceedings of the XIth International Conference on Infotelecommunications in Science and Education, 15–16 February 2022, St. Petersburg, Russian Federation*. St. Petersburg: Saint-Petersburg State University of Telecommunications Publ.; 2022. p.178–183. (in Russ.) EDN:USFELI
3. Keller T., Hanzo L. Orthogonal frequency division multiplex synchronization techniques for wireless local area networks. *Proceedings of the 7th International Symposium on Personal, Indoor, and Mobile Communications, PIMRC '96, 18 October 1996, Taipei, Taiwan*. IEEE; 1996. p.963–967. DOI:10.1109/PIMRC.1996.568424
4. Sandell M., Van De Beek J.J., Borjesson P.O. Timing and frequency synchronization in OFDM systems using the cyclic prefix. *Proceedings of the International Symposium on Synchronization, 14–15 December 1995, Essen, Germany*. 1995. p.16–19.
5. Chevillat P.R., Maiwald D., Ungerboeck G. Rapid Training of a Voiceband Data-Modem Receiver Employing an Equalizer with Fractional-T Spaced Coefficients. *IEEE Transactions on Communications*. 1987;35(9):869–876. DOI:10.1109/TCOM.1987.1096887
6. Van de Beek J.J., Sandell M., Isaksson M., Borjesson P.O. Low-complex frame synchronization in OFDM systems. *Proceedings of the 4th International Conference on Universal Personal Communications, ICUPS 95, 06–10 November 1995, Tokyo, Japan*. IEEE; 1995. DOI:10.1109/ICUPC.1995.497156

7. Moose P.H. A Technique for Orthogonal Frequency Division Multiplexing Frequency Offset Correction. *IEEE Transactions on Communications*. 1994;42(10):2908–2914. IEEE, 1994. DOI:10.1109/26.328961
8. ETSI ES 201 980 V2.2.1 (2005-10) V4.2.1 (2021-01). *Digital Radio Mondiale (DRM); System Specification*. 184 p.
9. Rec. ITU-R BS.1660-8 (06/2019). *Technical basis for planning of terrestrial digital sound broadcasting in the VHF band*. June 2019. p.33–49.
10. Guminsky O.A. *The model of the receiving and transmitting path of the DRM+ digital radio broadcasting system*. (in Russ.) URL: <https://cloud.mail.ru/public/AxE9/9nJmYeYHv> [Accessed 24.03.2025]
11. Dubov M.A., Polyaniin Yu.V., Stoyanov D.D., Bryukhanov Yu.A. Estimation of the probability of bit error in the reception of signals with quadrature modulation by non-reference methods. *Digital signal processing and its applications*. 2012;2(20): 173–177. (in Russ.) EDN:ZHJHRT
12. ETSI EN 302 245-1 V1.1.1. *Electromagnetic compatibility and Radio spectrum Matters (ERM); Transmitting equipment for the Digital Radio Mondiale (DRM) broadcasting service; Part 1: Technical characteristics and test methods*. January 2005.
13. ETSI EN 302 245 V2.2.1. *Transmitting equipment for the Digital Radio Mondiale (DRM) service; Harmonised Standard for access to radio spectrum*. May 2022.
14. Gu Q. *RF System Design of Transceivers for Wireless Communications*. Springer, 2005. 467 p.
15. Dubov M.A., Priorov A.L. Methodology for non-reference estimation of signal-to-noise ratio and bit error probability for signals with quadrature modulation. *Digital Signal Processing*. 2012;4:37–43. (in Russ.) EDN:PXXBYD
16. Kovalgin Yu.A. *Digital radio broadcasting: systems and technologies*. Moscow: Goriachaia liniia Telekom Publ.; 2021. 580 p. (in Russ.)
17. Sokolov S.A., Myshyanov S.V., Kovalgin Yu.A. Results of the DRM simulcast field trials in 87.5–108 MHz band. *Elektrosviaz*. 2021;4:23–29. (in Russ.) DOI:10.34832/ELSV.2021.17.4.003. EDN:OBGGGA
18. Kovalgin Yu.A., Sokolov S.A., Myshyanov S.V., et al. *Construction of an Experimental Analog-Digital DRM+/VHF FM Radio Broadcasting Zone in the 87.5–108 MHz Range: a Research Report*. Digiton Systems Publ.; 2019. 223 p. (in Russ.)
19. Guminsky O., Kovalgin Yu. Investigation and Evaluation of the Required Accuracy of Time Synchronization of Receiving Equipment in the E-Mode of the DRM Digital Radio Broadcasting System. *Proceedings of the XIIIth International Conference on Infotelecommunications in Science and Education, 27–28 February 2024, St. Petersburg, Russian Federation*. St. Petersburg: Saint Petersburg State University of Telecommunications Publ.; 2024. p.201–205. (in Russ.) EDN:WDRCWV

Статья поступила в редакцию 31.01.2025; одобрена после рецензирования 21.02.2025; принята к публикации 14.03.2025.

The article was submitted 31.01.2025; approved after reviewing 21.02.2025; accepted for publication 14.03.2025.

Информация об авторах:

ГУМИНСКИЙ
Олег Андреевич

аспирант кафедры телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0007-3560-3898>

КОВАЛГИН
Юрий Алексеевич

доктор технических наук, профессор, профессор кафедры телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-6753-8062>

Ковалгин Ю.А. является членом редакционного совета журнала «Труды учебных заведений связи» с 2016 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли. Kovalgin Yu.A. has been a member of the journal "Proceedings of Telecommunication Universities" Editorial Council since 2016, but has nothing to do with the decision to publish this article. The article has passed the review procedure accepted in the journal. The authors have not declared any other conflicts of interest.

Научная статья

УДК 621.391

<https://doi.org/10.31854/1813-324X-2025-11-2-32-40>

EDN:CTTRGR



Оптимизация распределения пользователей, мощности и расположения воздушной базовой станции на основе методов глубокого обучения с подкреплением

Тунг Зыонг Чан ✉, chan.tz@sut.ruАндрей Евгеньевич Кучерявый, akouch@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. С развитием информационных технологий и Интернета вещей возрастает спрос на более эффективные и гибкие мобильные сети. Будущие беспроводные системы должны обеспечивать не только высокую скорость и надежность соединения, но и быстрое восстановление связи в аварийных ситуациях. Наземные базовые станции (GBS, аббр. от англ. Ground Base Stations) обычно устанавливаются стационарно и ориентированы на длительное обслуживание, что ограничивает их эффективность при резком увеличении трафика или повреждении инфраструктуры. В таких условиях воздушные базовые станции (ВБС) становятся перспективным решением. Благодаря своей мобильности, доступной стоимости и возможности быстрого развертывания, они могут поддерживать работу наземных станций в условиях высокой плотности пользователей или в случае чрезвычайных ситуаций, когда GBS повреждены или уничтожены. Это делает их важным элементом сетей связи будущего.

Постановка задачи: разработка методов размещения ВБС в трехмерном пространстве и распределения пользователей и мощности среди последних с целью максимизации скорости передачи данных систем.

Цель работы: повышение скорости передачи данных систем с использованием ВБС для поддержки GBS за счет оптимального трехмерного положения ВБС, распределения пользователей между ВБС и GBS, а также распределения мощности среди пользователей.

Используемые методы. Исследования проводились с применением динамического подхода, при котором радиус покрытия GBS постепенно сокращается, а также алгоритма глубокого обучения с подкреплением. Анализ полученных **результатов** показал высокую эффективность предложенного метода и позволил добиться значительного увеличения скорости передачи данных в рамках поставленной задачи.

Научная новизна предложенного решения заключается в том, что совместная оптимизация размещения ВБС и распределения мощности в условиях ограниченных ресурсов позволила выявить зависимость между радиусом покрытия GBS и высотой полета ВБС: при большем радиусе покрытия GBS высота полета ВБС будет ниже, и наоборот. **Практическая значимость** заключается в возможности разработки методики планирования сетей связи общего пользования с использованием ВБС для поддержки GBS в условиях ограниченных ресурсов. Это позволяет обеспечить высокую суммарную скорость передачи данных и повысить надежность функционирования сети.

Ключевые слова: воздушная базовая станция, размещение, распределение мощности, распределения пользователей, GBS, глубокое обучение с подкреплением

Ссылка для цитирования: Чан Т.З., Кучерявый А.Е. Оптимизация распределения пользователей, мощности и расположения воздушной базовой станции на основе методов глубокого обучения с подкреплением // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 32–40. DOI:10.31854/1813-324X-2025-11-2-32-40. EDN:CTTRGR

Original research¹

<https://doi.org/10.31854/1813-324X-2025-11-2-32-40>

EDN:CTTRGR

User Association, Power Allocation, and Aerial Base Stations Placement in Hybrid Networks: A Deep Reinforcement Learning Approach

 **Tung D. Tran** ✉, chan.tz@sut.ru

 **Andrey E. Koucheryavy**, akouch@sut.ru

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance. With the development of information technologies and the Internet of Things, the demand for more efficient and flexible mobile networks is increasing. Future wireless systems must not only ensure high speed and reliability of connections but also enable quick recovery of communication in emergency situations. Ground base stations (GBS) are typically installed stationary and are designed for long-term service, which limits their efficiency during sudden increases in traffic or infrastructure damage. In such conditions, aerial base station (ABS) emerge as a promising solution. Due to their mobility, affordability, and the ability to deploy quickly, they can support the operation of ground stations in high user density conditions or in emergencies when GBS are damaged or destroyed. This makes them an essential element of future communication networks.

Problem Statement. Development of methods for the placement of ABS in three-dimensional space and the distribution of users and power among users with the goal of maximizing the data transmission speed of the systems.

Goal of the work. Increase the data transmission speed of systems using ABS to support GBS through optimal three-dimensional positioning of ABS, distribution of users between ABS and GBS, and power allocation among users.

Methods. The research was conducted using a dynamic approach, in which the coverage radius of the GBS is gradually reduced, along with the reinforcement learning algorithm. The analysis of the **results** showed the high effectiveness of the proposed method and allowed for a significant increase in data transmission speed within the framework of the task.

Scientific novelty. The scientific novelty of the proposed solution lies in the joint optimization of the placement of ABS and power allocation under resource constraints, which revealed a dependency between the coverage radius of GBS and the flight altitude of ABS. Specifically, as the coverage radius of GBS increases, the optimal flight altitude of ABS decreases, and vice versa.

Practical significance. The practical significance lies in the possibility of developing a methodology for planning public communication networks using ABS to support GBS under resource constraints. This approach makes it possible to ensure a high total data transmission rate and improve the reliability of network operation.

Keywords: aerial base station, placement, power allocation, user association, GBS, deep reinforcement learning

For citation: Tran T.D., Koucheryavy A.E. User Association, Power Allocation, and Aerial Base Stations Placement in Hybrid Networks: A Deep Reinforcement Learning Approach. *Proceedings of Telecommunication Universities*. 2025;11(2):32–40. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-32-40. EDN:CTTRGR

1. Введение

С быстрым развитием информационных технологий и Интернета вещей [1], а также с ростом числа устройств, чрезвычайный спрос на беспроводные услуги создает огромную проблему для мобильных сетей. Кроме того, будущие беспроводные сети должны быть гибкими и обладать возможностью оперативного восстановления связи. Поэтому

ожидается, что воздушные средства благодаря своей мобильности, гибкости развертывания и перемещения, станут незаменимой частью будущих систем беспроводной связи [2]. По сравнению с традиционными наземными сетями, где расположение наземных базовых станций (GBS, аббр. от англ. Ground Base Stations) жестко спроектировано в соответствии с долгосрочными потребностями тра-

фика на определенной территории, воздушные базовые станции (ВБС) могут улучшить качество обслуживания за счет гибкого изменения своего положения, что обеспечивает прямую видимость (LoS, аббр. от англ. Line of Sight) [3] соединения между ВБС и наземными пользователями в большинстве сценариев. В случае стихийных бедствий GBS могут быть уничтожены, или может резко возрасти трафик, например, во время спортивных мероприятий или фестивалей, что приводит к перегрузке. В таких случаях потенциальным решением являются ВБС, которые можно быстро развернуть для предоставления беспроводной связи наземным пользователям.

Эффективное размещение базовых станций является важной задачей как для традиционных GBS [4], так и для сетей ВБС. В то время как размещение GBS обычно ограничено природными условиями местности, ВБС могут быть развернуты более гибко, поскольку их местоположение не зависит от географических ограничений. Однако они сталкиваются с серьезными проблемами, в первую очередь с ограничением мощности и нестабильностью канала воздушно-наземной передачи из-за их непрерывного движения. Поэтому важно оптимизировать как размещение ВБС, так и распределение мощности.

В данной работе рассматривается задача распределения пользователей между GBS и ВБС, а также оптимизация распределения мощности для пользователей и размещения ВБС с целью максимизации скорости передачи данных в системе. Для решения данной проблемы предлагается динамический подход, при котором радиус покрытия GBS постепенно уменьшается. На основе радиуса покрытия GBS можно определить положение и количество пользователей, подключенных к GBS и ВБС. Для тех, кто подключен к GBS, мощность между ними распределяется одинаково. Для оставшихся пользователей применяется метод глубокого обучения с подкреплением – алгоритм глубокого детерминированного градиента политики (DDPG, аббр. от англ. Deep Deterministic Policy Gradient), который используется для оптимизации размещения ВБС и распределения мощности.

2. Модель системы и постановка задачи

Предполагается, что GBS находится в центре рассматриваемой зоны. Расположение GBS задается координатами (x_g, y_g, h_g) , где (x_g, y_g) – координаты GBS на поверхности земли, а h_g – высота GBS. Для поддержки наземной сети используется одна ВБС. Ее положение задано координатами (x_a, y_a, h_a) , где (x_a, y_a) – координаты ВБС на поверхности земли, а h_a – высота полета. В зависимости от мощности, конструкции, а также местных законов высота по-

лета должна находиться в определенном диапазоне. Всего в рассматриваемой зоне находится N пользователей. Расположение каждого пользователя n , где $n \in \{1, \dots, N\}$, задается координатами $(x_n, y_n, 0)$.

Для связи «воздух–земля» рассматривается вероятностная модель канала LoS [5], в которой вероятность LoS зависит от взаимного расположения между ВБС и пользователем, угла места между ними и окружающей среды.

Вероятность наличия LoS P^{LoS} между пользователем n рассчитывается по выражению:

$$P^{LoS} = \frac{1}{1 + a \exp(-b[\theta - a])}, \quad (1)$$

где a и b – постоянные значения, зависящие от окружающей среды (сельская, городская, густонаселенная и т. д.); θ – угол места между пользователем n и ВБС ($^\circ$), равен $\frac{180}{\pi} \times \arcsin(\frac{h_a}{d_{a,n}})$; $d_{a,n}$ – расстояние между ВБС и пользователем n , равно $\sqrt{(x_a - x_n)^2 + (y_a - y_n)^2 + h_a^2}$.

С другой стороны, вероятность соединения в условиях отсутствия прямой видимости (NLoS, аббр. от англ. Non-Line of Sight) будет равна:

$$P^{NLoS} = 1 - P^{LoS}.$$

Как следует из (1), вероятность LoS растет по мере увеличения угла места между пользователем и ВБС. Полагаем, что потери распространения между ВБС и пользователем n составляют β_n . С учетом наличия препятствий в воздушном канале связи, β_n обычно моделируются как случайная величина, связанная с вероятностью возникновения LoS- и NLoS-соединений.

Следовательно, β_n можно выразить как:

$$\beta_n = \begin{cases} \beta_0 d_{a,n}^{-\hat{\alpha}} & \text{LoS – соединение} \\ \kappa \beta_0 d_{a,n}^{-\hat{\alpha}} & \text{NLoS – соединение} \end{cases}, \quad (2)$$

где β_0 – это потеря сигнала на эталонном расстоянии 1 м для LoS-соединения; κ – обозначает дополнительные потери распространения для NLoS-соединения; $\hat{\alpha}$ – показатель затухания сигнала; и $d_{a,n}$ – расстояние между ВБС и пользователем n .

Таким образом, средний коэффициент усиления мощности канала $h_{a,n}^{BBS}$ между ВБС и пользователем n вычисляется по выражению [6]:

$$h_{a,n}^{BBS} = P^{LoS} \beta_0 d_{a,n}^{-\hat{\alpha}} + P^{NLoS} \kappa \beta_0 d_{a,n}^{-\hat{\alpha}}. \quad (3)$$

В данном исследовании для моделирования канала GBS используется NLoS-канал [7]. Это означает, что сигнал между GBS и пользователем может быть прегражден различными объектами, такими как здания, деревья или другие препятствия, блокирующие LoS между ними. В сценарии городского

макрорайона (UMa, *аббр. от англ.* Urban Macro), где GBS обычно размещается в условиях с множеством физических препятствий (например, высокие здания, плотная растительность), данный тип канала является наиболее типичным.

Потери распространения $PL_{g,n}^{GBS}$ (дБ) можно определить следующим образом:

$$PL_{g,n}^{GBS} = 32,4 + 20\log_{10}(f_g) + 30\log_{10}(d_{g,n}), \quad (4)$$

где f_g – частота работы GBS (ГГц), $d_{g,n}$ – расстояние от GBS до пользователя n (м).

3. Формулировка проблемы

Каждый пользователь должен быть назначен либо к ВБС, либо к GBS для осуществления нисходящей передачи данных. Чтобы определить состояние подключения пользователя, вычислим отношение сигнал / шум и помехи (SNR, *аббр. от англ.* Signal-to-Noise Ratio) в нисходящем канале. Пользователь считается подключенным к GBS или ВБС, если полученное значение SNR превышает заданный порог SNR_0 .

Предположим, что пользователь n обслуживается GBS, тогда полученное значение SNR_n определяется следующим образом:

$$SNR_{g,n}^{GBS} = \frac{P_{GBS} h_{g,n}^{GBS} g_n}{\sigma^2}, \quad (5)$$

где P_{GBS} – мощность передачи GBS; $h_{g,n}^{GBS}$ – коэффициент усиления канала между GBS и пользователем n , равный $\frac{1}{10^{(PL_{g,n}^{GBS}/10)}}$; g_n – коэффициент распределения мощности GBS для пользователя n ; σ^2 – мощность шума.

Аналогично, пользователь m обслуживается ВБС, тогда полученное значение SNR_m для него будет следующим:

$$SNR_{a,m}^{BBS} = \frac{P_{BBS} h_{a,m}^{BBS} p_m}{\sigma^2}, \quad (6)$$

где P_{BBS} – мощность передачи ВБС; $h_{a,m}^{BBS}$ – коэффициент усиления канала между ВБС и пользователем m , рассчитывается по формуле (3); p_m – коэффициент распределения мощности ВБС для пользователя m .

Предположим, что N – общее количество пользователей в рассматриваемой зоне, N_n – количество пользователей подключенных к GBS, а количество пользователей, подключенных к ВБС, равно N_m . Поскольку все пользователи одинаковы и между ними нет приоритета, будем предполагать, что полоса пропускания, выделяемая GBS и ВБС каждому пользователю при подключении, одинакова и равна B . Для упрощения математических вычислений полоса пропускания B нормализована до единицы (бит/с). Достижимая скорость передачи дан-

ных в нисходящем канале системы может быть рассчитана на основе формулы Шеннона [8]:

$$R = R_{GBS} + R_{BBS} = \sum_{n=1}^{N_n} \log_2(1 + SNR_{g,n}^{GBS}) + \sum_{m=1}^{N_m} \log_2(1 + SNR_{a,m}^{BBS}), \quad (7)$$

где $R_{GBS} = \sum_{n=1}^{N_n} \log_2(1 + SNR_{g,n}^{GBS})$ – общая достижимая скорость всех пользователей, подключенных к GBS; $R_{BBS} = \sum_{m=1}^{N_m} \log_2(1 + SNR_{a,m}^{BBS})$ – общая достижимая скорость всех пользователей, подключенных к ВБС.

Цель данной работы заключается в оптимальном распределении пользователей между ВБС и GBS для максимизации нисходящей скорости связи в системе. Задача включает в себя определение стратегии распределения пользователей, мощности и трехмерное размещение ВБС.

Она может быть сформулирована следующим образом:

$$\max_{\{x_a, y_a, h_a, p_m\}} R \quad (8)$$

при определенных условиях:

$$C1: \sum_{m=1}^{N_m} p_m = 1,$$

$$C2: \sum_{n=1}^{N_n} g_n = 1,$$

$$C3: N_m + N_n = N,$$

$$C4: x_{\min} \leq x_a \leq x_{\max},$$

$$C5: y_{\min} \leq y_a \leq y_{\max},$$

$$C6: h_{\min} \leq h_a \leq h_{\max}.$$

Ограничения C1 и C2 подразумевают, что общее распределение мощности пользователей при подключении к ВБС и GBS равно 1, соответственно. Ограничение C3 означает, что количество пользователей, к которым подключаются ВБС и GBS, должно быть равно количеству исходных пользователей в рассматриваемой зоне. Ограничения C4, C5 и C6 означают, что ВБС должны располагаться в рассматриваемой зоне, где $x_{\min}$, $y_{\min}$, $h_{\min}$ – нижние границы по осям, а $x_{\max}$, $y_{\max}$, $h_{\max}$ – верхние.

4. Алгоритм для решения задачи

4.1. Стратегия распределения пользователей

В беспроводных сетях распределение пользователей между GBS и ВБС играет ключевую роль в оптимизации общей скорости передачи данных. Одной из главных задач является определение границы зоны покрытия GBS таким образом, чтобы обслуживать оптимальное количество пользователей в зоне, одновременно снижая нагрузку на ВБС.

На практике зоны покрытия GBS и ВБС могут пересекаться, что создает дополнительную сложность при распределении пользователей. Чтобы не усложнять модель, в данном исследовании предполагается, что каждый пользователь может быть подключен только к одной из базовых станций – либо к GBS, либо к ВБС. Кроме того, чтобы минимизировать межканальные помехи, GBS и ВБС работают на отдельных частотных диапазонах. Для решения этой проблемы предлагается динамический подход, при котором радиус покрытия GBS изменяется. Пользователи, находящиеся за пределами зоны покрытия GBS, обслуживаются ВБС. На основе местоположения и количества пользователей, обслуживаемых GBS и ВБС, а также оптимального местоположения ВБС, достижимая скорость передачи данных в нисходящем канале системы может быть рассчитана с помощью формулы (7). Изменение радиуса покрытия GBS приводит к изменению количества пользователей, подключенных к GBS и ВБС, тем самым изменяя общую скорость передачи данных в нисходящем канале системы. Затем выбирается такое значение радиуса покрытия GBS, при котором общая скорость передачи данных максимальна.

4.2. Распределение мощности и трехмерное размещение ВБС

После определения местоположения и количества пользователей, подключенных к GBS или ВБС, на основе радиуса покрытия GBS, проводится анализ распределения мощности между пользователями и определяется оптимальное положение ВБС в трехмерном пространстве. Чтобы упростить задачу и сосредоточиться на оптимизации местоположения ВБС в трехмерном пространстве и распределении мощности среди пользователей, подключенных к ВБС, предполагается, что все пользователи, подключенные к GBS, получают равное распределение мощности. Для решения этой задачи применяется алгоритм DDPG. ВБС выступает в роли агента, и модель DDPG строится на основе элементов MDP (марковского процесса принятия решений): состояния, действий, вознаграждений и вероятности перехода. Задача ВБС – научиться оптимизировать свою позицию и распределение мощности, чтобы максимизировать общую скорость передачи данных в сети.

Состояния (S)

Состояние сети отражает текущее окружение ВБС, включая такие параметры, как положение ВБС в трехмерном пространстве и уровни мощности, выделенные различным пользователям в сети. Таким образом, состояние сети служит основой для принятия решений в системе ретрансляции на базе ВБС. В момент времени t состояние обозначается как s_t , и может быть записано в виде выражения:

$$s_t = \{x_a^t, y_a^t, h_a^t, p_m^t\}, \quad (9)$$

где (x_a^t, y_a^t, h_a^t) – положение ВБС во временном шаге t ; p_m^t – мощность, выделенная для пользователя m .

Действие (A)

На каждом шаге ВБС выбирает из пространства действий такое, которое включает и перемещение в новое состояние, и распределение мощности. Каждое действие приводит к переходу в новое состояние системы. Действие состоит из двух частей: перемещение в следующее состояние ВБС и распределение мощности. Перемещение ВБС включает движение вперед, назад, вправо, влево, вверх и вниз. Распределение мощности для каждого пользователя может увеличиваться или уменьшаться. Действие агента в момент времени t обозначается как a_t , и включает как изменение положения ВБС, так и регулировку мощности для пользователей.

Вознаграждение

Цель алгоритма DDPG – максимизация накопленного вознаграждения за время обучения (поставлена в уравнении (7)). Чтобы ее достичь, функция вознаграждения должна учитывать общую скорость передачи данных в системе, включая как GBS, так и ВБС. С другой стороны, после определения количества и расположения пользователей, подключенных к GBS или ВБС, на основе зоны покрытия GBS, суммарная скорость передачи данных в нисходящем канале для пользователей GBS остается неизменной. Следовательно, этот компонент можно исключить из функции вознаграждения.

Таким образом, вознаграждение r_t на каждом шаге времени t может быть вычислено на основе общей скорости передачи данных в нисходящем канале пользователей, подключенных к ВБС:

$$r_t = R_t^{\text{ВБС}} - \lambda \Delta N_t, \quad (10)$$

где $R_t^{\text{ВБС}}$ – общая скорость передачи данных в нисходящем канале пользователей, подключенных к ВБС; λ – штрафной коэффициент, когда ВБС не подключается ко всем пользователям; ΔN_t – количество неподключенных пользователей.

Вероятность перехода (P)

Вероятность перехода $P(s_{t+1}|s_t, a_t)$ из текущего состояния s_t , к следующему состоянию s_{t+1} при выполнении действия a_t определяется как:

$$P(s_{t+1}|s_t, a_t) = \begin{cases} 1, & s_{t+1} = s_t + a_t \\ 0, & \text{другие} \end{cases}. \quad (11)$$

Два основных компонента алгоритма DDPG – сеть Actor и сеть Critic – играют важную роль в оптимизации стратегии действий и оценке значений модели. DDPG использует архитектуру Actor – Critic для решения задачи оптимизации в непрерывном пространстве действий. Обе сети являются глубокими нейронными сетями (DNN, аббр. от англ. Deep

Neural Network) и работают совместно для обучения оптимальной стратегии.

Сеть Actor принимает состояние системы s_t в виде входных данных и генерирует действие a_t , которое, как ожидается, будет оптимальным для повышения общей скорости передачи данных.

Политика, параметризованная нейросетью с параметрами θ^π , определяет, какие действия должны быть предприняты агентом:

$$a_t = \pi(s_t | \theta^\pi) + \xi, \quad (12)$$

где $\pi(s_t | \theta^\pi)$ – политика агента, которая определяет оптимальные действия, включая перемещение ВБС и распределение мощности, на основе текущего состояния; θ^π – параметры сети Actor, которые обновляются в процессе обучения; ξ – шум.

Сеть Actor обучается с использованием DDPG, обновляя параметры θ^π так, чтобы максимизировать ожидаемое суммарное вознаграждение. Градиентная функция $\nabla_{\theta^\pi} J(\theta^\pi)$ обновления параметров сети Actor определяется в работе [9].

Сеть Critic оценивает действия, выбранные сетью Actor, и предоставляет обратную связь. Ее основная задача – вычислить ценность состояния или действия с использованием Q -функции, которая оценивает, насколько выгодно выбранное действие в данной ситуации. Таким образом, сеть Critic играет ключевую роль в обучении, помогая сети Actor выбирать более эффективные действия на основе актуальных оценок ценности.

В процессе обучения сеть Critic вычисляет целевое значение y_t , которое используется для корректировки ее параметров, минимизируя ошибку между текущей оценкой $Q(s_t, a_t | \theta^Q)$ и целевым значением:

$$y_t = r_t + \gamma Q(s_{t+1}, \pi(s_{t+1} | \theta^{\pi'}), \theta^{Q'}), \quad (13)$$

где r_t – это вознаграждение, полученное на шаге времени t ; $Q(s_{t+1}, \pi(s_{t+1} | \theta^{\pi'}), \theta^{Q'})$ – это оценка ценности для состояния s_{t+1} , вычисленная с использованием целевой сети Actor $\pi(s_{t+1} | \theta^{\pi'})$ и целевой сети Critic $Q(s_{t+1}, a_{t+1} | \theta^{Q'})$; γ – это коэффициент дисконтирования.

Сеть Critic использует функцию потерь для обновления своих параметров θ^Q , определяемую следующим образом:

$$L(\theta^Q) = \frac{1}{D} \sum_{j=1}^D (y_i - Q(s_j, a_j | \theta^Q))^2, \quad (14)$$

где D – общее количество образцов опыта в буфере.

Чтобы стабилизировать обучение, используется фактор мягкого обновления τ для актуализации параметров целевых сетей $\theta^{\pi'}$ и $\theta^{Q'}$:

$$\begin{aligned} \theta^{Q'} &\leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}, \\ \theta^{\pi'} &\leftarrow \tau \theta^\pi + (1 - \tau) \theta^{\pi'}. \end{aligned} \quad (15)$$

Процесс обучения DDPG включает сбор опыта из среды и его сохранение в буфер воспроизведения H . Случайная мини-партия извлекается из буфера воспроизведения для обновления сетей Critic и Actor. Первая обновляется с использованием ошибки, вторая – с использованием градиента функции действия. Целевые сети обновляются постепенно под воздействием фактора мягкого обновления τ , чтобы уменьшить колебания. Модели обучения и значения вознаграждения сохраняются для последующего анализа и применения. Этот процесс способствует оптимизации стратегии и анализу ее эффективности.

5. Результаты моделирования

В этом разделе приведены результаты моделирования, разработанного для проверки предлагаемой конструкции оптимизации (параметры моделирования представлены в таблице 2). Рассматривается 1000 пользователей, равномерно распределенных по территории размером 2000×2000 м. Высота полета ВБС находится в диапазоне от 200 до 1000 м в зависимости от местных правил и конструктивных особенностей.

ТАБЛИЦА 1. Параметры моделирования

TABLE 1. Simulation Parameters

Параметр	Значение	Параметр	Значение
P_{GBS}	40 дБм	$P_{ВБС}$	20 дБм
a	9,61	b	0,16
β_0	-30 дБ	κ	-20 дБ
f_g	3,5 ГГц	SNR_0	20 дБ
σ^2	-130 дБм	$\hat{\alpha}$	-2,3

Для экспериментального моделирования фактор дисконтирования вознаграждения составляет $\gamma = 0,99$, фактор мягкого обновления – $\tau = 0,005$, размер буфера воспроизведения – $H = 10^5$, общее количество образцов опыта в буфере – $D = 64$. Шум Орнштейна – Уленбека ξ следует нормальному распределению со средним значением 0 и дисперсией 0,5 [10]. В процессе обучения алгоритма DDPG используется 500 эпизодов, каждый из которых состоит из 100 шагов. Штрафной коэффициент – $\lambda = 10$.

В таблице 2 описано влияние радиуса покрытия GBS на высоту полета ВБС. В целом, по мере уменьшения радиуса покрытия GBS, высота полета ВБС будет постепенно увеличиваться, чтобы обеспечить связь со всеми оставшимися пользователями. Увеличение высоты полета также повышает вероятность соединения в пределах LOS, тем самым снижая потери при передаче и улучшая качество

обслуживания. Кроме того, увеличение высоты полета ВБС позволяет расширить зону покрытия, что помогает подключаться к удаленным пользователям.

ТАБЛИЦА 2. Влияние различных радиусов покрытия GBS на высоту полета ВБС

TABLE 2. Impact of GBS Coverage Radius on the Flight Altitude of an ABS

Радиус покрытия GBS (м)	600	700	750	800	900	1000
Высота полета ВБС (м)	810	612	560	512	477	450

Таблица 3 описывает влияние радиуса покрытия GBS на количество подключенных пользователей и общую скорость передачи данных для пользователей, подключенных к GBS и ВБС, а также для всей системы в целом. Очевидно, что при радиусе покрытия GBS, равном 750 м, общая скорость передачи данных системы достигает максимального значения.

ТАБЛИЦА 3. Радиус покрытия GBS и его влияние на количество пользователей и суммарную скорость передачи данных

TABLE 3. The Coverage Radius of the GBS and Its Effect on the Number of Users and Overall Data Transmission Rate

Радиус покрытия GBS (м)	600	700	750	800	900	1000
N_n (человек)	29	35	38	44	59	61
R_{GBS} (бит/с)	149,3	176,2	191,4	217,2	218,6	289,8
N_m (человек)	71	65	62	58	41	39
$R_{ВБС}$ (бит/с)	347,3	322,5	308,2	278,2	206,2	194
R (бит/с)	496,6	498,7	499,6	495,2	487,8	483,8

На рисунке 1 показаны распределения пользователей GBS и ВБС, при радиусе покрытия GBS, равном 750 м. Красный квадрат представляет трехмерное положение ВБС в пространстве, а синий ромб, расположенный в центре рассматриваемой зоны, – положение GBS. Синие и красные точки обозначают пользователей, подключенных к GBS и ВБС соответственно. Синий пунктирный круг обозначает радиус покрытия GBS.

На рисунке 2 показана скорость передачи данных системы при использовании методов глубокого Q -обучения (DQN, аббр. от англ. Deep Q Network) и DDPG. Последний превосходит DQN в задачах распределения ресурсов и оптимизации трехмерного местоположения ВБС благодаря своей способности работать в непрерывном пространстве. DQN применим только в пространстве дискретных действий [11], тогда как DDPG может регулировать положение ВБС и распределение мощности в соответствии с непрерывными значениями, что повышает общую скорость передачи данных. Кроме того, DDPG использует сети Critic и Actor, которые способствуют лучшему обучению в сложных условиях [12].

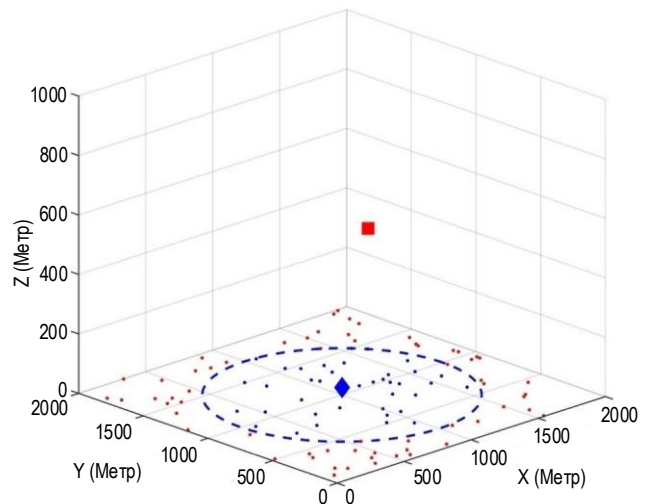


Рис. 1. Распределение пользователей, GBS и ВБС

Fig. 1. Distribution of Users, GBS, and ABS

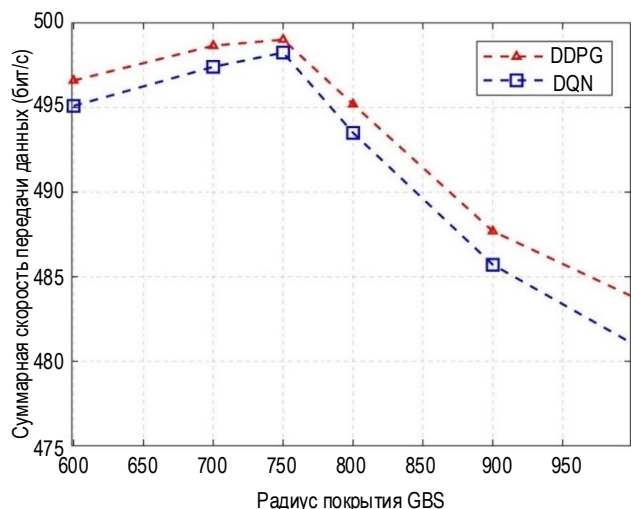


Рис. 2. Суммарная скорость передачи данных при использовании алгоритмов DQN и DDPG

Fig. 2. The Total Data Transmission Rate Achieved by Using DQN and DDPG Algorithms

6. Выводы

В данной статье предлагается модель и метод для изучения проблем распределения пользователей, мощности и оптимизации местоположения ВБС в гибридных сетях. Рассматривая радиус покрытия GBS как динамический параметр, можно определить количество и местоположение пользователей, подключенных к GBS или ВБС, в зависимости от каждого радиуса.

При распределении мощности пользователей, подключенных к GBS, всем пользователям выделяется одинаковая мощность. Для подключенных к ВБС рекомендуется использовать алгоритм DDPG для поиска оптимального положения ВБС и соответствующего распределения мощности. На основе анализа изменения суммарной скорости передачи

данных системы при различных радиусах покрытия GBS определяется оптимальный радиус, при котором достигается ее максимальное значение.

Кроме того, в статье также сравниваются алгоритмы DDPG и DQN. Результаты показывают, что DDPG превосходит DQN в предложенной задаче. Это можно объяснить разницей в том, как два алгоритма ищут решения: DDPG работает в непрерывном пространстве действий, что позволяет находить более оптимальные решения, тогда как DQN выполняет поиск только в дискретном пространстве, что приводит к ограничениям в точности и гибкости.

Особое внимание уделено зависимости между радиусом покрытия GBS и высотой полета ВБС. С увеличением радиуса покрытия GBS высота полета ВБС снижается, и наоборот. Это демонстрирует существование важного компромисса между этими параметрами, который должен быть учтен при оптимизации развертывания системы в условиях ограниченных ресурсов. Предложенная модель и методы могут быть полезны для разработки более эффективных сетевых решений в условиях ограниченных ресурсов (развертывание ВБС в удаленных или труднодоступных районах), что значительно улучшит качество обслуживания пользователей и повысит общую скорость передачи данных.

Список источников

1. Чан Т.З., Кучерявый А.Е. Оптимизация использования ресурсов воздушных базовых станций на основе методов искусственного интеллекта // Труды учебных заведений связи. 2025. Т. 11. № 1. С. 62–68. DOI:10.31854/1813-324X-2025-11-1-62-68. EDN:RVENV
2. Кузнецов К.А., Парамонов А.И., Мутханна А.С.А., Кучерявый А.Е. Модель и методы маршрутизации трафика в сети связи с использованием БПЛА // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 62–72. DOI:10.31854/1813-324X-2024-10-4-62-72. EDN:VYMCTD
3. Дунайцев Р.А., Бородин А.С., Кучерявый А.Е. Интегрированная сеть космос-воздух-земля-море как основа сетей связи шестого поколения // Электросвязь. 2022. № 10. С. 5–8. DOI:10.34832/ELSV2022.35.10.001. EDN:QCGIPI
4. Кучерявый А.Е., Парамонов А.И., Маколкина М.А., Мутханна А.С.А., Выборнова А.И., Дунайцев Р.А. и др. Трехмерные многослойные гетерогенные сверхплотные сети // Информационные технологии и телекоммуникации. 2022. Т. 10. № 3. С. 1–12. DOI:10.31854/2307-1303-2021-10-3-1-12. EDN:LHLYEM
5. Mozaffari M., Saad W., Bennis M., Debbah M. Unmanned Aerial Vehicle With Underlaid Device-to-Device Communications: Performance and Tradeoffs // IEEE Transactions on Wireless Communications. 2016. Vol. 15. Iss. 6. PP. 3949–3963. DOI:10.1109/TWC.2016.2531652
6. Ali M.A., Jamalipour A. UAV placement and power allocation in uplink and downlink operations of cellular network // IEEE Transactions on Communications. 2020. Vol. 68. Iss. 7. PP. 4383–4393. DOI:10.1109/TCOMM.2020.2983671. EDN:MEPFGQ
7. 3GPP TR 38.901. Study on channel model for frequencies from 0.5 to 100 GHz. 2018. URL: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3173> (Accessed 26.02.2025)
8. Shannon C.E. A Mathematical Theory of Communication // The Bell System Technical Journal. 1948. Vol. 27. Iss. 3. PP. 379–423. DOI:10.1002/j.1538-7305.1948.tb01338.x
9. Liu C.H., Chen Z., Tang J., Xu J., Piao C. Energy-Efficient UAV Control for Effective and Fair Communication Coverage: A Deep Reinforcement Learning Approach // IEEE Journal on Selected Areas in Communications. 2018. Vol. 36. Iss. 9. PP. 2059–2070. DOI:10.1109/JSAC.2018.2864373
10. Seid A.M., Boateng G.O., Anokye S., Kwantwi T., Sun G., Liu G. Collaborative Computation Offloading and Resource Allocation in Multi-UAV-Assisted IoT Networks: A Deep Reinforcement Learning Approach // IEEE Internet of Things Journal. 2021. Vol. 8. Iss. 15. PP. 12203–12218. DOI:10.1109/JIOT.2021.3063188
11. Mnih V., Kavukcuoglu K., Silver D., Rusu A.A., Veness J., Bellemare M.G., et al. Hassabis D. Human-level control through deep reinforcement learning // Nature. 2015. Vol. 518. PP. 529–533. DOI:10.1038/nature14236
12. Lillicrap T.P., Hunt J.J., Pritzel A., Heess N., Erez T., Tassa Y., et al. Continuous control with deep reinforcement learning // arXiv preprint arXiv:1509.02971. 2015. DOI:10.48550/arXiv.1509.02971

References

1. Tran T.D., Koucheryavy A.E. Resource Optimization of Airborne Base Stations Using Artificial Intelligence Methods. *Proceedings of Telecommunication Universities*. 2025;11(1):62–68. (in Russ.) DOI:10.31854/1813-324X-2025-11-1-62-68. EDN:RVENV
2. Kuznetsov K.A., Paramonov A.I., Muthanna A.S.A., Koucheryavy A.E. Model and Methods of Traffic Routing in a Communication Network Using UAVs. *Proceedings of Telecommunication Universities*. 2024;10(4):62–72. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-62-72. EDN:VYMCTD
3. Dunaitsev R.A., Borodin A.S., Koucheryavy A.E. Space-Air-Ground-Sea Integrated Networking as a Basis For 6G Networks. *Electrosvyaz*. 2022;10:5–8. (in Russ.) DOI:10.34832/ELSV2022.35.10.001. EDN:QCGIPI
4. Koucheryavy A., Paramonov A., Makolkina A., Muthanna A. S. A., Vybornova A., Dunaytsev R., et al. 3 Dimension Multilayer Heterogenous Ultra Dense Networks. *Telecom IT*. 2022;10(3):1–12. (in Russ.) DOI:10.31854/2307-1303-2021-10-3-1-12. EDN:LHLYEM

5. Mozaffari M., Saad W., Bennis M., Debbah M. Unmanned Aerial Vehicle With Underlaid Device-to-Device Communications: Performance and Tradeoffs. *IEEE Transactions on Wireless Communications*. 2016;15(6):3949–3963. DOI:10.1109/TWC.2016.2531652
6. Ali M.A., Jamalipour A. UAV placement and power allocation in uplink and downlink operations of cellular network. *IEEE Transactions on Communications*. 2020;68(7):4383–4393. DOI:10.1109/TCOMM.2020.2983671. EDN:MEPFGQ
7. 3GPP TR 38.901. *Study on channel model for frequencies from 0.5 to 100 GHz*. 2018. URL: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3173> [Accessed 26.02.2025]
8. Shannon C.E. A Mathematical Theory of Communication. *The Bell System Technical Journal*. 1948;27(3):379–423. DOI:10.1002/j.1538-7305.1948.tb01338.x
9. Liu C.H., Chen Z., Tang J., Xu J., Piao C. Energy-Efficient UAV Control for Effective and Fair Communication Coverage: A Deep Reinforcement Learning Approach. *IEEE Journal on Selected Areas in Communications*. 2018;36(9):2059–2070. DOI:10.1109/JSAC.2018.2864373.
10. Seid A.M., Boateng G.O., Anokye S., Kwantwi T., Sun G., Liu G. Collaborative Computation Offloading and Resource Allocation in Multi-UAV-Assisted IoT Networks: A Deep Reinforcement Learning Approach. *IEEE Internet of Things Journal*. 2021;8(15):12203–12218. DOI:10.1109/JIOT.2021.3063188
11. Mnih V., Kavukcuoglu K., Silver D., Rusu A.A., Veness J., Bellemare M.G., et al. Hassabis D. Human-level control through deep reinforcement learning. *Nature*. 2015;518:529–533. DOI:10.1038/nature14236
12. Lillicrap T.P., Hunt J.J., Pritzel A., Heess N., Erez T., Tassa Y., et al. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*. 2015. DOI:10.48550/arXiv.1509.02971

Статья поступила в редакцию 27.02.2025; одобрена после рецензирования 10.04.2025; принята к публикации 11.04.2025.

The article was submitted 27.02.2025; approved after reviewing 10.04.2025; accepted for publication 11.04.2025.

Информация об авторах:

**ЧАН
Тунг Зыонг**

аспирант кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0006-0080-9477>

**КУЧЕРЯВЫЙ
Андрей Евгеньевич**

доктор технических наук, профессор, заведующий кафедрой сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича
 <https://orcid.org/0000-0003-4479-2479>

Кучерявый А.Е. является членом редакционного совета журнала «Труды учебных заведений связи» с 2016 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.

Koucheryavy A.E. has been a member of the journal "Proceedings of Telecommunication Universities" Editorial Council since 2016, but has nothing to do with the decision to publish this article. The article has passed the review procedure accepted in the journal. The authors have not declared any other conflicts of interest.

Научная статья

УДК 004.932.2

<https://doi.org/10.31854/1813-324X-2025-11-2-41-48>

EDN:SCTROE



Интерактивная обработка изображений для робастного распознавания геометрических примитивов

Альфия Асхатовна Диязитдинова, a.diyazitdinova@psuti.ru

Поволжский государственный университет телекоммуникаций и информатики,
Самара, 443010, Российская Федерация

Аннотация

Распознавание геометрических примитивов используется при обработке изображений для решения задач, связанных с машинным обучением, сокращением области анализа и уменьшением вычислительной сложности. Одной из проблем распознавания примитивов является зависимость результатов обработки от таких внешних факторов как: широкий диапазон изменения яркости, контрастности, наличие помех, посторонних объектов или загрязнения. Отдельной задачей является определение геометрического положения примитива на изображении, которое определяется смещением, поворотом и масштабом или параметрами более сложной математической модели трансформации. Интерактивный способ обработки позволяет обеспечить робастность к пространственно-яркостным искажениям и различным помехам.

Целью настоящей статьи является повышение качества распознавания геометрических примитивов на изображениях за счет интерактивной обработки.

Сущность предлагаемого решения заключается в проведении двух этапов: этап предварительной обработки в интерактивном режиме и этап оценки геометрических параметров примитива с автоматическим удалением импульсных помех. На первом этапе выбирается порог для детектирования контура примитива и ограничение области анализа (выбор фрагмента) на изображении. Определение этих параметров проводится с помощью графического интерфейса в интерактивном режиме (например, изменение порога детектирования практически мгновенно отображает распознанные контуры на изображении). На втором этапе в соответствии с формой примитива выделяется область интереса, что убирает импульсные помехи (контурные точки, не принадлежащие примитиву), а по точкам в области интереса оцениваются параметры примитива методом наименьших квадратов. Разработанный алгоритм **имеет реализацию** в виде программы с графическим интерфейсом. **Эксперименты** для проверки разработанного алгоритма показали удовлетворительное распознавание геометрического примитива «окружность» на различных типах изображений, содержащих дорожный знак, частицу полимерного геля, феррулу оптического адаптера. **Научная новизна** решения состоит в возможности распознавания примитивов, которое робастно к пространственно-яркостным преобразованиям (масштабу, смещениям, неравномерности яркости и т. д.) и другим помехам.

Теоретическая значимость состоит в расширении возможностей методов распознавания за счет интерактивного выбора параметров на этапе предобработки.

Практическая значимость заключается в упрощении алгоритмов обработки изображений, которые используются при решении прикладных задач (подготовки данных машинного обучения, обработка методами оптической микрометрии), не требующих распознавания в режиме реального времени.

Ключевые слова: геометрический примитив, распознавание, робастность, интерактивный, изображение, контур

Ссылка для цитирования: Диязитдинова А.А. Интерактивная обработка изображений для робастного распознавания геометрических примитивов // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 41–48. DOI:10.31854/1813-324X-2025-11-2-41-48. EDN:SCTROE

Original article

<https://doi.org/10.31854/1813-324X-2025-11-2-41-48>

EDN:SCTROE

Interactive Image Processing for Robust Geometric Primitives Recognition

 Alphiya A. Diyazitdinova, a.diyazitdinova@psuti.ru

Povolzhskiy State University of Telecommunications and Informatics,
Samara, 443010, Russian Federation

Annotation

Relevance. Recognition of geometric primitives is used in image processing to solve problems related to the development of machine learning algorithms, reducing the analysis area and reducing computational complexity. One of the problems of primitives recognition is the resulting dependence from such external factors as: a wide range of changes in brightness, contrast and the interference caused by uneven lighting, foreign objects or pollution. A separate task is the geometric position estimation of the primitive in the image, which is defined by offsets, rotation and scale or parameters of a more complex mathematical transformation model. A wide class of tasks is not limited by the requirement of automatic processing in real time. Therefore, these problems can be solved by an interactive parameters setting. The interactive processing method ensures robustness to spatial-luminance distortions and various interferences.

The article purpose is to improve the quality of recognition of geometric primitives (using the example of a circle) in images through interactive (visually controlled by the operator) processing.

The proposed solution essence is two processing stages: the stage of pre-processing in interactive mode and the stage of estimating the geometric parameters of the primitive with automatic removal of impulse noise. At the first stage, a threshold is selected for detecting the contour of a primitive and limiting the analysis area (selecting a fragment) in the image. These parameters are determined by a graphical interface in interactive mode (for example, changing the detection threshold almost instantly displays recognized contours in the image). At the second stage, in accordance with the shape of the primitive, an area of interest is selected, which removes impulse noise (contour points that do not belong to the primitive), and the parameters of the primitive are estimated based on the point of the area of interest using the least squares method. The developed algorithm has an implementation as a program with a graphical interface. **Experiments** to test the developed algorithm showed satisfactory recognition of the geometric primitive "circle" on various types of images containing a road sign, a polymer gel particle, and an ferrule end face. **The scientific novelty** of the solution lies in the possibility of recognizing primitives, which is robust to spatial-brightness transformations (scale, displacements, brightness unevenness, etc.) and other noise.

The theoretical significance lies in expanding the capabilities of recognition methods (in particular, primitives such as "circles") through interactive selection of parameters at the preprocessing stage.

The practical significance lies in the simplicity of image processing algorithms that are used to solve applied problems (preparing machine learning data, processing by optical micrometry methods) that do not require real-time recognition.

Keywords: geometric primitive, recognition, robust, interactive, image, contour

For citation: Diyazitdinova A.A. Interactive Image Processing for Robust Geometric Primitives Recognition. *Proceedings of Telecommunication Universities*. 2025;11(2):41–48. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-41-48. EDN:SCTROE

1. Введение

Распознавание геометрических примитивов на изображении широко применяется при решении научных и прикладных задач. В системах автопи-

лотирования автомобильных средств геометрические примитивы соответствуют дорожным знакам [1, 2], в химической промышленности по изображениям с микроскопа оценивают концентрацию компонентов в растворах смесей [3, 4], в телеком-

муникациях по изображениям торцевых поверхностей оптических адаптеров принимаются решения о вносимом затухании мощности сигнала [5, 6] и т. д.

Общая схема распознавания геометрических примитивов заключается в предварительном детектировании контурных линий [7], делении изображения на фрагменты и вынесении решения о наличии или отсутствии геометрического примитива (или одного из нескольких возможных примитивов) на основе анализа контурных точек, содержащихся внутри фрагмента. Такая схема распознавания имеет множество особенностей, которые сильно влияют на результат.

Во-первых, детектирование контурных линий зависит от выбранного порога, по которому определяется, соответствует ли пиксель изображения контуру или не соответствует. На детектирование контуров влияет широкий диапазон изменения яркости, контрастности и наличия помех, обусловленных неравномерностью освещения, посторонними объектами и загрязнением.

Во-вторых, будет значительное количество фрагментов на изображении. Примитивы даже внутри одного изображения могут быть разного размера. Чтобы считывать геометрическое положение и размеры примитивов, которые определяются смещением, поворотом и масштабом или параметрами более сложной математической модели трансформации, необходимо анализировать фрагменты различных размеров, в том числе пересекающиеся между собой, так как внутри фрагмента может содержаться только часть примитивов, или наоборот, фрагмент захватывает не только примитив, но и «мешающие» контурные точки.

Эти особенности ведут к низкому качеству распознавания и предъявляют высокие требования к производительности вычислительной аппаратуры. Однако широкий класс прикладных задач не ограничен требованиями автоматического распознавания и обработкой в режиме реального времени. Проблема выбора порога для детектирования контуров и выбора области анализа может быть решена с помощью оператора (эксперта) интерактивным способом. Такой подход обеспечивает возможности разработки универсального алгоритма распознавания примитивов, робастного к пространственно-яркостным искажениям и различным помехам.

2. Обзор существующих работ

В работе [8] представлена методика распознавания отрезков прямых линий, формирующих схему контактных дорожек печатных плат. Основная особенность методики заключается в предварительной нормировке яркости изображения на

среднее значение и среднеквадратичное отклонение, что позволяет использовать одинаковый порог для детектирования контуров печатных плат. Методика ориентирована на обработку изображений, полученных в лабораторных условиях со стандартизированным освещением, поэтому использование ее для детектирования контуров на изображениях, снятых в условиях меняющейся освещенности, приводит к неудовлетворительным для практики результатам.

В работах [3, 4] рассматривается задача оптической микрометрии, в которой анализируется гранула полимера, снятая под микроскопом. Особенности обрабатываемых изображений являются:

- единственность объекта (гранулы полимера) на фото;
- наличие помех, возникающих из-за посторонних объектов.

Одной из основных проблем распознавания примитива, представляющего собой окружность, которая ограничивает гранулу, является «мешающие» контуры посторонних объектов. Решается она путем уменьшения размера изображения, на которых «мешающие контуры» перестают детектироваться (посторонние объекты много меньше гранул и с уменьшением размера становятся почти незаметными). Однако такой подход не является универсальным решением, которое могло бы обеспечить подавление контуров от загрязнений относительно крупных посторонних объектов.

В работах [5, 6] рассматривается задача распознавания границы «оболочка волокна-феррул» на изображениях торцевой поверхности оптического адаптера, которая приблизительно описывается окружностью. Для распознавания предложено интерактивным способом указывать четыре точки искомой границы, которые аппроксимируются по методу наименьших квадратов. Изображение торцевой поверхности получается в микроскопе в условиях стабильного освещения, что обеспечивает возможность использования заранее выбранного порога для детектирования контура. Недостатком предложенного способа является длительное время выбора точек и необходимость повторения процедуры, если точки были указаны с высокой погрешностью (точки не совпадают с искомой границей).

В работе [9] рассматривается задача сегментации изображений, в которой авторы утверждают, что автоматическое распознавание возможно построить для изображений объектов, полученных при определенных условиях освещенности и имеющих заранее известную сцену (например, однотипные объекты на однородном фоне), а сами геометрические объекты должны быть известны заранее и меняться в узком диапазоне. В противном случае обработка должна вестись интерак-

тивным образом – оператор должен выделить область анализа, тем самым обеспечивая стабильность результатов.

Таким образом, в существующих работах есть как утверждения, так и аргументы, приведенные в пользу необходимости использования интерактивной обработки изображений для получения качественного результата распознавания геометрических примитивов.

3. Описание процедуры интерактивной обработки и алгоритма распознавания

Интерактивная обработка позволяет решить две проблемы:

- выбор порога для детектирования контура геометрического примитива;
- ограничение области анализа на изображении.

На рисунках 1 и 2 показаны примеры влияния порога на детектирование контуров. При выборе низкого порога наблюдается значительное количество контурных точек, которые будут приводить к увеличению погрешности оценки геометрических размеров и положения примитивов (см. рисунки 1b и 2b). Порог, значение которого выбирается с учетом визуального контроля, позволяет достичь необходимого качества детектирования контуров, позволяющего проводить последующую автоматическую обработку (см. рисунки 1c и 2c).

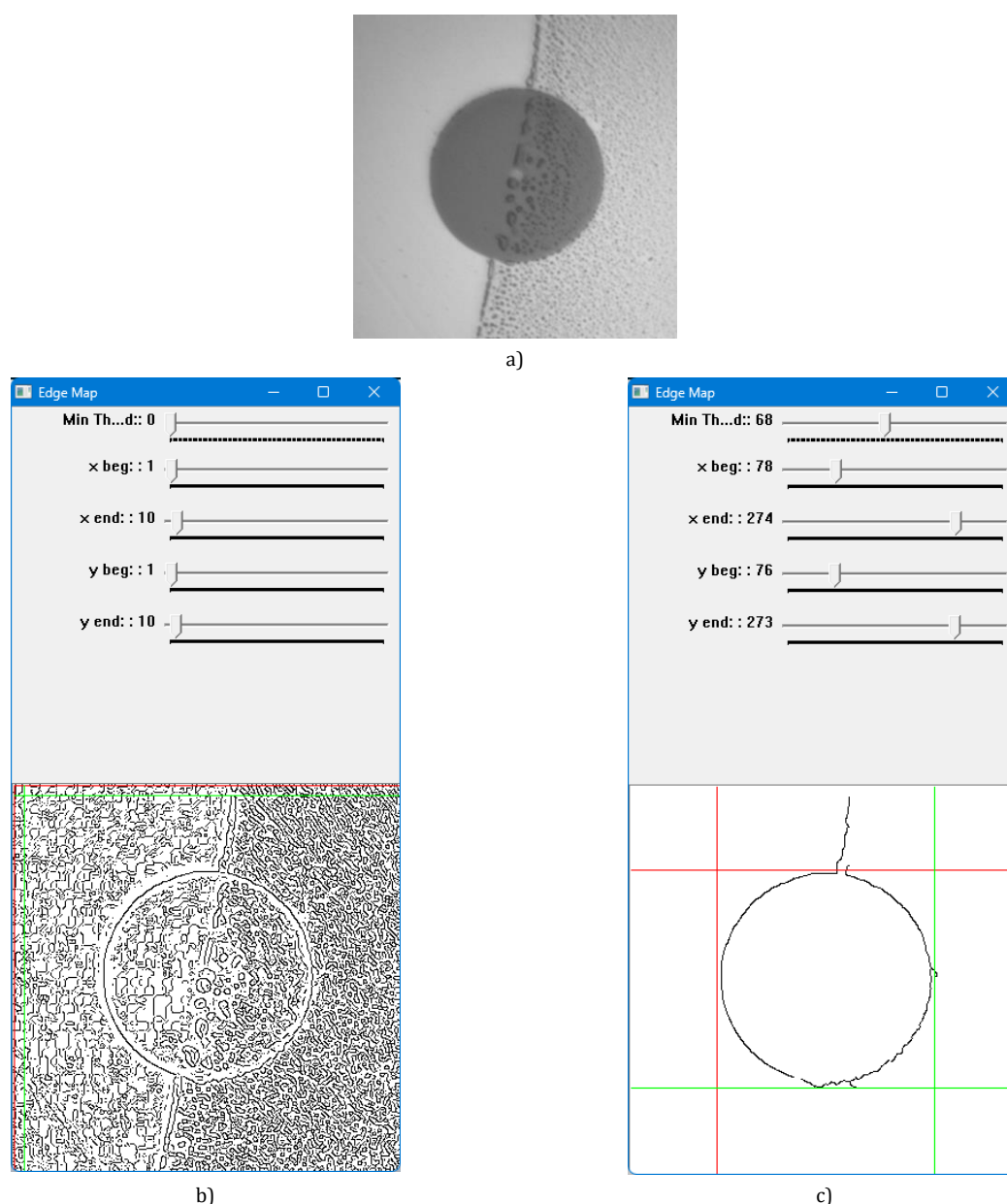
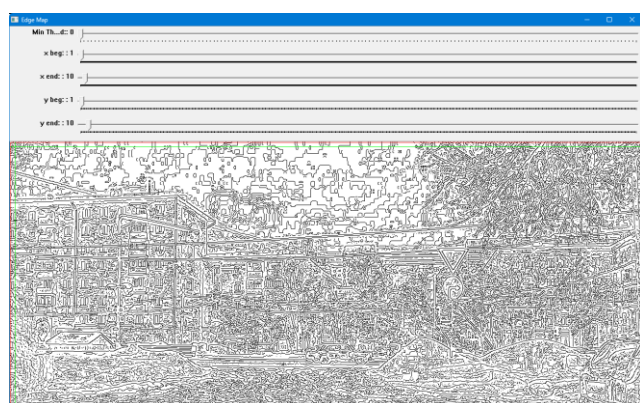


Рис. 1. Детектирование контура на изображении загрязненной торцевой поверхности оптического адаптера [10]: исходное изображение (a), контурные линии с низким порогом (b), с порогом, выбранным оператором (c)

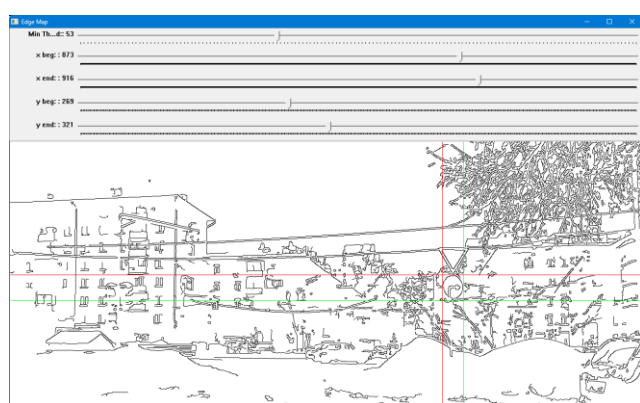
Fig. 1. Contour Detection on Dirty Ferrule End Face Image [10]: Source (a), Contour by Low Threshold (b), by Operator Defined Threshold (c)



a)



b)



c)

Рис. 2. Детектирование контура дорожного знака с изображения видеорегистратора [11]: исходное изображение (а), контурные линии с низким порогом (б), с порогом, выбранным оператором (с)

Fig. 2. Road Sign Contour Detection from Digital Video Recorder Image [11]: Source (a), Contour by Low Threshold (b), by Operator Defined Threshold (c)

В работе рассматривается распознавание геометрических примитивов, представляющих собой окружности. Алгоритм распознавания примитива разрабатывался с учетом его формы.

Исходными данными для работы алгоритма являются:

- 1) изображение с детектированными контурами;
- 2) прямоугольная область, определяемая двумя точками (x_b, y_b) , (x_e, y_e) (рисунок 3).

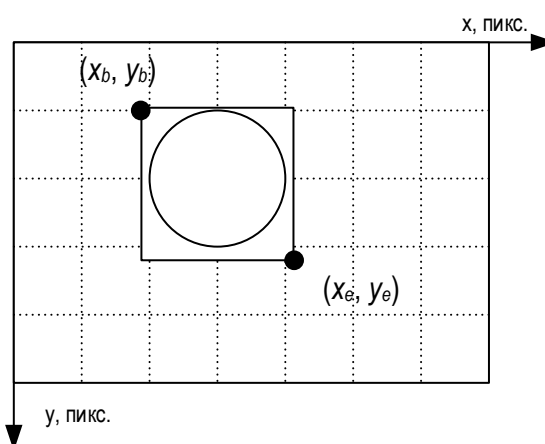


Рис. 3. Схематичное описание прямоугольной области с примитивом

Fig. 3. Rectangular Area with a Primitive

Алгоритм состоит из трех этапов.

Этап 1. Расчет координат четырех точек (рисунок 4):

$$\begin{aligned} x_1 &= (x_b + x_e) / 2 & y_1 &= y_b \\ x_2 &= x_e & y_2 &= (y_b + y_e) / 2 \\ x_3 &= (x_b + x_e) / 2 & y_3 &= y_e \\ x_4 &= x_b & y_4 &= (y_b + y_e) / 2 \end{aligned}$$

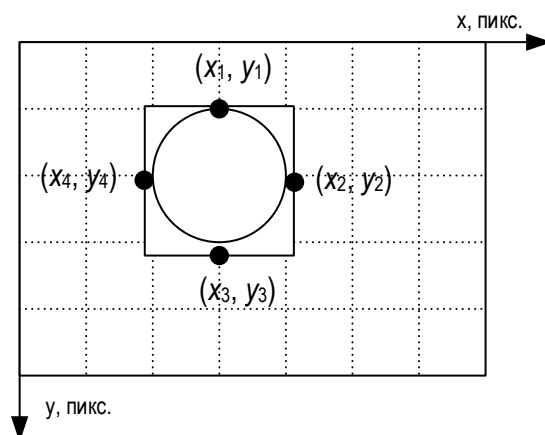


Рис. 4. Расположение точек $(x_1, y_1) \dots (x_4, y_4)$ относительно искомого примитива

Fig. 4. Location of Points $(x_1, y_1) \dots (x_4, y_4)$ Relatively a Primitive

Этап 2. Аппроксимация этих точек окружностью по методу наименьших квадратов [12].

Координаты центра и радиус окружности определяются по формулам:

$$x_0 = \frac{W_1 \cdot N_{22} - W_2 \cdot N_{12}}{\det(N)}, \quad (1)$$

$$y_0 = \frac{W_1 \cdot N_{11} - W_2 \cdot N_{21}}{\det(N)}, \quad (2)$$

$$R^2 = x_0^2 + y_0^2 + (S_{xx} + S_{yy} - 2(x_0 S_x + y_0 S_y)/n), \quad (3)$$

где $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$, $\det(N) = N_{11} \cdot N_{22} - N_{12} \cdot N_{21}$,

$$\begin{aligned}
N_{11} &= 2(S_{xx} - (S_x \cdot S_x) / n), \quad N_{12} = 2(S_{xy} - (S_x \cdot S_y) / n), \\
N_{21} &= 2(S_{xy} - (S_x \cdot S_y) / n), \quad N_{22} = 2(S_{yy} - (S_y \cdot S_y) / n), \\
W_1 &= S_{xxx} + S_{xyy} - (S_{xx} \cdot S_x) / n - (S_{yy} \cdot S_y) / n, \\
W_2 &= S_{xxy} + S_{yyx} - (S_{xx} \cdot S_y) / n - (S_{yy} \cdot S_x) / n, \\
S_x &= \sum_{i=1}^n x_i, \quad S_y = \sum_{i=1}^n y_i \\
S_{xx} &= \sum_{i=1}^n x_i^2, \quad S_{yy} = \sum_{i=1}^n y_i^2, \quad S_{xy} = \sum_{i=1}^n x_i \cdot y_i, \\
S_{xxx} &= \sum_{i=1}^n x_i^3, \quad S_{yyy} = \sum_{i=1}^n y_i^3, \quad S_{xxy} = \sum_{i=1}^n x_i^2 \cdot y_i, \\
S_{xyy} &= \sum_{i=1}^n x_i \cdot y_i^2,
\end{aligned}$$

n – количество точек ($n = 4$).

Этап 3. Уточнение параметров окружности.

Для уточнения использовались точки контура, которые удовлетворяют неравенству:

$$(k \cdot R)^2 \leq (x - x_0)^2 + (y - y_0)^2 \leq (m \cdot R)^2, \quad (4)$$

где k, m – множители, определяющие область вокруг окружности (рисунок 5).

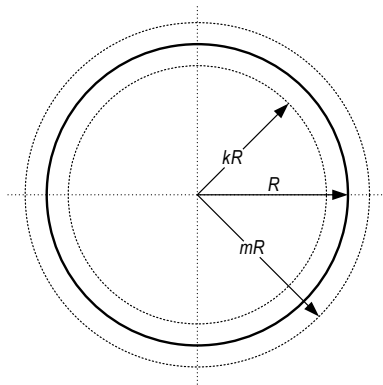


Рис. 5. Геометрический смысл множителей k, m

Fig. 5. Geometric Sense of Parameters k, m

Множители выбирались экспериментально. При проведении экспериментов $k = 0,95$, $m = 1,01$. Точки контура, удовлетворяющие неравенству (4), аппроксимируются окружностью по выражениям (1–3). Рассчитанные параметры определяют положение и размеры искомого геометрического примитива.

4. Результаты

Алгоритм был апробирован на трех видах изображений (рисунок 6) для распознавания примитивов, ограничивающих:

- 1) дорожный знак (см. рисунок 2);
- 2) частицу полимерного геля [4];
- 3) торцевую поверхность феррула (см. рисунок 1).

Как можно видеть, алгоритм распознавания, дополненный этапом интерактивного выбора порогов и области анализа, позволяет решать задачу робастного распознавания геометрических примитивов, отличающихся размеров, на различных видах изображений, полученных в условиях широкого диапазона изменения яркости, контрастности, неравномерности освещенности и при влиянии мешающих факторов (посторонних объектов, загрязнениях и т. д.).

Также интерактивная обработка обеспечивает контроль в режиме реального времени, и в случае неудовлетворительного распознавания процедуру можно повторить. Это удобно не только оператору, но и разработчику. Так как нет необходимости проводить глубокий анализ причин, которые привели к таким результатам.

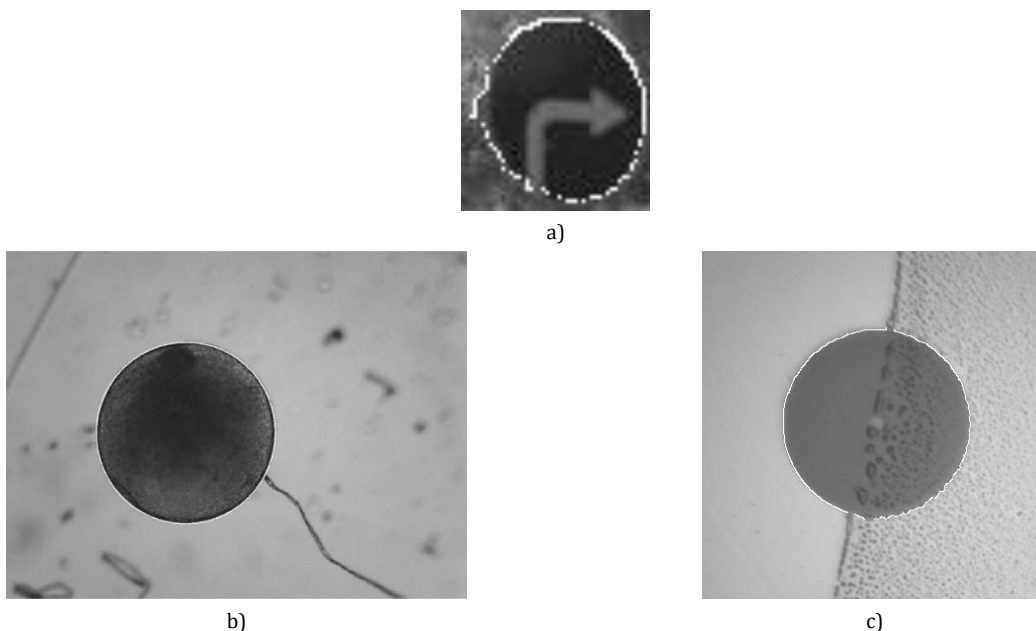


Рис. 6. Примеры распознавания геометрических примитивов для дорожного знака (а), частицы полимерного геля (б), и феррула (с)

Fig. 6. Geometric Primitives Recognition of Road Sign (a), Polymer Gel (b) and Ferrule End Face (c)

Заключение

Интерактивный способ позволяет существенно упростить алгоритмы обработки изображений, которые используются при решении прикладных задач, не требующих автоматического распознавания в режиме реального времени. Интерактивный способ позволяет практически мгновенно контролировать визуальные результаты, что обеспечивает удобство взаимодействия оператора с программой обработки. Преимущество предложенного подхода заключается в сокращении времени разработки, отладки и последующей модернизации программных средств, ориентированных

на решение прикладных задач, связанных с подготовкой данных для машинного обучения, вынесением решений методами оптической микрометрии (по изображениям с микроскопов), широко применяемых в телекоммуникациях и химической промышленности.

Дальнейшее исследование по данной теме будет вестись в направлении увеличения видов примитивов (окружности, треугольники, прямоугольники и т. д.), а также совместного применения с алгоритмами, повышающими стабильность оценки параметров, и алгоритмами машинного обучения для сокращения времени поиска областей анализа.

Список источников

1. Шульга Т.Э., Солопекин Д.А. Распознавание дорожных знаков российского образца с использованием нейронных сетей // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2024. № 2. С. 85–94. DOI:10.24143/2072-9502-2024-2-85-94. EDN:LTLMXA
2. Rajesh R., Rajeev K., Suchithra K., Lekhesh V.P., Gopakumar V., Ragesh N.K. Coherence vector of Oriented Gradients for traffic sign recognition using Neural Networks // Proceedings of the International Joint Conference on Neural Networks (San Jose, USA, 31 July – 05 August 2011). IEEE, 2011. PP. 907–910. DOI:10.1109/IJCNN.2011.6033318. EDN:PKQGGB
3. Атавуллаева Ш.Ш., Ферапонтов Н.Б., Тробов Х.Т., Турсунова Г.Х., Джуряева Р.А. Определение состава растворов смесей электролитов методом оптической микрометрии // Universum: химия и биология. 2024. № 11-2(125). С. 46–51. DOI:10.32743/UniChem.2024.125.11.18355. EDN:HYSJJG
4. Токмачев М.Г. Методика оценки объема гранулы полимерного геля в оптической микрометрии // Компьютерная оптика. 2023. Т. 47. № 6. С. 968–971. DOI:10.18287/2412-6179-CO-1271. EDN:OJTKGA
5. Bourdine A.V., Pashin S.S., Zaitseva E.S., Vasilets A.A., Antonov S.A. Fast and simple method for estimation of the insertion loss at the connection of singlemode optical fibers with contaminated ferrule end faces // Proceedings of the XVIIth International Scientific and Technical Conference "Optical Technologies for Telecommunications" (Kazan, Russian Federation, 19–21 November 2019). SPIE, 2020. Vol. 11516. P. 1151610. DOI:10.1117/12.2566456. EDN:GGNEOJ
6. Пашин С.С. Исследование потенциальных возможностей оценивания коэффициента передачи основной моды на основе анализа перекрытия радиального распределения полей в дискретном представлении // Инфокоммуникационные технологии. 2021. Т. 19. № 2. С. 172–178. DOI:10.18469/ikt.2021.19.2.05. EDN:QIKRUK
7. Canny J. A Computational Approach to Edge Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986. Vol. PAMI-8. Iss. 6. PP. 679–698. DOI:10.1109/TPAMI.1986.4767851
8. Chochia P.A. Image segmentation via contour tracking in application to the analysis of the photographs of electronic microcircuits // Journal of Communications Technology and Electronics. 2010. Vol. 55. Iss. 12. PP. 1466–1473. DOI:10.1134/S1064226910120193
9. Boykov Y., Jolly M.-P. Interactive Organ Segmentation Using Graph Cuts // Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2000, Pittsburgh, USA, 11–14 October 2000). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2000. Vol. 1935. PP. 276–286. DOI:10.1007/978-3-540-40899-4_28
10. Пашин С.С. Разработка методов контроля параметров передачи разъемных соединений оптических волокон кабелей связи. Дис. ... канд. техн. наук. Самара: Поволжский государственный университет телекоммуникаций и информатики, 2022. 221 с. EDN:KBOFZL
11. Konushin A., Shakhuro V. Traffic sign recognition // Graphics and Media Lab. UTL: <https://graphics.cs.msu.ru/projects/traffic-sign-recognition.html> (дата обращения 05.01.2025)
12. Каретин А.Н. МНК для аппроксимации данных окружностью. 2010. URL: https://mykaralw.narod.ru/articles/mnk_circle.pdf (дата обращения 05.01.2025)

References

1. Shulga T.E., Solopekin D.A. Recognition of Russian-style road signs using neural networks. *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*. 2024;2:85–94. (in Russ.) DOI:10.24143/2072-9502-2024-2-85-94. EDN:LTLMXA
2. Rajesh R., Rajeev K., Suchithra K., Lekhesh V.P., Gopakumar V., Ragesh N.K. Coherence vector of Oriented Gradients for traffic sign recognition using Neural Networks. *Proceedings of the International Joint Conference on Neural Networks, 31 July – 05 August 2011, San Jose, USA*. IEEE; 2011. p.907–910. DOI:10.1109/IJCNN.2011.6033318. EDN:PKQGGB
3. Atavullaeva Sh., Ferapontov N., Trobov Kh., Tursunova G., Djuraeva R. Determination of the composition of solutions of electrolyte mixtures by optical pyrometry. *Universum: Chemistry and Biology*. 2024;11-2(125): 46–51 (in Russ.) DOI:10.32743/UniChem.2024.125.11.18355. EDN:HYSJJG

4. Tokmachev M.G. A method for estimating the polymer gel granule volume by an optical technique. *Computer Optics*. 2023;47(6):968–971. (in Russ.) DOI:10.18287/2412-6179-CO-1271. EDN:OJTKGA
5. Bourdine A.V., Pashin S.S., Zaitseva E.S., Vasilets A.A., Antonov S.A. Fast and simple method for estimation of the insertion loss at the connection of singlemode optical fibers with contaminated ferrule end faces. *Proceedings of the XVIIth International Scientific and Technical Conference "Optical Technologies for Telecommunications"*, 19–21 November 2019, Kazan, Russian Federation, vol.11516. SPIE; 2020. P. 1151610. DOI:10.1117/12.2566456. EDN:GGNEOJ
6. Pashin S.S. Research potential possibilities of estimating the basic mode transmission coefficient based on analysis overlapping of the radial distribution fields in discrete representation. *Infocommunication Technologies*. 2021;19(2):172–178. (in Russ.) DOI:10.18469/ikt.2021.19.2.05. EDN:QIKRUK
7. Canny J. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1986:PAMI-8(6):679–698. DOI:10.1109/TPAMI.1986.4767851
8. Chochia P.A. Image segmentation via contour tracking in application to the analysis of the photographs of electronic microcircuits. *Journal of Communications Technology and Electronics*. 2010;55(12):1466–1473. DOI:10.1134/S1064226910120193
9. Boykov Y., Jolly M.-P. Interactive Organ Segmentation Using Graph Cuts. *Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2000, 11–14 October 2000, Pittsburgh, USA. Lecture Notes in Computer Science*, vol.1935. Berlin, Heidelberg: Springer, 2000. p.276–286. DOI:10.1007/978-3-540-40899-4_28
10. Pashin S.S. *Development of methods for monitoring transmission parameters of detachable connections of optical fibers of communication cables*. PhD Thesis. Samara: Povolzhskiy State University of Telecommunications and Informatics Publ.; 2022, 221 p. (in Russ.) EDN:KBOFZL
11. Konushin A., Shakhuro V. Traffic sign recognition. *Graphics and Media Lab*. (in Russ.) URL: <https://graphics.cs.msu.ru/projects/traffic-sign-recognition.html> [Accessed 05.01.2025]
12. Karetin A.N. *MNC for approximation of data by a circle*. 2010. (in Russ.) URL: https://mykaralw.narod.ru/articles/mnk_circle.pdf [Accessed 05.01.2025]

Статья поступила в редакцию 03.02.2025; одобрена после рецензирования 13.02.2025; принята к публикации 28.03.2025.

The article was submitted 03.02.2025; approved after reviewing 13.02.2025; accepted for publication 28.03.2025.

Информация об авторе:

**ДИЯЗИТДИНОВА
Альфия Асхатовна**

кандидат технических наук, доцент кафедры информационных систем и технологий Поволжского государственного университета телекоммуникаций и информатики

 <https://orcid.org/0000-0001-8940-4543>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.

Научная статья

УДК 539.143

<https://doi.org/10.31854/1813-324X-2025-11-2-49-55>

EDN:YRLSIN



Оптимальный выбор диаметра лазерного луча в атмосферных сетях связи распределенного и мобильного типа

Афлатун Масим Маммадов, aflatunmasimoglu@gmail.com

Азербайджанский государственный экономический университет,
Баку, 1001, Азербайджан

Аннотация

Актуальность. Стремительное развитие технологии интернета вещей привело к экспоненциальному росту количества различных интеллектуальных технических средств, присоединяемых к интернету, что, в свою очередь, способствовало появлению большого объема данных, подлежащих передаче по системам связи. Электромагнитные сети связи обладают недостаточным потенциалом для решения данной задачи, наиболее эффективными здесь оказываются оптические системы передачи информации по атмосферным каналам. Однако оптические атмосферные системы связи подвержены влиянию атмосферных факторов, т. к. из-за поглощения, рассеяния и дифракции лазерный луч ослабляется по мощности и подвергается уширению. В исследуемой предметной области известны работы, посвященные расходимости луча, где рассматривается вопрос зависимости величины максимальной расходимости луча от таких показателей, как мощность лазерного источника, геометрическая длина канала, длина волны оптической радиации. Вместе с тем, в указанных работах рассматривается один канал лазерной сети связи, а вопрос о выборе диаметра лазерного луча во всех каналах лазерной многоканальной атмосферной сети не обсуждается.

Цель. Сформулирована задача оптимального выбора диаметра лазерного пучка в многоканальной системе атмосферной оптической связи распределенного типа с учетом расходимости луча. В отличие от известных работ, в которых решается задача оптимизации с учетом влияния расходимости луча, решение поставленной цели охватывает как стационарные, так и мобильные варианты реализации каналов сети всей многоканальной системы. **Решение** поставленной задачи осуществлено путем формирования единого целевого функционала и дальнейшей оптимизации для выявления оптимальной взаимосвязи между величиной радиуса луча на входе приемника канала и радиусом луча на выходе излучателя при подаче на рассматриваемый атмосферный канал с учетом возможности его уширения. Показано, что такое расширение диаметра лазерного луча по выявленному оптимальному закону позволяет достичь максимума средней величины интенсивности лазерного луча, переданного по атмосферному каналу ко всем приемникам системы. Проведенные **модельные исследования** предложенной методики учета уширения диаметра луча подтвердили возможность получения оптимального соотношения между основными показателями многоканальной лазерной атмосферной сети.

Научная новизна. Разработана математическая модель оптимизации выбора диаметра луча в многоканальной лазерной атмосферной сети.

Теоретическая и практическая значимость. Оптимизация предложенной модели позволила получить рациональное соотношение между основными показателями многоканальной лазерной атмосферной сети распределенного типа, что может найти применение при построении подобных систем.

Ключевые слова: оптимизация, расходимость луча, моделирование, атмосферная связь, интенсивность луча

Ссылка для цитирования: Маммадов А.М. Оптимальный выбор диаметра лазерного луча в атмосферных сетях связи распределенного и мобильного типа // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 49–55. DOI:10.31854/1813-324X-2025-11-2-49-55. EDN:YRLSIN

Original research

<https://doi.org/10.31854/1813-324X-2025-11-2-49-55>

EDN:YRLSIN

Optimal Choice of Laser Beam Diameter in Atmospheric Communication Networks of Distributed and Mobile Type

 Aflatun M. Mammadov, aflatunmasimoglu@gmail.com

Azerbaijan State Economic University,
Baku, 1001, Azerbaijan

Annotation

Relevance. The rapid development of the Internet of Things technology has led to an exponential increase in the number of different intelligent hardware connected to the Internet, which, in turn, has contributed to the emergence of a large amount of data to be transmitted over communication systems. Electromagnetic communication networks have insufficient potential to solve this problem, and optical information transmission systems via atmospheric channels are one of the real options for solving this problem. However, optical atmospheric communication systems are subject to the influence of atmospheric factors, i.e. due to absorption, scattering, and diffraction, the laser beam is weakened in power and broadens. In well-known works on the divergence of the beam, the question of the dependence of the maximum divergence of the beam on such indicators as the power of the laser source, the geometric length of the channel, and the wavelength of optical radiation is investigated. At the same time, in these works, one channel of the laser communication network is considered, and the question of choosing the diameter of the laser beam in all channels of the laser multichannel atmospheric network is not discussed.

Purpose. The problem of optimal choice of the diameter of a laser beam in a multichannel atmospheric optical communication system of a distributed type is formulated, taking into account the divergence of the beam.

The essence of the proposed solution, unlike well-known works that solve the optimization problem taking into account the influence of beam divergence, the solution to this goal covers both stationary and mobile versions of the network channels of the entire multichannel system. The task was solved by forming a single target functional and further optimizing it in order to identify the optimal relationship between the radius of the beam at the input of the channel receiver and the radius of the beam at the output of the radiator when applied to the atmospheric channel under consideration, taking into account the possibility of its broadening. It is shown that such an expansion of the diameter of the laser beam, according to the optimal law identified, makes it possible to achieve a maximum of the average intensity of the laser beam transmitted through the atmospheric channel to all receivers of the system. The conducted **model studies** of the proposed method for accounting for beam diameter broadening confirmed the possibility of obtaining an optimal ratio between the main indicators of a multichannel laser atmospheric network.

Scientific novelty. A mathematical model for optimizing the choice of beam diameter in a multichannel laser atmospheric network has been developed.

Theoretical and practical significance. Optimization of the proposed model made it possible to obtain an optimal ratio between the main indicators of a multi-channel laser atmospheric network of a distributed type, which can be used in the construction of such systems.

Keywords: optimization, beam divergence, modeling, atmospheric coupling, beam intensity

For citation: Mammadov A.M. Optimal Choice of Laser Beam Diameter in Atmospheric Communication Networks of Distributed and Mobile Type. *Proceedings of Telecommunication Universities*. 2025;11(2):49–55. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-49-55. EDN:YRLSIN

Введение

Развитие сетей связи диктует необходимость создания коммуникационных структур, отличающихся высокими показателями пропускания, без-

опасности, малым потреблением энергии, способностью создания связи со многими устройствами. Примером тому являются сети 5G, обеспечивающие высокую скорость коммуникации по сравнению с

существующими беспроводными сетями связи [1, 2]. Как отмечается в работах [3, 4], в следующем поколении мобильной связи (6G) терагерцовые антенны будут реализованы в виде малогабаритных деталей, и здесь могут быть применены средства открытой атмосферной связи.

Согласно [5–8], интернет вещей является концепцией, позволяющей контролировать ресурсы и осуществлять коммуникацию в реальном времени в различных сферах, таких как бизнес, промышленность, медицина и т. д. Быстрое развитие интернета вещей привело к экспоненциальному росту количества различных интеллектуальных технических средств, присоединяемых к интернету, что, в свою очередь, привело к появлению большого массива данных, подлежащих передаче по системам связи. Для решения данной задачи электромагнитные сети связи обладают недостаточным потенциалом, и одним из реальных вариантов для устранения указанной проблемы являются оптическая система передачи информации по атмосферным каналам [9–10]. Вместе с тем, такие оптические системы связи подвержены влиянию атмосферных факторов, т. к. из-за поглощения, рассеяния и дифракции лазерный луч ослабляется по мощности и подвергается уширению по диаметру [11, 12]. В связи с малой высотой трассы прохождения лазерного луча считается, что турбулентность атмосферы отсутствует. Расходимость луча привела к появлению относительно большого количества работ по оптимизации таких систем [13–20]. В этих работах исследуется вопрос зависимости величины максимальной расходимости луча от таких показателей, как мощность лазерного источника, геометрическая длина канала, длина волны оптической радиации. Вместе с тем, в указанных работах рассматривается один канал лазерной сети связи, а вопрос о выборе диаметра лазерного луча во всех каналах лазерной многоканальной атмосферной сети не обсуждается. В настоящей работе сформулирована и решена задача оптимального выбора диаметра лазерного луча во взаимосвязи с основными параметрами многоканальной сети атмосферной оптической связи, результаты которой могут быть использованы как в стационарных, так и мобильных реализациях сети.

Материалы и методы

Прежде всего рассмотрим существующие модели, используемые для оценки влияния уширения луча на показатели систем оптической атмосферной связи [21]. При создании систем оптической атмосферной связи учет расхождения луча имеет особую важность, т. к. из-за дифракции света оптический луч подвергается уширению, не достигая апертуры приемника. Как результат, появляются потери сигнала из-за расходимости, по причине

превышения диаметра луча размера апертуры приемника. Кроме этих потерь, также существуют потери нацеливания луча из-за смещения последнего относительно оптической оси нацеливания на приемник. Указанные потери диктуют необходимость проведения анализа изменения интенсивности луча по радиальному направлению поперечного сечения луча.

Согласно [22], для характеристики лазерного луча на расстоянии $z = L$ используются такие показатели, как параметр кривизны $\varphi_0(L)$ и отношение Френеля $\Omega_0(L)$, определяемые как:

$$\varphi_0(L) = 1 - \frac{L}{C_0}, \quad (1)$$

$$\Omega_0(L) = \frac{2L}{kR_0^2}, \quad (2)$$

где R_0 – радиус луча на определенной дистанции, вычисляемый как расстояние от центральной точки поперечного сечения до такой радиальной точки, где интенсивность уменьшается до величины $\left(\frac{1}{e}\right) \cdot I_0$, I_0 – максимальная интенсивность луча в центре; C_0 – радиус кривизны; k – волновое число, определяемое с помощью выражения:

$$k = \frac{2\pi}{\lambda}, \quad (3)$$

где λ – длина волны.

При этом уширение оптического луча оценивается как:

$$\theta = \frac{\lambda}{\pi R_B}, \quad (4)$$

где R_B – минимальная величина радиуса луча, определяемая выражением:

$$R_B = \frac{R_0}{\left[\left(\frac{kR_0^2}{2C_0}\right)^2 + 1\right]^{1/2}}. \quad (5)$$

Согласно [21] интенсивность оптического луча на радиальном расстоянии r от оптической оси определяется как:

$$I(r, L) = \frac{I_0 R_0^2}{R^2(L)} \exp\left[-\frac{2r^2}{R^2(L)}\right], \quad (6)$$

где $R(L)$ – радиус луча на дистанции L , определяемый выражением:

$$R(L) = R_0[\varphi^2(L) + \Omega_0^2(L)]^{1/2}. \quad (7)$$

Согласно [21] мощность оптического луча (P_0) и интенсивность оптической волны при $r = 0$ связаны соотношением:

$$I(0, L) = \frac{2P_0}{\pi R^2(L)}. \quad (8)$$

При этом, на практике выполняется условие $\varphi_0 \gg \Omega_0$, вследствие чего уравнение принимает следующий вид [3]:

$$R(L) = R_0 + L\theta. \quad (9)$$

С учетом вышеизложенных исходных материалов задача исследования формируется следующим образом: рассматривается система оптической атмосферной связи, в которой показатель L (расстояние от передатчика до приемника) изменяется в пределах $L_{\min}$ до $L_{\max}$. Это условие может быть реализовано при рассмотрении двух типов систем.

Первый тип системы. Распределенная стационарная система оптической атмосферной связи, где использован один передатчик и n количество стационарных приемников (рисунок 1). В этой структуре L изменяется в пространстве.

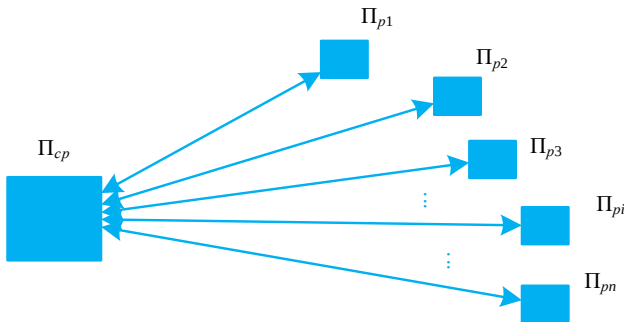


Рис. 1. Структура системы атмосферной связи, где показатель L изменяется в пространстве

Fig. 1. The Structure of the Atmospheric Communication System, Where the Parameter L Varies in Space

Второй тип системы. Система мобильной оптической атмосферной связи, где показатель L изменяется во времени. Например, неподвижный передатчик поддерживает связь с мобильными приемниками, расстояние между которыми изменяется во времени, т. е. $L = L(t)$ (рисунок 2).

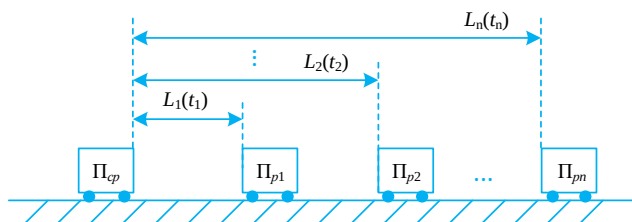


Рис. 2. Структура оптической мобильной атмосферной связи, где приемник, удаляясь от передатчика в моменты $t_1, t_2, \dots, t_n$, занимает позиции $\Pi_{p1}, \Pi_{p2}, \dots, \Pi_{pn}$ на расстояниях $L_1(t_1), L_2(t_2), \dots, L_n(t_n)$

Fig. 2. The Structure of Optical Mobile Atmospheric Communication, Where the Receiver, Moving Away from the Transmitter at Moments $t_1, t_2, \dots, t_n$, Occupies Positions $\Pi_{p1}, \Pi_{p2}, \dots, \Pi_{pn}$ at Distances $L_1(t_1), L_2(t_2), \dots, L_n(t_n)$

При этом приемник оповещает передатчик о текущем месторасположении. Допускается, что за время сеанса связи мобильный приемник проходит

дистанцию, длина которой намного меньше расстояния между передатчиком и приемником, т. е. может быть не учтена. Следует определить такое оптимальное соотношение между показателями r, L, R_0 и θ , при котором интенсивность луча достигла бы максимума в среднем на входах всех приемников.

Для решения поставленной задачи на базе выражения (6) определим функционал цели F :

$$F = \frac{1}{L_{\max}} \int_{L_{\min}}^{L_{\max}} \frac{I_0 R_0^2}{R^2(L)} \exp \left[-\frac{2r^2(L)}{R^2(L)} \right] dL. \quad (10)$$

При этом имеется ввиду, что показатель r может быть изменен в зависимости от L . Вычислим $R(L)$, при котором F достиг бы максимальной величины. Согласно [5] оптимальная функция $R(L)$ при этом удовлетворяет условию:

$$\frac{d \left\{ \frac{I_0 R_0^2}{R^2(L)} \exp \left[-\frac{2r^2(L)}{R^2(L)} \right] \right\}}{dR(L)} = 0. \quad (11)$$

Из условия (11) получаем:

$$-\frac{I_0 R_0^2}{R^2(L)} + \frac{I_0 R_0^2}{R^2(L)} \cdot \frac{2r^2(L)}{R^2(L)} = 0. \quad (12)$$

Проведя сокращения в (12), находим:

$$R(L) = \sqrt{2}r(L). \quad (13)$$

Исходя из (13), функционал (10) достигает максимума, что легко проверить, вычислив показатель Лагранжа (α):

$$\alpha = \frac{d^2 \left\{ \frac{I_0 R_0^2}{R^2(L)} \exp \left[-\frac{2r^2(L)}{R^2(L)} \right] \right\}}{dR(L)^2}, \quad (14)$$

знак которого оказывается отрицательным.

С учетом выражений (9) и (13) получаем:

$$R_0 + L\theta = \sqrt{2}r(L) \quad (15)$$

или

$$r(L) = \frac{R_0 + L\theta}{\sqrt{2}}. \quad (16)$$

Таким образом, при условии (13) целевой функционал (10) достигает максимума. Вычисленная оптимизационная функция (13) совместно с выражением (9) позволяет определить оптимальный закон уширения диаметра луча, при котором средняя величина интенсивности могла бы достичь максимальной величины. В отличие от результатов известных работ по оптимизации систем атмосферной оптической связи с учетом расходимости лазерного луча, полученное выражение позволяет оптимизировать многоканальную систему по фактору уширения луча путем реализации условия (16), которое, в свою очередь, приводит к оптимизации функционала цели F (10).

Модельные исследования

Проведем модельные расчеты на базе формулы (15). Допустим, что исходный радиус лазерного луча равен 1 мм, т. е. $R_{ис} = 1$ мм.

Вычислим R_0 . Исходную интенсивность луча обозначим как $I_{0исх}$. При радиусе луча R_0 его интенсивность уменьшается до величины $\frac{I_{0исх}}{e}$. Интенсивность луча определим по выражению:

$$I_0 = \frac{P}{S}, \quad (17)$$

где P – сила света; S – площадь поперечного сечения луча.

Исходная интенсивность луча находится в следующем соотношении с указанными выше показателями:

$$kI_{0исх} = \frac{P}{S_{исх}}, \quad (18)$$

где k – коэффициент пропорциональности.

При достижении радиуса R_0 справедливо следующее соотношение:

$$\frac{kI_{0max}}{e} = \frac{P}{S(L)}, \quad (19)$$

где $S(L)$ – площадь поперечного сечения луча.

Из (18) и (19) получим следующее соотношение:

$$\frac{P}{S_{исх}} = \frac{eP}{S(L)}. \quad (20)$$

Из (20) находим площадь поперечного сечения луча на расстоянии (L) :

$$S(L) = eS_{исх}. \quad (21)$$

Из геометрии известно, что:

$$S(L) = \pi R_0^2. \quad (22)$$

Сравнивая (21) и (22) получим:

$$\pi R_0^2 = eS_{исх}, \quad (23)$$

или

$$\begin{aligned} R_0 &= \sqrt{\frac{e}{\pi} S_{исх}} = \sqrt{\frac{2,71}{3,14} S_{исх}} = \\ &= \sqrt{\frac{2,71 \cdot 3,14 r^2}{3,14}} = 1,66r. \end{aligned} \quad (24)$$

Для конкретного расчета примем следующие модельные данные: $\theta = 10 \mu rad$; $L = 100$ м. С учетом выражения (16) вычислим оптимальный радиус:

$$r(L) = 1,66 + 100 \cdot 10 \mu rad = 1,66 + 2 \text{ мм} = 3,66 \text{ мм}.$$

Таким образом, согласно вышеизложенному при оптимальном радиусе $r(L) = 3,66$ мм, а также при $L = 100$ мм; $\theta = 10 \mu rad$; $R_{ис} = 1$ мм, по критерию (10) на входы приемников в среднем поступит оптический сигнал максимальной интенсивности.

Обсуждение

Таким образом, поставлена и решена задача оптимизации атмосферной оптической связи многоканального типа, решение которой позволило осуществить оптимальный выбор диаметра луча. Полученное решение может быть распространено как на стационарные, так и мобильные варианты систем. В отличие от известных работ, в которых решается задача оптимизации одного канала связи с учетом влияния уширения луча, решение поставленной задачи, охватывающей всю многоканальную систему по данному фактору, осуществлено путем формирования единого целевого функционала и дальнейшей оптимизации с целью выявления оптимальной взаимосвязи между величиной радиуса луча на входе некоторого фиксированного приемника канала и радиусом луча на выходе излучателя с учетом его возможного уширения до подачи на этот канал.

Заключение

Показано, что уширение диаметра лазерного луча до подачи его в конкретный атмосферный канал по выявленному оптимальному закону позволяет достичь максимума средней величины интенсивности лазерного луча по каналам системы.

Разработан математический аппарат и модель оптимизации многоканальной лазерной атмосферной сети по фактору уширения диаметра луча.

Оптимизация предложенной модели учета уширения диаметра луча позволила получить оптимальное соотношение между основными показателями многоканальной лазерной атмосферной сети.

Список источников

1. Masud M., Gaba G.S., Choudhary K., Alrooba R., Hossain M.S. A Robust and Lightweight Secure Access Scheme for Cloud Based E-Healthcare Services // Peer-to-Peer Networking and Applications. 2021. Vol. 14. Iss. 5. PP. 3043–3057. DOI:10.1007/s12083-021-01162-x. EDN:HOLGNG
2. Zhang Y., Ma X., Zhang J., Hossain M.S. Edge Intelligence in the Cognitive Internet of Things: Improving Sensitivity and Interactivity // IEEE Network. 2019. Vol. 33. Iss. 3. PP. 58–64. DOI:10.1109/MNET.2019.1800344

3. Harada R., Shibata N., Kaneko S., Imai T., Kani J.I., Yoshida T. Adaptive Beam Divergence for Expanding Range of Link Distance in FSO With Moving Nodes Toward 6G // *IEEE Photonics technology letters*. 2022. Vol. 34. Iss. 20. PP. 1061–1064. DOI:10.1109/LPT.2022.3199789
4. Sahoo P.K., Yadav A.K., Prajapati Y.K., Tripathi R. Optimum APD Gain Evaluation of FSO System for Inter-building Laser Communication Application // *Select Proceedings of VCAS: Advances in VLSI, Communication, and Signal Processing. Lecture Notes in Electrical Engineering*. Singapore: Springer, 2020. Vol. 587. PP. 307–314. DOI:10.1007/978-981-32-9775-3_29
5. Ding J., Nemati M., Ranaweera C., Choi J. IoT Connectivity Technologies and Applications: A Survey // *IEEE Access*. 2020. Vol. 8. PP. 67646–73. DOI:10.1109/ACCESS.2020.2985932. EDN:PTUAHG
6. Liu T., Lu D. The application and development of IoT // *Proceedings of the International Symposium on Information Technologies in Medicine and Education (Hokodate, Hokkaido, 03–05 August 2012)*. IEEE, 2012. Vol. 2. PP. 991–994. DOI:10.1109/ITiME.2012.6291468
7. Masud M., Alazab M., Choudhary K., Gaba G.S. 3P-SAKE: Privacy-Preserving and Physically Secured Authenticated Key Establishment Protocol for Wireless Industrial Networks // *Computer Communications*. 2021. Vol. 175. PP. 82–90. DOI:10.1016/j.comcom.2021.04.021. EDN:ODUELB
8. Gaba G.S., Kumar G., Kim T.-H., Monga H., Kumar P. Secure Device-to-Device communications for 5G enabled Internet of Things applications // *Computer Communication*. 2021. Vol. 169. PP. 114–128. DOI:10.1016/j.comcom.2021.01.010. EDN:URLIZU
9. Huang L., Liu S., Dai P., Li M., Chang G.-K., Shi Y. Unified Performance Analysis of Hybrid FSO/RF System with Diversity Combining // *Journal of Lightwave Technology*. 2020. Vol. 38. Iss. 24. PP. 6788–6800. DOI:10.1109/JLT.2020.3018125. EDN:LTGOED
10. Bloom S., Hartley W. The last-mile solution: Hybrid FSO Radio. 2002. URL: <https://www.semanticscholar.org/paper/The-last-mile-solution%3A-Hybrid-FSO-Radio-Bloom-Hartley/d3ff29c18df55d5fb8baf4372adddc2399ceb121> (Accessed 28.03.2025)
11. Guiomar F.P., Lorences-Riesgo A., Ranzal D., Rocco F., Sousa A.N., Fernandes M.A., et al. Adaptive Probabilistic Shaped Modulation for High-Capacity Free-Space Optical Links // *Journal of Lightwave Technology*. 2020. Vol. 38. Iss. 23. PP. 6529–6541. DOI:10.1109/JLT.2020.3012737. EDN:XYPJOZ
12. Rangan S., Rappaport T.S., Erkip E. Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges // *Proceedings of the IEEE*. 2014. Vol. 102. Iss. 3. PP. 366–385. DOI:10.1109/PROC.2014.2299397
13. Zhao Z., Liao R., Zhang Y. Impact of Laser beam diverging angle on free-space optical communications // *Proceedings of the Aerospace Conference (Big Sky, Montana, USA, 05–12 March 2011)*. IEEE, 2011. DOI:10.1109/AERO.2011.5747410
14. Killinger D. Free-Space Optics for Laser Communication Through the Air // *Optics and Photonics News*. 2002. Vol. 13. Iss. 10. PP. 36–42. DOI:10.1364/OPN.13.10.000036
15. Farid A., Hranilovic S. Outage Capacity Optimization for Free-Space Optical Links with Pointing Errors // *Journal of Lightwave Technology*. 2007. Vol. 25. Iss. 7. PP. 1702–1710. DOI:10.1109/JLT.2007.899174
16. Ren Y., Dang A., Luo B., Guo H. Capacities for Long-Distance Free-Space Optics Links Under Beam Wander Effects // *IEEE Photonics Technology Letters*. 2010. Vol. 22. Iss. 14. PP. 1069–1071. DOI:10.1109/LPT.2010.2050468. EDN:ODHBUJ
17. Peppas K.P., Datsikas C.K. Average Symbol Error Probability of General-Order Rectangular Quadrature Amplitude Modulation of Optical Wireless Communication Systems Over Atmospheric Turbulence Channels // *Journal of Optical Communications and Networking*. 2010. Vol. 2. Iss. 2. PP. 102–110. DOI:10.1364/JOCN.2.000102
18. Arnon S. Effects of atmospheric turbulence and building sway on optical wireless communication systems // *Optics Letters*. 2003. Vol. 28. Iss. 2. PP. 129–131. DOI:10.1364/OL.28.000129
19. García-Zambrana A., Castillo-Vázquez B., Castillo-Vázquez C. Asymptotic error-rate analysis of FSO links using transmit laser selection over gamma-gamma atmospheric turbulence channels with pointing errors // *Optics Express*. 2012. Vol. 20. Iss. 3. PP. 2096–2109. DOI:10.1364/OE.20.002096.
20. Mitsev T., Kolev N., Ivanov H., Dimitrov K. Optimum Divergence of the Transmitter Optical Radiation in FSO Systems // *Proceedings of the 6th International Conference of Environmental Science and Technology (ICEST 2012, Veliko Tarnovo, Bulgaria, 28–30 June 2012)*. 2012.
21. Singh H., Miglani R., Mittal N., Gaba G., Masud M., Aljhdali S. Design and analysis of commercially viable free-space optical communication link for diverse beam divergence profiles // *Frontiers in Physics*. 2021. Vol. 9. DOI:10.3389/fphy.2021.778734. EDN:UHUYWI
22. Andrews L.C., Phillips R.L. *Laser Beam Propagation through Random Media*. Bellingham, Washington: SPIE, 2005. 783 p. DOI:10.1117/3.626196. EDN:YCTRZM

References

1. Masud M., Gaba G.S., Choudhary K., Alroobaea R., Hossain M.S. A Robust and Lightweight Secure Access Scheme for Cloud Based E-Healthcare Services. *Peer-to-Peer Networking and Applications*. 2021;14(5):3043–3057. DOI:10.1007/s12083-021-01162-x. EDN:HOLGNG
2. Zhang Y., Ma X., Zhang J., Hossain M.S. Edge Intelligence in the Cognitive Internet of Things: Improving Sensitivity and Interactivity. *IEEE Network*. 2019;33(3):58–64. DOI:10.1109/MNET.2019.1800344
3. Harada R., Shibata N., Kaneko S., Imai T., Kani J.I., Yoshida T. Adaptive Beam Divergence for Expanding Range of Link Distance in FSO With Moving Nodes Toward 6G. *IEEE Photonics technology letters*. 2022;34(20):1061–1064. DOI:10.1109/LPT.2022.3199789
4. Sahoo P.K., Yadav A.K., Prajapati Y.K., Tripathi R. Optimum APD Gain Evaluation of FSO System for Inter-building Laser Communication Application. *Select Proceedings of VCAS: Advances in VLSI, Communication, and Signal Processing. Lecture Notes in Electrical Engineering*, vol.587. Singapore: Springer; 2020. p.307–314. DOI:10.1007/978-981-32-9775-3_29

5. Ding J., Nemati M., Ranaweera C., Choi J. IoT Connectivity Technologies and Applications: A Survey. *IEEE Access*. 2020;8: 67646–73. DOI:10.1109/ACCESS.2020.2985932. EDN:PTUAHG
6. Liu T., Lu D. The application and development of IoT. *Proceedings of the International Symposium on Information Technologies in Medicine and Education, Hokodate, Hokkaido, 03–05 August 2012, vol.2*. IEEE; 2012. p.991–994. DOI:10.1109/ITIME.2012.6291468
7. Masud M., Alazab M., Choudhary K., Gaba G.S. 3P-SAKE: Privacy-Preserving and Physically Secured Authenticated Key Establishment Protocol for Wireless Industrial Networks. *Computer Communications*. 2021;175:82–90. DOI:10.1016/j.comcom.2021.04.021. EDN:ODUELB
8. Gaba G.S., Kumar G., Kim T.-H., Monga H., Kumar P. Secure Device-to-Device communications for 5G enabled Internet of Things applications. *Computer Communication*. 2021;169:114–128. DOI:10.1016/j.comcom.2021.01.010. EDN:URLIZU
9. Huang L., Liu S., Dai P., Li M., Chang G.-K., Shi Y. Unified Performance Analysis of Hybrid FSO/RF System with Diversity Combining. *Journal of Lightwave Technology*. 2020;38(24):6788–6800. DOI:10.1109/JLT.2020.3018125. EDN:LTGOED
10. Bloom S., Hartley W. *The last-mile solution: Hybrid FSO Radio*. 2002. URL: <https://www.semanticscholar.org/paper/The-last-mile-solution%3A-Hybrid-FSO-Radio-Bloom-Hartley/d3ff29c18df55d5fb8baf4372adddc2399ceb121> [Accessed 28.03.2025]
11. Guiomar F.P., Lorences-Riesgo A., Ranzal D., Rocco F., Sousa A.N., Fernandes M.A., et al. Adaptive Probabilistic Shaped Modulation for High-Capacity Free-Space Optical Links. *Journal of Lightwave Technology*. 2020;38(23):6529–6541. DOI:10.1109/JLT.2020.3012737. EDN:XYPJOZ
12. Rangan S., Rappaport T.S., Erkip E. Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges. *Proceedings of the IEEE*. 2014;102(3):366–385. DOI:10.1109/JPROC.2014.2299397
13. Zhao Z., Liao R., Zhang Y. Impact of Laser beam diverging angle on free-space optical communications. *Proceedings of the Aerospace Conference, 05–12 March 2011, Big Sky, Montana, USA*. IEEE; 2011. DOI:10.1109/AERO.2011.5747410
14. Killinger D. Free-Space Optics for Laser Communication Through the Air. *Optics and Photonics News*. 2002;13(10): 36–42. DOI:10.1364/OPN.13.10.000036
15. Farid A., Hranilovic S. Outage Capacity Optimization for Free-Space Optical Links with Pointing Errors. *Journal of Lightwave Technology*. 2007;25(7):1702–1710. DOI:10.1109/JLT.2007.899174
16. Ren Y., Dang A., Luo B., Guo H. Capacities for Long-Distance Free-Space Optics Links Under Beam Wander Effects. *IEEE Photonics Technology Letters*. 2010;22(14):1069–1071. DOI:10.1109/LPT.2010.2050468. EDN:ODHBUJ
17. Peppas K.P., Datsikas C.K. Average Symbol Error Probability of General-Order Rectangular Quadrature Amplitude Modulation of Optical Wireless Communication Systems Over Atmospheric Turbulence Channels. *Journal of Optical Communications and Networking*. 2010;2(2):102–110. DOI:10.1364/JOCN.2.000102
18. Arnon S. Effects of atmospheric turbulence and building sway on optical wireless communication systems. *Optics Letters*. 2003;28(2):129–131. DOI:10.1364/OL.28.000129
19. García-Zambrana A., Castillo-Vázquez B., Castillo-Vázquez C. Asymptotic error-rate analysis of FSO links using transmit laser selection over gamma-gamma atmospheric turbulence channels with pointing errors. *Optics Express*. 2012;20(3): 2096–2109. DOI:10.1364/OE.20.002096.
20. Mitsev T., Kolev N., Ivanov H., Dimitrov K. Optimum Divergence of the Transmitter Optical Radiation in FSO Systems. *Proceedings of the 6th International Conference of Environmental Science and Technology, ICEST 2012, 28–30 June 2012, Veliko Tarnovo, Bulgaria*. 2012.
21. Singh H., Miglani R., Mittal N., Gaba G., Masud M., Aljahdali S. Design and analysis of commercially viable free-space optical communication link for diverse beam divergence profiles. *Frontiers in Physics*. 2021;9. DOI:10.3389/fphy.2021.778734. EDN:UHUYWI
22. Andrews L.C., Phillips R.L. *Laser Beam Propagation through Random Media*. Bellingham, Washington: SPIE; 2005. 783 p. DOI:10.1117/3.626196. EDN:YCTRZM

Статья поступила в редакцию 27.01.2025; одобрена после рецензирования 25.02.2025; принята к публикации 01.03.2025.

The article was submitted 27.01.2025; approved after reviewing 25.02.2025; accepted for publication 01.03.2025.

Информация об авторе:

МАММАДОВ
Афлатун Масим

преподаватель кафедры цифровой коммуникаций и прикладной информатики
Азербайджанского государственного экономического университета

 <https://orcid.org/0009-0000-5355-0493>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.

Научная статья

УДК 621.396.6

<https://doi.org/10.31854/1813-324X-2025-11-2-56-66>

EDN:KVNRSR



Разработка расчетной модели радиотрассы систем дальнемагистральной связи декаметрового диапазона для анализа параметров ионосферного канала

Алексей Игоревич Рыбаков ✉, rybakov.ai@sut.ru

Григорий Алексеевич Фокин, fokin.ga@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. В системах дальнемагистральной связи декаметрового (ДКМ) диапазона на количество битовых ошибок одновременно влияют два ключевых фактора: соотношение сигнал / шум и степень фазовых искажений, вызванных доплеровским сдвигом, возникающим из-за случайного движения неоднородностей ионосферы. Проблема повышения помехоустойчивости таких систем осложняется тем, что даже при высоком уровне сигнала на входе демодулятора прием может затрудняться фазовыми искажениями, которые приводят к резкому увеличению количества битовых ошибок и ухудшению коэффициента BER. Несмотря на достаточное число классических работ, проблема повышения помехоустойчивости современных отечественных систем ДКМ радиосвязи в заданных сценариях функционирования с использованием современных методов и средств цифровой обработки сигналов остается актуальной и востребованной.

Объектом исследования являются современные отечественные системы ДКМ радиосвязи, которые зачастую демонстрируют известные недостатки, включая низкую адаптивность к изменениям в ионосфере и проблемы с интерференцией сигналов. Например, система Р-016 имеет ограничения по диапазону частот, что делает ее менее эффективной в условиях варьирования ионосферных характеристик, которые влияют на уровень сигналов. Прототипы также могут иметь проблемы с обработкой сигналов, что приводит к возникновению битовых ошибок до 10^{-3} даже при отсутствии заметных помех.

Предметом исследования являются модели и методы функционирования радиотрассы ДКМ радиосвязи.

Задачей исследования является оценка влияния различных факторов, таких как изменение длины преамбулы и использование адаптивных фильтров, на помехоустойчивость системы. Анализ полученных **результатов** показывает, что увеличение длины преамбулы в таких системах способствует повышению помехоустойчивости дальнемагистральной связи. **Научная новизна** заключается в усовершенствовании существующих расчетных моделей радиотрасс в ДКМ диапазоне путем применения комплекса параметров, включающего задаваемое для данного сеанса связи отношение сигнал / шум в радиотрассе и для повышения точности расчета напряженности поля в точке приема, пересчитанные значения критических частот по прогнозам концентрации электронов, а также доплеровский сдвиг для каждого слоя ионосферы.

Практическая значимость результата заключается в повышении помехоустойчивости существующих систем ДКМ радиосвязи в ионосферных условиях распространения.

Ключевые слова: дальнемагистральная система связи, отношение сигнал / шум, помехоустойчивость, декаметровая радиосвязь, эквалайзер, преамбула, адаптивные фильтры

Ссылка для цитирования: Рыбаков А.И., Фокин Г.А. Разработка расчетной модели радиотрассы систем дальнемагистральной связи декаметрового диапазона для анализа параметров ионосферного канала // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 56–66. DOI:10.31854/1813-324X-2025-11-2-56-66. EDN:KVNRSR

Original research

<https://doi.org/10.31854/1813-324X-2025-11-2-56-66>

EDN:KVNRSR

Development of a Radio Wave Propagation Model for Long-Haul Decameter-Band Communication Systems in the Decameter Range for the Analysis of Ionospheric Channel Parameters

Алексей Игоревич Рыбаков ✉, rybakov.ai@sut.ru

Григорий Алексеевич Фокин, fokin.ga@sut.ru

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance. In long-haul decameter-band (LDB) communication systems, the number of bit errors is simultaneously influenced by two key factors: the signal-to-noise ratio and the degree of phase distortions caused by the Doppler shift, which arises from the random movement of ionospheric inhomogeneities. The challenge of improving the interference resistance of such systems is complicated by the fact that even with a high signal level at the demodulator's input, reception can be hindered by phase distortions, leading to a sharp increase in the number of bit errors and a degradation of the BER (bit error rate) coefficient. Despite the sufficient number of classical works, the problem of enhancing the interference resistance of modern domestic LDB radio communication systems under specified operating scenarios using contemporary methods and digital signal processing tools remains relevant and in demand, which lead to a sharp increase in the number of bit errors and a deterioration in the BER coefficient. Despite a sufficient number of classical works, the problem of improving the noise immunity of modern domestic LDB or short-wave (HF) radio communication systems in specified operating scenarios using modern methods and means of digital signal processing (DSP) remains relevant and in demand.

The object of the research is modern domestic decameter-band radio communication systems, which often exhibit known shortcomings, including low adaptability to changes in the ionosphere and issues with signal interference. For example, the R-016 system has limitations in its frequency range, making it less effective under varying ionospheric conditions that affect signal levels. Prototypes may also have problems with signal processing, which leads to bit errors of up to 10^{-3} even in the absence of noticeable interference.

The subject of the study is the models and methods of functioning of LDB decameter-band communication lines.

The objective of the research is to evaluate the impact of various factors, such as changes in preamble length and the use of adaptive filters, on the interference resistance of the system. The analysis of the obtained **results** shows that increasing the preamble length in such systems contributes to enhancing the interference resistance of long-haul communication. **The scientific novelty** lies in the improvement of existing radio propagation models in the decameter range by applying a complex of parameters that includes the signal-to-noise ratio set for a specific communication session in the radio line, and to increase the accuracy of field strength calculations at the receiving point, recalculated values of critical frequencies based on electron concentration forecasts, as well as Doppler shift for each layer of the ionosphere. **The practical significance** of the results lies in the enhancement of the interference resistance of existing decameter-band communication systems in ionospheric propagation conditions.

Keywords: long-range communication system, signal-to-noise ratio, interference immunity, decameter radio communication, equalizer, preamble, adaptive filters

For citation: Rybakov A.I., Fokin G.A. Development of a Radio Wave Propagation Model for Long-Haul Decameter-Band Communication Systems in the Decameter Range for the Analysis of Ionospheric Channel Parameters. *Proceedings of Telecommunication Universities*. 2025;11(2):56–66. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-56-66. EDN:KVNRSR

1. Введение

Радиосвязь в декаметровом (ДКМ) диапазоне квалифицируется как высокочастотная (ВЧ) коротковолновая (КВ) связь [1]. На протяжении многих лет существования ДКМ радиосвязи считалось, что новые виды связи в конечном итоге ее заменят [2]. Порой действительно автоматизированные сети связи строятся на основе высокоэффективных кабельных и радиорелейных линий. Проблемы связи в малонаселенных, удаленных или труднопроходимых регионах могут быть решены с помощью спутниковых систем связи. Однако вопрос о полном вытеснении ДКМ линий радиосвязи в ближайшем будущем практически не поднимается ни в одной стране мира. Напротив, несмотря на бурное развитие микроволновых средств дальней связи, включая спутниковые системы, в последние годы возрастает интерес к технической реконструкции ДКМ линий радиосвязи [2]. Это основано на оценке ее технического потенциала, экономической эффективности и стратегической роли, как важного резерва. В условиях катастрофической ионизации атмосферы ДКМ линии радиосвязи, вероятно, испытывают такие же трудности, как и другие радиотехнические системы, но при этом они адаптируются и восстанавливаются значительно быстрее [3]. Однако надежность и помехоустойчивость ДКМ линий радиосвязи может быть недостаточно высокой из-за эффектов замираний сигнала при ионосферном распространении и наличия «молчащих зон». Это может привести к существенным потерям информации, если команды управления для крупной региональной системы передаются исключительно по ДКМ линиям радиосвязи. Радиосвязь в ДКМ диапазоне, наряду со спутниковой, продолжает оставаться одним из видов межконтинентальной связи, представляя собой экономичный способ организации дальней радиосвязи для относительно небольшого числа каналов [2, 3]. Это отличает их от систем связи в микроволновых диапазонах, которые, напротив, становятся эффективными только при одновременной передаче нескольких сотен–тысяч телефонных каналов.

Для обеспечения надежной передачи информации на обширных территориях используют различные системы и технические средства связи, которые в совокупности обеспечивают высокую степень надежности информационных коммуникаций в нашей стране [4]. Радиосвязь в ДКМ диапазоне (длина волны 10–100 м) играет ключевую роль как для внутренней, так и для международной связи, включая зональную, подвижную и производственно-диспетчерскую связь общего и ведомственного назначения. Она предоставляет следующие услуги: магистральная связь, радиовещание, служебные линии для наземных станций спутниковой связи, авиационная связь «земля–воздух», мор-

ская связь «берег–судно», связь между судами в морском флоте, радиосвязь Гидрометеослужбы и другие [3]. Следует отметить, что радиосвязь в ДКМ диапазоне часто называют КВ из-за общепринятого обозначения волн длиной от 10 до 100 м как коротких волн.

В исследуемой дальнемагистральной системе связи на количество битовых ошибок одновременно влияют два ключевых фактора: соотношение сигнал / шум и степень фазовых искажений, вызванных доплеровским сдвигом, возникающим из-за случайного движения неоднородностей ионосферы [4]. Это приводит к тому, что даже при высоком уровне сигнала на входе демодулятора прием может затрудняться из-за фазовых искажений, что негативно сказывается на качестве связи [5]. Объектом настоящего исследования являются современные отечественные системы ДКМ радиосвязи. Примерами систем прототипов являются: ПИРС, Р-016 [3], Сердолик и Mesa [6], каждая из которых имеет свои недостатки.

ПИРС (поточковая измерительная радиосистема) обладает ограниченной помехоустойчивостью, что приводит к высокому уровню битовых ошибок в условиях изменчивой ионосферы, особенно при увеличении расстояний передачи. Также ПИРС характеризуется недостатком гибкости в настройках частот.

Система Р-016 [3] имеет низкую устойчивость к фазовым искажениям, возникающим из-за доплеровского сдвига, что ограничивает ее эффективность в динамических условиях. Кроме того, она не обеспечивает необходимую разностную обработку сигналов, что также влияет на качество связи.

Система Сердолик [6] демонстрирует высокую чувствительность к радиочастотным шумам, имея уровень битовых ошибок, достигающий 10^{-4} без адекватной фильтрации сигналов. Также следует отметить, что Сердолик не способен эффективно адаптироваться к изменениям в условиях приема.

Система Mesa [6] часто сталкивается с проблемой «интерференции» от соседних каналов, что ухудшает качество связи. На практике это может привести к потере целых пакетов данных, что также увеличивает количество битовых ошибок.

Предметом исследования выступают модели и методы функционирования радиолиний ДКМ радиосвязи. Их исследование позволяет оценить влияния различных факторов, таких как изменение длины преамбулы и использование адаптивных фильтров, на помехоустойчивость системы.

Для достижения цели повышения помехоустойчивости систем дальнемагистральной связи ДКМ диапазона формализуем далее следующие частные задачи исследования:

1) разработка расчетной модели радиотрассы систем дальнемагистральной связи ДКМ диапазона для определения рекомендуемых параметров работы модели ионосферного канала;

2) разработка эквалайзера и конфигурационных файлов его настройки для обеспечения доставки данных за сеанс радиосвязи в ДКМ радиолинии с заданной помехоустойчивостью с использованием программной реализации средствами имитационного моделирования в среде MatLab для модели рэлеевского канала;

3) экспериментальная апробация комплексной модели ДКМ радиолинии, включающей расчетную модель радиотрассы, эквалайзер с конфигурационными файлами настройки и полевые измерения профиля ДКМ радиолинии, для научно обоснованного выбора параметров передачи на основе отношения сигнал / шум с учетом различных длин преамбул.

Комплексной научной задачей исследования является разработка, программная реализация и экспериментальная апробация эквалайзера с конфигурационными файлами настройки, учитывающая расчетную модель радиотрассы и полевые измерения профиля ДКМ радиолинии.

Материал настоящего исследования организован далее следующим образом. В разделе 2 приводится анализ состояния проблемы. Раздел 3 посвящен разработке расчетной модели радиотрассы в ДКМ диапазоне. Разработка эквалайзера и конфигурационных файлов его настройки для обеспечения доставки данных за сеанс радиосвязи представлена в разделе 4. Результаты экспериментальной апробации комплексной модели ДКМ радиолинии описаны в разделе 5. Выводы и направления дальнейших исследований приводятся в заключении.

2. Анализ состояния проблемы

Известно, что ДКМ радиосвязь подвержена влиянию ионосферных возмущений и насыщена сигналами помеховых станций, поэтому к ней предъявляются высокие требования к надежности, электромагнитной совместимости и эффективному использованию частотного спектра [5, 7]. В материалах исследования Brandon T. Hunt, David B. Naab, Thomas Cameron Sego и Tom V. Holschuh [8, 9] показана возможность доработки известной модели К. Ватерсона для применения технологии многоантенного приема в условиях многолучевого распространения, специфичных для ДКМ радиоканала. Современные исследования коллектива авторов Zhaoji Zhang, Zhu Jin, Ying Li, Guanghui Song и Yizhuo Wang [9, 10] ориентированы на решение вопросов расширения спектра сигнала с помощью множественных несущих и применения метода Уолша для повышения помехоустойчивости и защищенности в радиоканале.

Научные исследования отечественных ученых, посвященные улучшению качества передачи в ДКМ радиолиниях, проводятся в различных технических вузах. Особого внимания заслуживают исследования А.А. Сидоренко и Д.А. Мартышевой [6, 11, 12] из Владимирского государственного университета имени А.Г. и Н.Г. Столетовых, которые посвящены проблеме подавления межсимвольной интерференции через применение новых алгоритмов и увеличению эффективности использования помехоустойчивых турбокодов. Кроме того, в контексте актуальности темы исследования рассматриваются экспериментальные работы по эквалайзерам на основе нейронных сетей с обучением в многолучевом радиоканале, выполненные на физическом факультете МГУ им. М.В. Ломоносова Д.Р. Валиулиным и П.Н. Захаровым [13].

Современные ДКМ линии радиосвязи сталкиваются с рядом технических недостатков, которые требуют комплексного решения и адаптации параметров в зависимости от конкретных условий. К основным проблемам можно отнести неопределенность состояния ионосферы, что приводит к значительным колебаниям ее характеристик в пределах диапазона частот от 10 до 100 МГц, а также сложности с точным прогнозированием этих характеристик на временной промежуток, соответствующий предстоящему сеансу радиосвязи, который может составлять от нескольких часов до суток [12]. Переменяющиеся неоднородности концентрации электронов в ионосфере, которые могут изменяться на уровне от 10^9 до 10^{12} электронов на кубический метр, а также колебания их уровня в различных слоях ионосферы, приводят к изменению высоты переотражения электромагнитной волны. Обычно высота переотражения варьируется от 30 до 1000 км и определяется условиями радиосигнала. Такие изменения вызывают флуктуации мощности огибающей сигнала, которые могут варьироваться от 1 до 10 дБ, и доплеровские смещения на приемной стороне, достигающие десятков герц, что негативно сказывается на качестве связи [14].

Для минимизации указанных недостатков ДКМ радиолиний целесообразно эффективно использовать методы цифровой обработки сигналов при выборе вариантов адаптации параметров передачи. Применение современных алгоритмов, таких как адаптивные фильтры и методы пространственной обработки, могут значительно улучшить стабильность и помехоустойчивость систем связи к динамическим условиям ионосферы. Например, использование эквалайзеров, способных компенсировать флуктуации в диапазоне от 10 до 20 дБ [15], позволяет значительно повысить устойчивость связи и качество передаваемых данных. Это, в свою очередь, ведет к значительному улучшению показателей помехоустойчивости и общей надежности

радиосистем, что становится особенно важным для успешного функционирования радиосвязи на больших расстояниях. Учитывая изложенные выше проблемы, в настоящей работе предполагается провести уточнение существующей методики функционирования ДКМ линий радиосвязи, акцентируя внимание на разработке эквалайзера с конфигурационными файлами, адаптированными к высоким флуктуациям и условиям работы в изменчивой ионосфере. Это позволит не только улуч-

шить надежность передачи данных, но и обеспечить более высокий уровень устойчивости системы к внешним воздействиям, что является ключевым для эффективного использования радиосвязи в данном диапазоне. На рисунке 1 представлена структурная схема типового адаптивного комплекса радиосвязи, которая иллюстрирует компоненты и процессы, позволяющие повысить помехоустойчивость системы, тем самым повысить качество связи в условиях изменения параметров ионосферы.

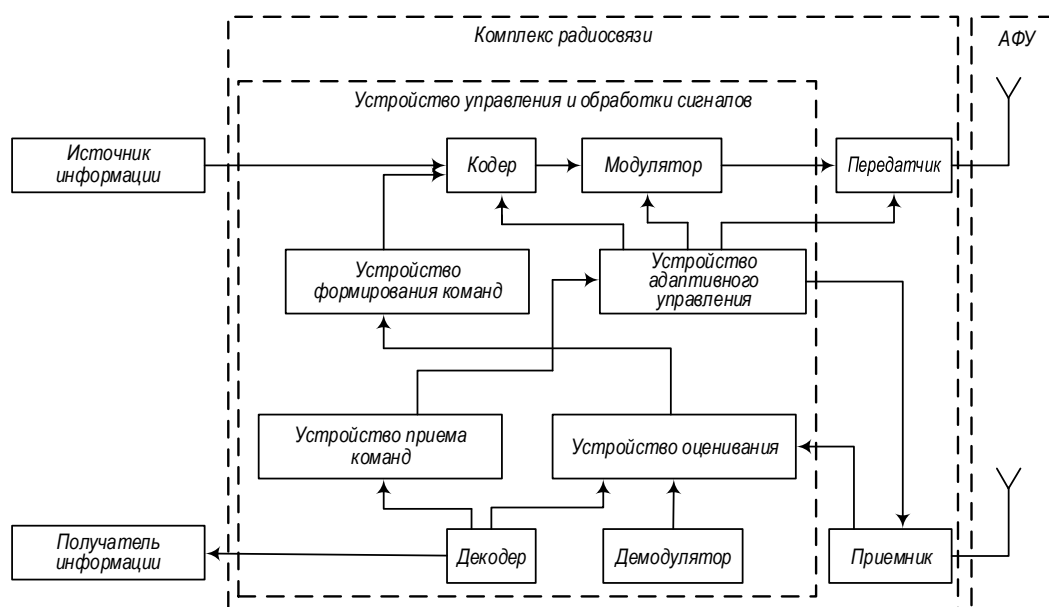


Рис. 1. Структурная схема адаптивного комплекса радиосвязи

Fig. 1. Structural Diagram of an Adaptive Radio Communication System

Для реализации схемы адаптивного комплекса радиосвязи (см. рисунок 1) в настоящее время используются современные персональные электронно-вычислительные машины, работающие под управлением адаптивных алгоритмов [15, 16]. Подобная реализация открывает новые возможности для более точного и быстрого анализа состояния ДКМ радиолиний, позволяя автоматизировать алгоритмы на основе методов конечных элементов и численные методы решения дифференциальных уравнений.

Данный подход особенно актуален в условиях динамической и изменчивой ионосферы, где необходимо учитывать большое количество переменных и факторов, способных повлиять на уровень электромагнитного поля (ЭМП). Однако он характеризуется следующими недостатками. Во-первых, не учитываются факторы, связанные с распространением радиоволн по радиотрассе, а также проблемы суперпозиции сигналов в точке приема [15, 16]. Во-вторых, не принимаются во внимание энергетические характеристики радиолиний, включая

диаграммы направленности антенн, уровень необходимой мощности и текущее отношение сигнал / шум в динамически изменяющейся ДКМ радиолинии. С учетом указанных ранее недостатков, в данной работе для повышения помехоустойчивости систем дальнемагистральной связи в ДКМ диапазоне решается частная научная задача по разработке расчетной модели радиотрассы. Эта модель предназначена для определения рекомендуемых параметров работы ионосферного канала, что будет способствовать адаптации работы систем дальнемагистральной связи в условиях данного диапазона.

3. Разработка расчетной модели ДКМ радиолинии

3.1. Постановка задачи

Обычно при работе с ДКМ радиолинией заранее задаются ключевые энергетические параметры радиоаппаратуры, такие как мощность передатчика – $P_{\text{прд}}$, коэффициент усиления (КУ) антенн, и др., что позволяет в итоге разрабатывать технические решения по настройке сигнально-кодовой конструк-

ции (СКК) [12]. Далее приводится расчетная модель радиотрассы систем дальнемагистральной связи ДКМ диапазона, учитывающая параметры конфигурационных файлов и настроек эквалайзера.

Модель К. Ватерсона [5] предназначена для расчета уровней ЭМП в зависимости от условий распространения радиоволн, учитывая факторы, такие как рельеф местности, атмосфера и время суток. Данная модель широко используется для оценки дальности связи и качества сигнала ДКМ радиолиний, так как позволяет исследовать различные сценарии, делит условия функционирования радиотрасс на различные категории (например, плодородные и сложные зоны), что делает ее полезной для практического применения.

Модель А.Н. Казанцева уточняет более специфические параметры, таких как углы падения и отражения радиоволн [4, 5]. Данная модель используется для оценки уровня ЭМП на конкретных маршрутах связи, адаптируясь к местным условиям, что делает ее особенно полезной в сложных ландшафтах.

Объединение моделей А.Н. Казанцева и К. Ватерсона представляется научно обоснованным подходом для уточнения текущих значений уровней ЭМП в точках приема в зависимости от комплекса факторов. Доводами в пользу данного подхода являются следующие обстоятельства. Во-первых, объединение позволяет создавать многоуровневые модели, которые смогут учитывать как статистические, так и детализированные факторы, что особенно важно в условиях, где географические и погодные аспекты играют значительную роль. Во-вторых, совмещение моделей способствует разработке алгоритмов, адаптивных к сценариям конкретных регионов, позволяя более точно планировать размещение антенн и маршруты связи. Для успешного использования моделей их необходимо объединить с учетом калибровки и последующей апробации на экспериментально измеренных профилях ДКМ радиолиний [2].

Разрабатываемая расчетная модель дает возможность определить мощность сигнала на входе приемника в зависимости от параметров передающего и приемного оборудования, основываясь на заданных параметрах СКК (скорость передачи, полосу частот, требуемое отношение сигнал / шум). Расчетная модель, представляемая далее, разработана для того, чтобы определить интервал по времени, который позволяет проводить сеансы связи и выполнять испытания по проверке применяемого адаптационного метода. В результате проведения расчета требуемая мощность передатчика для обеспечения уровня битовых ошибок 3×10^{-3} , определена на уровне 10 Вт.

Данные значения обусловлены двумя факторами:

1) значение уровня битовых ошибок в 3×10^{-3} указывает на допустимый уровень ошибок в системе, который считается приемлемым для ряда радиокommunikационных приложений [17]; в разрабатываемой далее модели планируется обеспечить баланс между качеством передачи и мощностью передатчика;

2) уровень 10 Вт выбран на основе анализа параметров ионосферного канала, включая влияние шумов, затухание сигналов и другие факторы, способные ухудшить качество связи [18]; при этой мощности удастся на практике достичь наилучшей в заданных условиях помехоустойчивости (мощность может быть снижена до 5 Вт; такое снижение приводит к уменьшению уровня принимаемого сигнала на входе демодулятора на 3 дБ).

Рассматриваемая далее модель представляет собой обобщенную вариацию известных моделей расчета А.Н. Казанцева и К. Ватерсона [4, 5] с использованием данных с ионозонда. В разработанной модели рассчитаны точки отражения от ионосферы и от поверхности Земли. Далее проводится сопоставление полученных точек с прогнозом по статистическим данным модели IRI (*аббр. от англ. International Reference Ionosphere*). Также точки отражения могут сравниваться с данными, полученными с аппаратуры ионозондов о критических частотах в точках отражения. Для протяженной радиотрассы точки отражения находятся в разных частях планеты. Из-за физической особенности среды распространения в нестационарном канале связи Солнце по-разному ионизирует слои ионосферы. Максимальное возможное количество скачков составляет не более шести, т. к. при максимальном их количестве волна обогнет планету полностью. В дальнейшем осуществляется формализация расчетной модели каналов (ДКМ) радиолинии, которая будет включать в себя детальную математическую формулировку процессов передачи информации через радиоканал, учет влияния различных факторов, таких как помехи и искажения, а также применение соответствующих алгоритмов обработки сигналов для повышения помехоустойчивости системы.

3.2. Формализация

Распространение радиоволн в атмосфере сопровождается потерей части электромагнитной энергии при прохождении через слои атмосферы D и E , а также для ДКМ волн, отражающихся от слоя F_2 . Положительная рефракция может вернуть волну обратно на Землю.

Поясним операции вычисления требуемых характеристик сигнала на обобщенной модели рас-

чета параметров ДКМ радиолнии [1, 14]. Предположим, что значение диэлектрической проницаемости $\varepsilon = 1$, при $q_n \gg 90^\circ$. По принятым условиям запишем формулу расчета угла падения на нижнюю границу ионосферы, при котором волна вновь направляется в сторону земной поверхности [14]:

$$\sin \varphi = \sqrt{e_{n(homp)_1}}, \quad (1)$$

где φ – угол входа луча в ионосферу; $e_{n(homp)_1}$ – относительная диэлектрическая проницаемость слоя ионосферы, где произошло отражение.

Условие отражения можно представить следующим образом [16]:

$$\sin \varphi = 1 - \sqrt{81 \frac{N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right) N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right)}{f^2 (\text{кГц})}}. \quad (2)$$

В формуле (2) концентрация электронов $N\varepsilon$ соответствует высоте отражения $h_{\text{отр}} = h_0 + z_{\text{отр}}$. Из условия отражения (2) следует, что волна с частотой f будет отражена на высоте $h_{\text{отр}}$ в ионосфере со следующей характеристикой:

$$f = \frac{9 \sqrt{N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right)}}{\cos \varphi}. \quad (3)$$

В случае нормального падения волны $\cos \varphi = 1$ отражение описывается следующим образом:

$$f = 9 \sqrt{N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right)} = f_0. \quad (4)$$

С увеличением частоты отражение происходит на все больших высотах, заканчиваясь на высоте с максимальной концентрацией свободных зарядов. Наивысшая частота, при которой волна еще отражается от ионосферы при нормальном падении, называется критической и определяется выражением:

$$f_{\text{кр}} = 9 \sqrt{N\varepsilon_{\text{max}}} = f_{0\text{max}}. \quad (5)$$

На частотах $f > f_{\text{кр}}$ диэлектрическая проницаемость ионосферы будет положительной $\varepsilon > 0$ и ионосфера будет радиопрозрачной. Критические частоты относятся к ДКМ диапазону $f_{\text{кр}} = 3 \div 30$ МГц. При известных средних значениях $N\varepsilon_{\text{max}}$ и полученных радиопрогнозов модели IRI: в дневное время критическая частота соответствует диапазону $f_{\text{кр}} = 11 \div 13$ МГц, в ночное – $f_{\text{кр}} = 4 \div 6$ МГц. В средних широтах Земли в дневное время критическая частота равна $f_{\text{кр}} = 6 \div 8$ МГц, а в ночное – $f_{\text{кр}} = 2,5 \div 4,5$ МГц. Исключением являются районы Крайнего Севера в период полярного сияния, поскольку D -слой, благодаря своей высокой плотности, поглощает сигнал вместо его отражения. Плот-

ность нейтральных, неионизированных частиц, которые составляют основную массу в этом регионе, в D -слое в 1000 раз выше, чем в E - и F -слоях.

В случае, когда отражение волны происходит при наклонном падении на ионосферу ($\varphi > 0$), частота радиоволны определяется по закону секанса [1, 14]:

$$f = \frac{9 \sqrt{N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right)}}{\cos \varphi} = f_0 \sec \varphi. \quad (6)$$

На больших расстояниях между точкой излучения и точкой приема справедливо следующее выражение для рабочей частоты [15]:

$$f = K f_0 \sec \varphi_0, \quad (7)$$

где при расстояниях свыше 500 км необходимо учитывать сферичность Земли, и для этого вводят поправочный коэффициент K .

Из формул (2 и 7) получим выражение для поправочного коэффициента:

$$K = \sqrt{\frac{1 + \frac{2z_0}{(R_3 + h_0) \cos \varphi_0}}{\frac{z_0}{(R_3 + h_0)}}} \sec \varphi_0, \quad (8)$$

где z_0 – расстояние от нижней границы ионосферы до высоты с максимальной концентрацией свободных зарядов; с увеличением угла падения волна проникает глубже в ионосферу.

При заданном угле падения с ростом рабочей частоты f отражение происходит на все больших высотах. Наибольшая частота, на которой волна еще испытывает отражение при данном угле падения, называется максимальной.

С увеличением максимальной рабочей частоты отражение происходит на все больших высотах и определяется выражением [16]:

$$f = \frac{9 \sqrt{N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right)}}{\cos \varphi} = K \sqrt{N\varepsilon \left(\frac{\text{эл}}{\text{см}^3}\right)} \sec \varphi. \quad (9)$$

Из-за кривизны земной поверхности и удаленности нижней границы ионосферы от поверхности Земли (60 км) угол падения на ионосферу не может достигать 90° [17]. Максимальный угол падения (10) соответствует траектории, проходящей по касательной к поверхности Земли [18].

$$\sin \varphi_{\text{max}} = \frac{R_3}{R_3 + h_0}, \quad (10)$$

где h_0 – нижняя граница ионосферы.

Наиболее высокая частота при максимальном угле входа в атмосферу, которая еще может отражаться в ионосфере, называется предельной [3]:

$$f_{k_{\max}} N\varepsilon = \frac{9\sqrt{N\varepsilon \left(\frac{\varepsilon_l}{\text{см}^3}\right)}}{\cos \varphi}, f_{k_{\max}} = 4f_{\text{кр}}. \quad (11)$$

Максимальный угол падения волны не может превышать $81,5^\circ$, а частота $f_{k_{\max}}$ не может превышать 30 МГц. В ионосфере отражаются волны длиннее 10 м, т. е. те, частота которых меньше 30 МГц. Соответственно, если частота несущей выше 30 МГц, радиоволна выходит за пределы ионосферы. В реальной ионосфере, например, при появлении спорадических слоев E_s отражение может наступить при гораздо больших частотах, в том числе и метрового диапазона. На границе раздела сред должны выполняться условия равенства тангенциальных составляющих векторов E и H (при отсутствии поверхностных токов в этих слоях), на основании чего для вертикально поляризованной волны можно составить систему двух уравнений [3, 17]:

$$E_{m_{\text{пад}}} \cos \varphi - E_{m_{\text{отр}}} \cos \varphi = E_{m_{\text{пр}}} \cos \psi, \quad (12)$$

$$H_{m_{\text{пад}}} + H_{m_{\text{отр}}} = H_{m_{\text{пр}}}. \quad (13)$$

Коэффициент отражения волны R определяется как отношение амплитуд $E_{m_{\text{пад}}}/E_{m_{\text{отр}}}$.

Пусть свойства Земли близки к идеальному диэлектрику. Тогда из (13) следует [16]:

$$E_{m_{\text{пр}}} = \frac{E_{m_{\text{пад}}} + E_{m_{\text{отр}}}}{\sqrt{\varepsilon'_k}}. \quad (14)$$

Подставим формулу (14) в выражение (12) и поделим результат на $E_{m_{\text{пад}}}$. Далее, перейдя к углу скольжения $\theta = \frac{\pi}{2} - \varphi$, с помощью (15) исключим угол ψ так, что справедливо выражение:

$$\cos \psi = \sqrt{1 - \sin^2 \psi} = \sqrt{\frac{\varepsilon'_k - \cos^2 \theta}{\varepsilon'_k}}, \quad (15)$$

из которого можно записать формулу коэффициента отражения:

$$R_B = \frac{\varepsilon'_k \sin \theta - \sqrt{\varepsilon'_k - \cos^2 \theta}}{\varepsilon'_k \sin \theta + \sqrt{\varepsilon'_k - \cos^2 \theta}}. \quad (16)$$

Если проводимость почвы $\sigma \neq 0$, то ε'_k является комплексной величиной, из чего следует комплексный характер членов выражения (16), поэтому R_B можно представить в виде [19]:

$$R_B = |R_B|e^{-i\beta_B}, \quad (17)$$

откуда следует, что при взаимодействии радиоволны с проводящей поверхностью Земли появляется сдвиг фаз между падающей и отраженной волнами на угол β_B . Данный вывод позволяет далее

сделать заключение о том, что коэффициент отражения можно представить в виде вещественной величины:

$$R_B = \frac{\varepsilon' \sin \theta - \sqrt{\varepsilon' - \cos^2 \theta}}{\varepsilon' \sin \theta + \sqrt{\varepsilon' - \cos^2 \theta}} R_B = |R_B|e^{-i\beta_B}, \quad (18)$$

Расчет напряженности ЭМП E_c ионосферной волны на ДКМ радиолинии при помощи расчетной методики позволяет проводить оценку уровней ЭМП в регионе ДКМ радиолиний для средних широт. При этом используются типовые данные радиопрогнозов и характеристики подстилающей поверхности (модель IRI), а также заданные энергетические характеристики передатчика и параметры диаграмм направленности антенн. В основу расчетов положены формулы (16, 18) [14, 16].

Напряженность поля в точке приема определяется выражением:

$$E = \frac{\sqrt{30P_1 G_1}}{r_l} \frac{1}{2} \frac{1 + |\dot{R}|}{2} |\dot{R}|^{(n-1)(n-1)} e^{-\Gamma}, \quad (19)$$

где P_1 – мощность передатчика, подводимого к антенне; G_1 – коэффициент усиления антенны; r_l – дальность распространения, равная длине луча ионосферной волны для n -скачковой радиотрассы; множители ослабления ионосферной волны определяются модулем коэффициента отражения от Земли $|\dot{R}|$ через полный коэффициент поглощения в ионосфере Γ .

Множитель $1/2$ в формуле (19) характеризует уменьшение мощности, переносимой волной, на 6 дБ. Общий интегральный коэффициент поглощения $e^{-\Gamma}$ представляет собой сумму частных интегральных коэффициентов поглощения в слоях D, E, F_1 , через которые проходит волна, и в слое F_2 , от которого волна отражается. Расчет полного коэффициента поглощения Γ произведен по методу А.Н. Казанцева. При отражении от слоя E полный коэффициент поглощения Γ на частоте излучения передатчика f определяется по критическим частотам слоя E с использованием выражения [14, 19]:

$$\Gamma = \frac{3(f_0^E)^2 \sec \varphi_0^E}{(f + f_L)^2} + 4 \frac{\cos^2 \varphi_0^E}{f_0^E} f \sqrt{\frac{f}{(f + f_L)}}. \quad (20)$$

Для совершенствования типовой методики расчета уровня напряженности на территории распространения ДКМ радиотрассы была проведена оценка уровня ЭМП от передатчика с антенной, обладающей заданными характеристиками диаграммы направленности. Расчеты на n -скачковой радиотрассе выполняются не в одной точке, а на всей территории региона, и результаты представляются графически в двумерной системе координат. За счет многократных отражений в ионосфере и последующих отражений от земной поверхности

дальняя связь может достигать расстояний в тысячи (даже десятки тысяч) километров при сравнительно небольших мощностях передатчика. Измерения осуществляются с помощью специализированных станций высотного зондирования. На основе полученных данных составляются графики суточного изменения максимально применимых частот (МПЧ), которые рассылаются по различным местам. Оптимальная частота выбирается равной уровню $f = 0,85 \text{ МПЧ}$. Представленный на рисунке 2 график изменения МПЧ в зависимости от времени суток получен по данным модели IRI. Численно полученный коэффициент перед f_0 определяет высоту максимума ионизации слоя F_2 в предположении параболического распределения ионизации вблизи максимума и без учета запаздывания волны в нижележащих слоях ионосферы. Верхняя линия демонстрирует значения МПЧ (значения f_i), взятые из модели IRI, а нижняя линия показывает оптимальные значения МПЧ (значения f_{opt}).

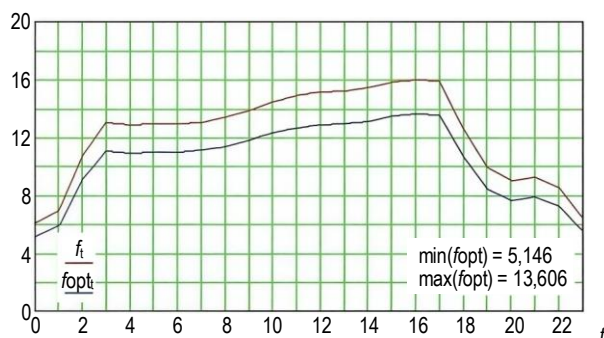


Рис. 2. График изменения МПЧ

Fig. 2. Schedule of Changes in the Maximum Applicable Frequencies

При действующей высоте слоя F_2 основные рабочие частоты, полученные по модели IRI, оцениваются по формуле:

$$f = 0,85f_0. \quad (21)$$

Пересчет МПЧ при работе с ДКМ радиолинией позволяет определить рабочие диапазоны частот для передачи сигналов, учитывая факторы, влияющие на распространение радиоволн [19, 20]. Этот процесс требует учета множества факторов, которые могут существенно влиять на распространение радиоволн, таких как условия ионосферы, время суток, климатические изменения.

Основной задачей расчета является выбор таких частот, которые будут отражаться и пропускаться через многослойную ионосферу в заданных условиях. Это позволяет обеспечить максимальную помехоустойчивость радиосвязи, поскольку правильно выбранные частоты обеспечивают наилучшее взаимодействие с различными слоями ионосферы. Кроме того, правильный расчет применимых частот позволяет установить максимальную

дальность. Это критично для обеспечения стабильной связи, поскольку выбранные частоты должны минимизировать потери сигнала на больших расстояниях, что в конечном итоге повышает помехоустойчивость ДКМ радиолиний. Анализ частот также помогает избежать потенциальных проблем с интерференцией, возникающей из-за передачи на тех же частотах, что и другие радиосигналы. Наконец, расчет максимально применимых частот предоставляет практическое руководство для настройки передающего и приемного оборудования с целью повышения помехоустойчивости ДКМ радиолиний. Это включает в себя использование антенн и усилителей, которые могут быть адаптированы для работы с ДКМ радиолинией.

3.3. Перспективы использования расчетной модели ДКМ радиолинии

Представленная расчетная модель ДКМ радиолинии решает задачу пересчета МПЧ для научно обоснованного выбора подходящих диапазонов передачи сигналов в текущий момент времени. Данная расчетная модель используется, чтобы решить комплексную задачу повышения помехоустойчивости систем дальнемагистральной связи ДКМ радиолиний.

При разработке и конфигурации эквалайзера предложенная в настоящей работе расчетная модель ДКМ радиолинии решает задачу научно-обоснованного выбора частот, которые в данном сценарии эффективно отражаются и пропускаются ионосферой. Обоснованный подбор таких частот позволяет решить задачу достижения максимальной дальности связи в декаметровом диапазоне. Конфигурационные файлы эквалайзера, которые в свою очередь также определяются диапазоном пересчитанной частоты, при этом решают задачу обеспечения максимальной помехоустойчивости. Имитационное моделирование в MatLab позволяет проверить различные конфигурации эквалайзера в условиях рэлеевского канала, и, таким образом, научно обосновать его настройку.

Последующая экспериментальная апробация комплексной модели ДКМ радиолинии также учитывает предложенную в настоящей работе расчетную модель ДКМ радиолинии и конфигурацию эквалайзера с различными длинами преамбул. Полученные средствами имитационного моделирования значения отношения сигнал / шум на входе приемника учитывают модель ДКМ радиолинии и определяют вариативный порядок выбора длины преамбулы. Экспериментальная апробация выбранных длин преамбул в функционирующей системе ДКМ радиолинии для заданного сценария подтверждает практическую применимость предложенного подхода.

Закключение

Предложенная расчетная модель для систем дальнемагистральной радиосвязи в декаметровом диапазоне актуальна и востребована в условиях изменяющейся ионосферы, требующей учета множества переменных и факторов. В рамках будущих исследований будет продемонстрирован алгоритм применения данной модели для настройки эквалайзера на приемной стороне ДКМ радиолинии. Модель будет использоваться для комплексного решения задачи повышения помехоустойчивости

систем дальнемагистральной радиосвязи. При разработке и конфигурации эквалайзера данная модель позволит определить оптимальные частоты, которые в данном сценарии будут эффективно отражаться и пропускаться ионосферой. Обоснованный выбор таких частот способствует максимизации дальности связи в выбранном диапазоне. Конфигурационные файлы эквалайзера, зависящие от диапазона выбранной частоты, будут направлены на обеспечение максимальной помехоустойчивости системы дальнемагистральной радиосвязи.

Список источников

1. Скляр Б. Цифровая связь. М.: Издательский дом Вильямс, 2003. 1104 с.
2. Качнов А.И., Пенкин А.А., Рыбаков А.И. Разработка мобильной системы информационного обеспечения с использованием каналов метеорной связи // V Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфокоммуникаций в науке и образовании» (Санкт-Петербург, Российская Федерация, 10–11 марта 2016 г.). СПб.: СПбГУТ, 2016. С. 177–181. EDN:WZIKIL
3. Воробьев О.В., Рыбаков А.И. Вариант реализации двунаправленной связи в системе метеорной связи. Описание программно-аппаратного комплекса СМС // VI Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфокоммуникаций в науке и образовании» (Санкт-Петербург, Российская Федерация, 01–02 марта 2017 г.). СПб.: СПбГУТ, 2017. С. 128–133. EDN:ZECEUT
4. Черный Ф.Б. Распространение радиоволн. М.: Сов. Радио, 1972. 464 с.
5. Грудинская Г.П. Распространение радиоволн. М.: Высшая школа, 1975, 280 с.
6. Мартышевская Д.А., Полушин П.А. Моделирование сверточного метода обработки сигналов при межсимвольной интерференции // XXI Международная научная конференция студентов и молодых ученых «Современные техника и технологии» (СТТ-2015, Томск, Российская Федерация, 5–9 октября 2015 г.). Томск: ТПУ, 2015. С. 43–45.
7. Попов Н.А., Пятаков А.И. Расчет территориального распределения уровней сигнала на многоскачковых радиотрассах от КВ передатчика средней мощности с антенной “Ромб” // Автоматизация процессов управления. 2009. № 15. С. 65–73. EDN:KHPEFN
8. Нарышкин Е.М., Серков В.П. Волновая служба и антенные устройства. Ч. 1. Теория электромагнитного поля и распространение радиоволн. М.: Воениздат, 1982. 288 с.
9. Hunt B.T., Naab D.B., Sego T.C., Holschuh T.V., Moradi H., Farhang-Boroujeny B. Hunt Examining the Performance of MIL-STD-188-110D Waveform 0 Against FBMC-SS Over Skywave HF Channels // IEEE Transactions on Vehicular Technology. 2022. Vol. 71. Iss. 11. PP. 11637–11649. DOI:10.1109/TVT.2022.3189762
10. Zhang Z., Jin Z., Li Y., Song G., Wang Y. Multi-Stage Receiver of MIL-STD-188-110D Waveform 0 for High-Frequency Communication // IEEE Communications Letters. 2024. Vol. 28. Iss. 5. PP. 1166–1170. DOI:10.1109/LCOMM.2024.3375601
11. Самойлов А.Г., Сидоренко А.А. Применения кодов РС в каскаде с двоичными кодами с целью повышения эффективности борьбы с независимыми ошибками // Проектирование и технология электронных средств. 2014. № 3. С. 2–7.
12. Сидоренко А.А. Адаптивное помехоустойчивое кодирование // Материалы X международной научно-технической конференции «Перспективные технологии в средствах передачи информации». Владимир, Суздаль, 2013. Т. 1. С. 152–154.
13. Валиуллин Д.Р., Захаров П.Н. Экспериментальные исследования эквалайзера на основе нейронных сетей с обучением в многолучевом радиоканале // Журнал радиоэлектроники. 2017. № 12. С. 4. EDN:YNSPDZ
14. Шеннон К. Работы по теории информации и кибернетики. М.: Изд-во иностранной литературы, 1963. 829 с.
15. Зюко А.Г., Фалько А.И., Панфилов И.П. и др. Помехоустойчивость и эффективность систем передачи информации. М.: Радио и связь, 1985. 271 с.
16. Зюко А.Г., Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи. М.: Радио и связь, 1999. 432 с.
17. Блейхут Р.Э. Теория и практика кодов, контролирующих ошибки. Пер. с англ. М.: Мир, 1986. 576 с.
18. Кларк Дж., Кейн Дж. Кодирование с исправлением ошибок в системах цифровой связи. Пер. с англ. М.: Мир, 1987.
19. Прокис Дж. Цифровая связь. Пер. с англ. М.: Радио и связь, 2000.
20. Феер К. Беспроводная цифровая связь. Методы модуляции и расширения спектра. Пер. с англ. М.: Радио и связь, 2001. 518 с.

References

1. Sklyar B. *Digital communication*. Moscow: Villeyams Publishing House, 2003. 1104 p. (in Russ.)
2. Kachnov A., Penkin A.A., Rybakov A. The Development of Mobile Information Provision Systems Using Meteor Communication Channels. *Proceedings of the Vth International Scientific, Technical and Scientific-Methodological Conference at Actual Problems of Information Communications in Science and Education, 10–11 March 2016, St. Petersburg, Russian Federation*. St. Petersburg: The Bonch-Bruевич Saint Petersburg State University of Telecommunications Publ.; 2016. p.177–181. (in Russ.) EDN:WZIKIL

3. Vorobyov O., Rybakov A. Option Enable Bidirectional Communication In SMS (System Meteor Communication). Description Software-Hardware Complex SMS. *Proceedings of the VIth International Scientific, Technical and Scientific-Methodological Conference at Actual Problems of Information Communications in Science and Education, 01–02 March 2017, St. Petersburg, Russian Federation*. St. Petersburg: The Bonch-Bruевич Saint Petersburg State University of Telecommunications Publ.; 2017. p.128–133. (in Russ.) EDN:ZECEUT
4. Cherny F.B. *Propagation of Radio Waves*. Moscow: Sov. Radio Publ.; 1972. 464 p. (in Russ.)
5. Grudinskaya G.P. *Propagation of Radio Waves*. Moscow: Vysshaya shkola Publ.; 1975. 280 p. (in Russ.)
6. Martyshevskaya D.A., Polushin P.A. Modeling of Convolutional Signal Processing Method with Intersymbol Interference. *Proceedings of the Conference of Students and Young Scientists at Modern Engineering and Technology, 5–9 October 2015, Tomsk, Russian Federation*. Tomsk: Tomsk Polytechnic University Publ.; 2015. p. 43–45. (in Russ.)
7. Popov N.A., Pyatakov A.I. Calculation of the Territorial Distribution of Signal Levels on Multi-Jump Radio Paths from a Medium-Power HF Transmitter with a Rhomb Antenna. *Automation of Control Processes*. 2009;15:65–73. (in Russ.) EDN:KHPEFN
8. Naryshkin E.M., Serkov V.P. *Wave Service and Antenna Devices. Part 1. Theory of the Electromagnetic Field and Propagation of Radio Waves*. Moscow: Voenizdat Publ.; 1982. 288 p. (in Russ.)
9. Hunt B.T., Haab D.B., Sego T.C., Holschuh T.V., Moradi H., Farhang-Boroujeny B. Hunt Examining the Performance of MIL-STD-188-110D Waveform 0 Against FBMC-SS Over Skywave HF Channels. *IEEE Transactions on Vehicular Technology*. 2022;71(11):11637–11649. DOI:10.1109/TVT.2022.3189762
10. Zhang Z., Jin Z., Li Y., Song G., Wang Y. Multi-Stage Receiver of MIL-STD-188-110D Waveform 0 for High-Frequency Communication. *IEEE Communications Letters*. 2024;28(5):1166–1170. DOI:10.1109/LCOMM.2024.3375601
11. Samoilov A.G., Sidorenko A.A. Application of RS codes in a cascade with binary codes in order to increase the effectiveness of combating independent errors. *Design and Technology of Electronic Means*. 2014;(3):2–7. (in Russ.)
12. Sidorenko A.A. Adaptive noise-resistant coding. *Proceedings of the Xth International Scientific and Technical Conference "Promising Technologies in the Media of Information Transmission", vol.1*. Vladimir, Suzdal; 2013. p.152–154. (in Russ.)
13. Valiullin D.R., Zakharov P.N. Experimental studies of an equalizer based on neural networks with training in a multipath radio channel. *Journal of Radio Electronics*. 2017;12:4. (in Russ.) EDN:YNSPDZ
14. Shannon K. *Works on Information Theory and Cybernetics*. Moscow: Izdatelstvo inostranoi literatury Publ.; 1963. 829 p. (in Russ.)
15. Ziuko A.G., Falko A.I., Panfilov I.P., et al. *Noise Immunity and Efficiency of Information Transmission Systems*. Moscow: Radio i svyaz Publ.; 1985. 271 p. (in Russ.)
16. Ziuko A.G., Klovsky D.D., Korzhik V.I., Nazarov M.V. *Theory of Electrical Communication*. Moscow: Radio i svyaz Publ.; 1999. 432 p. (in Russ.)
17. Bleihut R.E. *Theory and Practice of Error Controlling Codes*. Addison-Wesley Publishing Company; 1983. 500 p.
18. Clark-Jr. G.C., Cain J.B. *Error-Correcting Coding in Digital Communication Systems*. New York and London: Plenum Press; 1981.
19. Prokis J. *Digital Communications*. McGraw-Hill Science; 2000.
20. Feer K. *Wireless Digital Communication. Methods of Modulation and Spectrum Expansion*. Prentice Hall; 1995.

Статья поступила в редакцию 01.03.2025; одобрена после рецензирования 14.03.2025; принята к публикации 03.04.2025.

The article was submitted 01.03.2025; approved after reviewing 14.03.2025; accepted for publication 03.04.2025.

Информация об авторах:

РЫБАКОВ
Алексей Игоревич

старший преподаватель кафедры беспроводных технологий и систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0003-0147-4462>

ФОКИН
Григорий Алексеевич

доктор технических наук, доцент, профессор кафедры радиосвязи и вещания Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-5358-1895>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

Обзорная статья

УДК 004.75

<https://doi.org/10.31854/1813-324X-2025-11-2-67-82>

EDN:QCIUMV



Перспективная архитектура сетей, определяемых знаниями (KDN)

Филип Сергеевич Блан, blan.fs@sut.ru

Василий Сергеевич Елагин ✉, v.elagin@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

В данной статье рассматривается концепция и архитектура сетей, определяемых знаниями – новой парадигмы управления сетями, которая интегрирует искусственный интеллект и машинное обучение для обеспечения интеллектуального и адаптивного поведения сети.

Актуальность исследования обусловлена ограничениями традиционных и программно-определяемых сетей в условиях современных вызовов, таких как экспоненциальный рост трафика, динамичные условия и увеличение операционных затрат. Рассматриваемые в исследовании сети вводят плоскость знаний, что позволяет оптимизировать распределение ресурсов, автоматизировать принятие решений и повышать безопасность в режиме реального времени. Несмотря на то, что сегодня большой популярностью пользуется технология SDN (Software-Defined Network), в которой централизованная функция управления позволяет обзирать все процессы, происходящие в сети. В свое время ее появление действительно оказалось прорывом, и сейчас некоторые эксперты склоняются к тому, что следующим этапом эволюции сетей станет Knowledge-Defined Network – сеть, определяемая знаниями, действующая на основе алгоритмов машинного обучения. Маршрутизация, распределение ресурсов, виртуализация сетевых функций (Network Functions Virtualization, NFV), цепочка сервисных функций (Service Function Chaining, SFC), обнаружение аномалий, анализ загруженности сети – все эти пункты способна взять на себя KDN.

Цель исследования заключается в изучении структурных и функциональных особенностей сетей, определяемых знаниями, а также – в анализе взаимодействия пяти логических плоскостей: данных, управления, мониторинга, знаний и приложений – для достижения высокой степени автоматизации и адаптации.

Методы включают анализ научной литературы, концептуальное моделирование и сравнительную оценку архитектур определяемой знаниями сети и программно-определяемой сети.

Результаты. В ходе исследования была проанализирована архитектура сетей, определяемых знаниями, и определено, что интеграция плоскости знаний в сеть позволяет добиться значительного повышения автоматизации и адаптивности.

Новизна. Проведенное исследование является одной из первых попыток провести системный анализ концепции сетей, определяемых знаниями, в контексте русскоязычной научной литературы. Работа восполняет существующий пробел в отечественной науке, предлагая уникальный взгляд на возможности сетей, определяемых знаниями, с учетом специфики локальных условий и применения

Теоретическая значимость работы заключается в создании основы для изучения и интеграции методов машинного обучения в системы управления сетями.

Ключевые слова: сети, определяемые знаниями, программно-определяемые сети, автоматизация управления, машинное обучение, интеллектуальное управление

Ссылка для цитирования: Блан Ф.С., Елагин В.С. Перспективная архитектура сетей, определяемых знаниями (KDN) // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 67–82. DOI:10.31854/1813-324X-2025-11-2-67-82. EDN:QCIUMV

Review research

<https://doi.org/10.31854/1813-324X-2025-11-2-67-82>

EDN:QCIUMV

Future Architecture of Knowledge-Defined Networking (KDN)

✉ Filip S. Blan, blan.fs@sut.ru

✉ Vasily S. Elagin, v.elagin@sut.ru

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

In this paper, the concept and architecture of Knowledge-Defined Networking (KDN) are explored as a new paradigm of network management that integrates artificial intelligence and machine learning to enable intelligent and adaptive network behavior.

The relevance of the research is driven by the limitations of traditional and Software-Defined Networking (SDN) systems in the face of modern challenges such as exponential traffic growth, dynamic conditions, and rising operational costs. KDN introduces a knowledge plane that optimizes resource allocation, automates decision-making, and enhances security in real-time. Despite the fact that today the SDN (Software-Defined Network) technology is very popular, in which the centralized control function allows to review all processes occurring in the network. At the time, its appearance really turned out to be a breakthrough, and now some experts are inclined to believe that the next stage of network evolution will be the Knowledge-Defined Network - a network defined by knowledge, operating on the basis of machine learning algorithms. Routing, resource allocation, network function virtualization (NFV), service function chaining (Service Function Chaining, SFC), anomaly detection, network load analysis - all these points can be taken on by KDN. **The study aims** to examine the structural and functional features of KDN and analyze the interaction of its five logical planes – data, control, monitoring, knowledge, and applications – to achieve a high degree of automation and adaptability. The research **methods** include literature analysis, conceptual modeling, and a comparative evaluation of KDN and SDN architectures.

The results. The study analyzed the architecture of KDN, comprising five logical planes: data, control, monitoring, knowledge, and applications. The findings demonstrate that integrating the knowledge plane significantly enhances automation and adaptability within the network.

The novelty of this work lies in being one of the first attempts to conduct a systematic analysis of the Knowledge-Defined Networking (KDN) concept in the context of Russian-language scientific literature. The research addresses an existing gap in domestic science, offering a unique perspective on KDN capabilities considering local conditions and applications.

The theoretical significance of the work lies in establishing a foundation for the study and integration of machine learning methods into network management systems.

Keywords: Knowledge-Defined Networking, Software-Defined Networking, management automation, machine learning, intelligent management

For citation: Blan F.S., Elagin V.S. Future Architecture of Knowledge-Defined Networking (KDN). *Proceedings of Telecommunication Universities*. 2025;11(2):67–82. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-67-82. EDN:QCIUMV

Введение

Развитие сетевых технологий прошло несколько ключевых этапов, начиная с традиционных фиксированных архитектур, где управление сетью требовало ручной настройки каждого устройства, до по-

луавтоматических систем, предоставляющих централизованное управление. Первые системы отличались низкой гибкостью и высокой сложностью обслуживания, что делало их малоэффективными для масштабных и динамичных сценариев. С появлением концепции программно-определяемых се-

тей (SDN, аббр. от англ. Software-Defined Networking) в начале 2000-х гг. произошло отделение плоскости управления от плоскости данных, что позволило повысить уровень автоматизации и упростить эксплуатацию [1]. Однако в современных реалиях и SDN оказываются ограничены в своей способности адаптироваться к быстро меняющимся условиям и резкому росту трафика, поэтому встал вопрос о переходе от традиционных архитектур к более интеллектуальным и автоматизированным системам [2].

В настоящее время значительную часть сетевой нагрузки составляют запросы на передачу медиа-файлов (аудио-, фото-, видео-) в высоком разрешении, что приводит к формированию огромных объемов передаваемых и хранимых данных. Недостаточные возможности автоматизации в существующих решениях, особенно в динамичных и высоконагруженных средах, приводят к росту операционных затрат, снижению эффективности обслуживания сети и увеличению рисков эксплуатации.

В условиях возросших требований к скорости, точности и гибкости обработки данных, традиционные сети с фиксированной конфигурацией сталкиваются с серьезными ограничениями, поэтому концепция сетей, определяемых знаниями (KDN, аббр. от англ. Knowledge-Defined Networking), становится одним из наиболее перспективных направлений эволюции сетевых технологий. Возможности машинного (ML, аббр. от англ. Machine Learning) и глубокого обучения (DL, аббр. от англ. Deep Learning) в интеллектуальном слое, известном как плоскость знаний, позволяют эффективно влиять на все части системы [3]. Эта плоскость, являющаяся операционным центром сети, обрабатывает данные для автономного принятия решений и адаптации работы в режиме реального времени.

Цель данной статьи – предоставить обзор сетей, определяемых знаниями, рассмотрев их архитектуру, потенциальные преимущества, недостатки и направления дальнейшего развития. В первую очередь, анализируются архитектурные особенности KDN и их ключевые достоинства, включая автоматизацию и адаптивность, которые снижают затраты, нагрузку на сетевых администраторов и повышают эффективность эксплуатации сети. Также обсуждаются ограничения, связанные с необходимостью значительных вычислительных ресурсов и проблемами интеграции с существующими сетевыми инфраструктурами. Наконец, в статье представлены перспективы внедрения KDN в различные сетевые сценарии.

Обзор сетей, определяемых знаниями

Главным отличием KDN является их способность использовать знания для автоматизации и интеллектуального управления сетью, что делает такие сети более гибкими и адаптивными по сравнению

с традиционными подходами. Если же в традиционных сетях управление основывается на жестко заданных правилах, определяемых администраторами, в KDN подход смещается от информационно-ориентированного к знание-ориентированному. Знание в контексте KDN представляет собой обработанную с помощью машинного обучения информацию, полученную на основе данных, собираемых из сети. Знания представляют собой более высокий уровень абстракции, который формируется через анализ данных, накопленный опыт и обучение [4].

Архитектура KDN организована таким образом, чтобы на каждом уровне, от сбора данных до их анализа и применения, обеспечивать непрерывный цикл получения, обработки и использования знаний. Особенно важна интеграция знаний в плоскости мониторинга и контроля, которые позволяют KDN адаптироваться к изменяющимся условиям и автономно оптимизировать свою работу.

Архитектура KDN включает в себя пять логических плоскостей (рисунок 1), расширяя типичную структуру SDN.

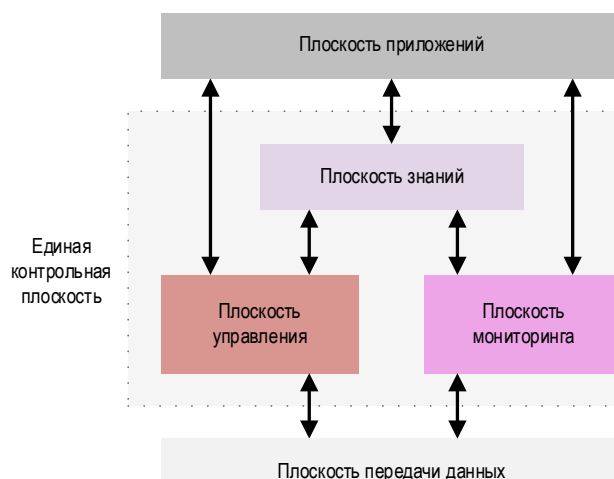


Рис. 1. Схематическая блок-диаграмма архитектуры KDN

Fig. 1. Schematic Block Diagram of the KDN Architecture

Плоскость передачи данных (Data Plane) отвечает за обработку и передачу данных, следуя правилам, заданным *плоскостью управления* (Control Plane). Здесь осуществляется управление маршрутизацией и настройкой сетевых устройств, передавая правила и инструкции в *плоскость передачи данных*. *Плоскость мониторинга* (Management Plane, встречается так же Measurement Plane) обеспечивает сбор данных, контроль состояния сети и конфигурацию сетевых устройств. *Плоскость знаний* (Knowledge Plane) генерирует, обрабатывает и предоставляет знания на основе анализа данных для интеллектуального управления сетью. Далее *плоскость приложений* (Application Plane) предоставляет платформу для взаимодействия сетевых приложений с инфраструктурой и определения высокоуровневых политик.

Следует отметить, что плоскости управления, мониторинга и знаний могут быть функционально интегрированы в единую контрольную плоскость. Такое объединение позволяет проводить параллели с архитектурой SDN, где логически централизованный уровень выполняет схожие задачи координации, наблюдения и регулирования состояния сети [5].

Архитектура KDN отличается высокой степенью согласованности взаимодействия всех пяти плоскостей, что обеспечивает способность системы мгновенно адаптироваться к изменениям внешних и внутренних условий. Это позволяет достичь значительного уровня автоматизации и внедрения интеллектуальных механизмов управления.

Далее каждая из составляющих плоскостей рассмотрена подробно.

Плоскость передачи данных

В KDN плоскость передачи данных, обеспечивая передачу информации между конечными устройствами, выполняет функции, аналогичные традиционным сетям и SDN. Однако существует и несколько ключевых отличий. Прежде всего, настройка и отслеживание состояния плоскости передачи данных в KDN выполняются плоскостью мониторинга, тогда как реализация политик и пра-

вил находится под управлением плоскости контроля. Из-за этого в KDN, по сравнению с SDN, значительно больший объем данных передается в единую контрольную плоскость [6].

Эта увеличенная нагрузка требует от устройств маршрутизации в KDN (физических или виртуальных коммутаторов, маршрутизаторов, точек доступа и т. д.) больше пропускной способности и вычислительной мощности, нежели чем от схожих устройств в SDN. Базовая структура такого устройства, а также взаимодействия плоскости данных с другими элементами системы, представлены на рисунке 2.

Следует отметить, что устройства маршрутизации в идеальных KDN не обладают автономным интеллектом для принятия решений, так как все правила формируются и задаются логически централизованным контроллером. Несмотря на то, что данный подход представляется оптимальным с теоретической точки зрения, его реализация в легаси-системах (унаследованных системах) с устаревшей инфраструктурой вызывает значительные трудности.

Гибридная система, которая упрощает интеграцию KDN в существующие системы, должна содержать дополнительный локальный агент управления. Функциональная структура такого узла в KDN показана на рисунке 3.

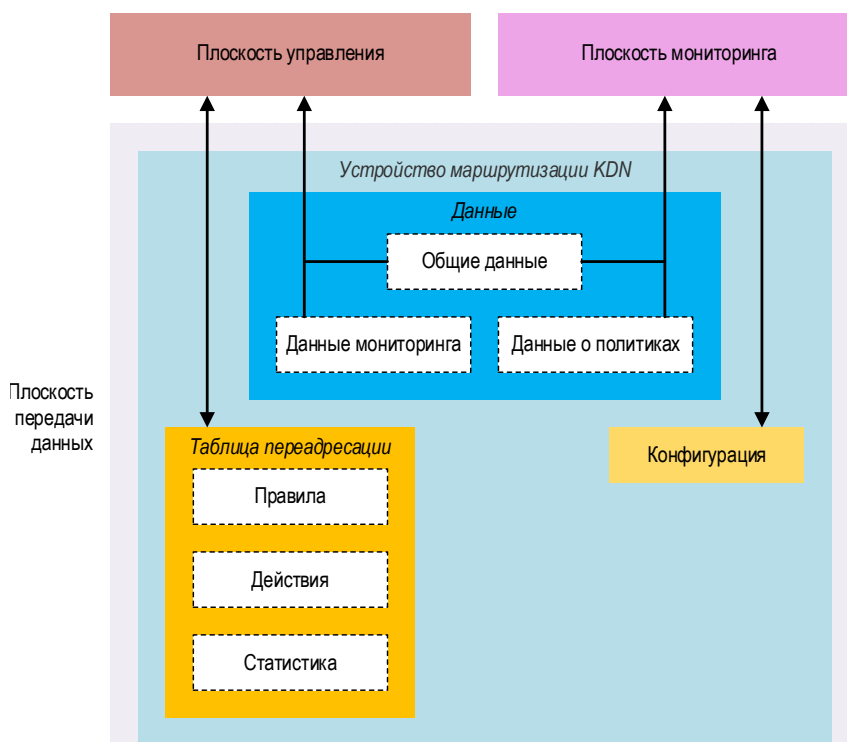


Рис. 2. Схематическая блок-диаграмма плоскости передачи данных и устройства маршрутизации KDN

Fig. 2. Schematic Block Diagram of the Data Plane and KDN Forwarding Device

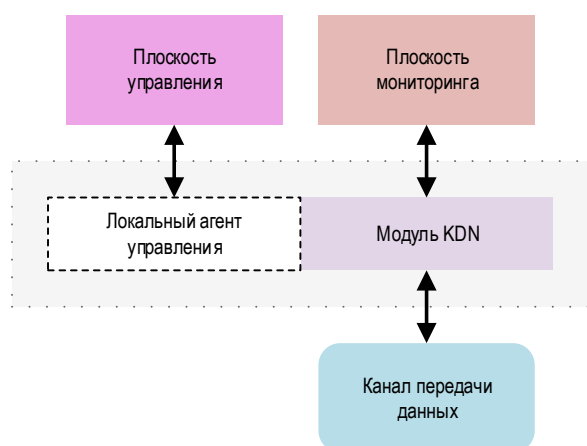


Рис. 3. Функциональная структура гибридного узла в KDN

Fig. 3. The Structure of a Hybrid Node in KDN

Локальный агент управления осуществляет обработку пакетов с использованием управляющих сообщений, полученных от плоскости управления, в то время как модуль KDN настраивает коммутатор на основе конфигураций, поступающих из плоскости мониторинга. Кроме того, модуль KDN собирает и передает данные обратно в упомянутые выше плоскости.

Если управляющий канал (канал, соединяющий гибридный узел и плоскость управления) доступен, тогда локальный агент находится в пассивном состоянии и перенаправляет полученные управляющие сообщения от внешнего контроллера в модуль KDN внутри узла. А в случае потери связи с централизованным контроллером, локальный агент управления берет на себя роль активного резервного контроллера, который переключается на использование заранее определенных политик и правил. В таком случае нормальная работа узла, включая отправку и прием данных, а также перенаправление пакетов, поступающих от других узлов, может быть обеспечена через использование канала передачи данных.

Плоскость управления

Плоскость управления обеспечивает маршрутизацию и обработку данных с использованием знаний, полученных из сетевых данных. В отличие от традиционных сетей, где контроль децентрализован и ограничен физической инфраструктурой, или SDN, где управление логически централизовано, контрольная плоскость KDN интегрирует элементы ML и анализа данных для создания интеллектуальных моделей управления сети.

Плоскость управления, состоящая из одного или более SDN-контроллера, координирует передачу информации между плоскостью данных и другими элементами, опираясь на глобальное представление о топологии и состоянии сети для эффективного управления потоками информации. Данные,

поступающие в плоскость управления, через плоскость знаний подвергаются высокоуровневому анализу с применением методов ML, что позволяет более точно прогнозировать поведение сети и оптимизировать ее производительность. Плоскость знаний анализирует телеметрию, исторические данные и текущие сетевые параметры для формирования стратегий и политик управления. Эти политики передаются обратно в плоскость управления, которая применяет их для обновления правил маршрутизации, управления трафиком и распределения ресурсов.

Плоскость управления KDN может быть реализована в рамках различных архитектур: централизованной, распределенной или гибридной. Следует отметить, что все узлы в представленных архитектурах должны соответствовать ранее описанной гибридной структуре узлов (см. рисунок 3).

Централизованная архитектура. Данная модель, представленная на рисунке 4а, сохраняет основную концепцию архитектуры SDN с логически централизованной плоскостью управления, где все действия узлов полностью управляются SDN-контроллером [7]. Такая архитектура отличается простотой, удобством управления и высокой гибкостью. Однако ее недостатком является повышенная задержка, обусловленная удаленной связью между узлами и контроллером, логически и физически отделенным от плоскости данных. Хотя такая схема эффективна для небольших сетей, ее производительность снижается по мере увеличения масштаба и нагрузки инфраструктуры. Контроллер может столкнуться с нехваткой ресурсов при обработке большого числа запросов от устройств, что приведет к ухудшению показателей. Кроме того, наличие единой точки отказа делает модель уязвимой и ограничивает ее надежность.

Распределенная архитектура. Традиционная архитектура SDN с логически централизованной плоскостью управления, разделяющей плоскости управления и данных, демонстрирует низкую производительность в междоменных, особенно в гетерогенных сетях. Для улучшения управления в таких сетях была предложена архитектура с логически распределенной плоскостью управления, где контроль распределяется между несколькими контроллерами [8], как показано на рисунке 4b. В данной модели, вместо централизованного SDN-контроллера, элементы передачи взаимодействуют с локальным контроллером, который обеспечивает управление в пределах своей зоны ответственности. В отличие от логически централизованной архитектуры изменения в сети не транслируются автоматически всем контроллерам – обновления передаются только при необходимости, например, при междоменных сервисах, когда контроллеры взаимодействуют друг с другом напрямую для обмена информацией.

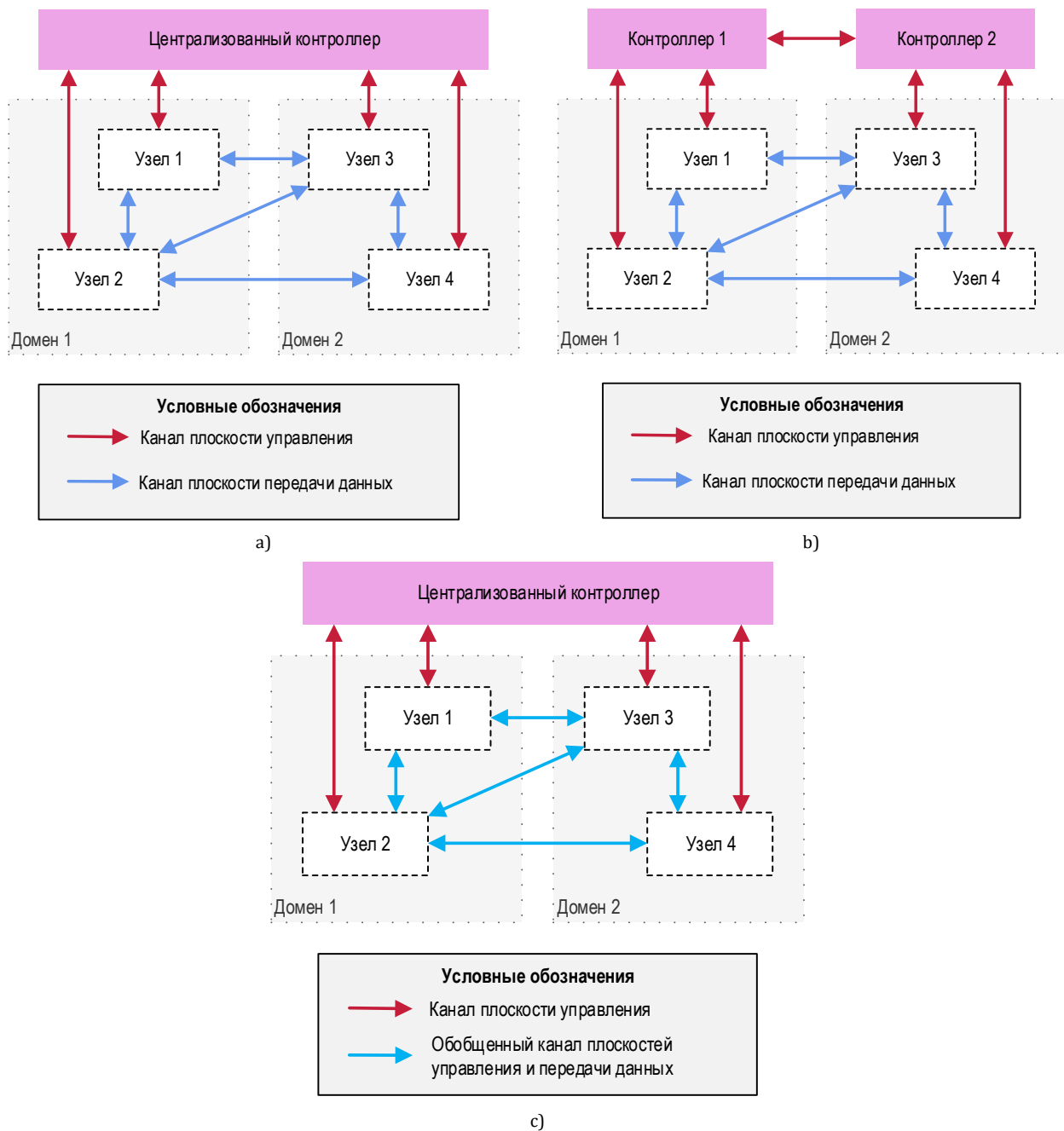


Рис. 4. Архитектура плоскости управления KDN: централизованная (а), распределенная (б) и гибридная (с)

Fig. 4. Architecture of the KDN Control Plane: Centralized (a), Distributed (b) and Hybrid (c)

Гибридная архитектура. Гибридная архитектура управления сочетает в себе черты централизованной и распределенной архитектуры, а контроль над сетью осуществляется совместно с плоскостями данных и управления [9]. В зависимости от условий сети, гибридная архитектура может гибко регулировать тип управления, автоматически переключая узлы и контроллеры (рисунок 4с). Коммуникация плоскости управления включает как прием управляющих сообщений от централизованного контроллера, так и обмен сообщениями

между узлами. Элементы плоскости передачи данных также участвуют в принятии решений и управлении сетью, что обеспечивает более высокую масштабируемость и устойчивость к сбоям.

Наряду с централизованной, распределенной и гибридной архитектурами существуют и альтернативные подходы к реализации управляющей плоскости KDN. Такие архитектуры могут включать многоуровневую организацию управления или механизмы контекстно-ориентированного взаимо-

действия, что позволяет более эффективно координировать обмен данными и поддерживать высокую производительность сети в условиях масштабирования и динамических изменений топологии.

Плоскость мониторинга

Основной целью плоскости мониторинга, которая функционирует параллельно с контроллером (управляющей плоскостью) в архитектуре KDN, является наблюдение за сетевыми устройствами, сбор данных и информации, обнаружение и

настройка параметров качества обслуживания (QoS, аббр. от англ. Quality of Service) в сети.

Как уже упоминалось ранее, в KDN функции мониторинга отделены от логики управления, и для стандартных сценариев (сбора данных, сетевого мониторинга и настройки) используются отдельные протоколы. Это упрощает устранение неисправностей в случае сбоев по сравнению с традиционной архитектурой SDN.

Схема плоскости мониторинга в архитектуре KDN представлена на рисунке 5.

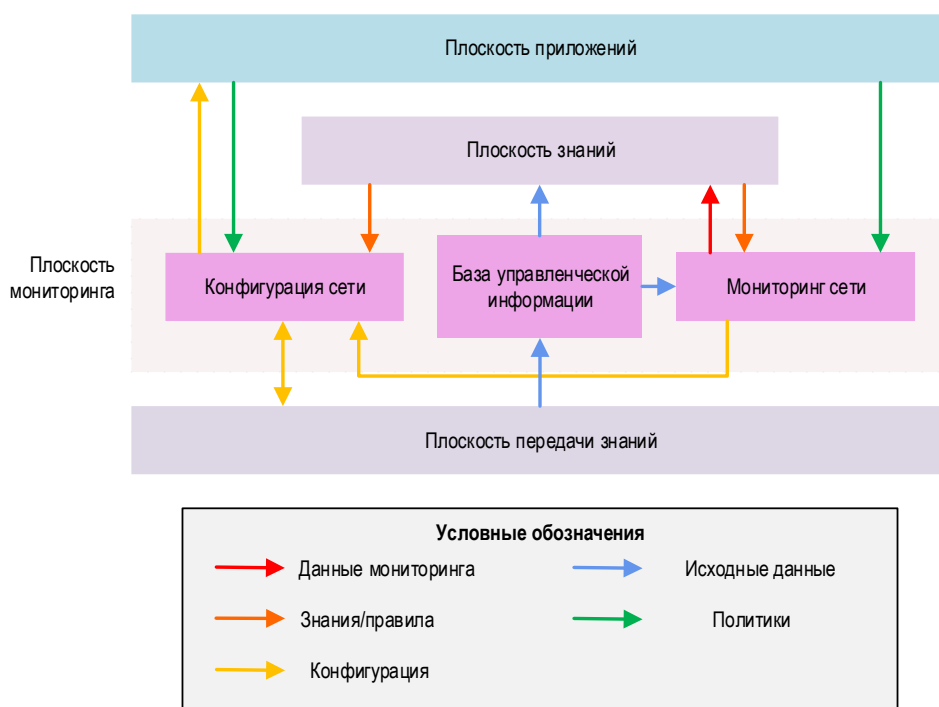


Рис. 5. Схематическая блок-диаграмма плоскости мониторинга KDN

Fig. 5. Schematic Block Diagram of the KDN Management Plane

В архитектуре KDN плоскость управления предоставляет унифицированный интерфейс для сбора данных, конфигурирования устройств и мониторинга сети. База управленческой информации отвечает за сбор и управление сетевыми данными, необходимыми для выполнения задач мониторинга и генерации знаний. Исходные данные из плоскости данных могут либо использоваться непосредственно в модуле мониторинга, либо передаваться в плоскость знаний для их дальнейшей обработки и генерации знаний.

Модуль сетевого мониторинга генерирует конфигурационные правила, сравнивая имеющиеся данные или знания с политиками сети. Анализ состояния сети может осуществляться как по запросу, так и непрерывно, а его результаты также могут направляться в плоскость знаний для генерации новых правил или знаний. Важно отметить, что, хотя это не показано на схеме, если ре-

зультаты мониторинга передаются в плоскость знаний, то могут быть созданы дополнительные знания или правила, которые, в свою очередь, могут повлиять на политики плоскости приложений.

Сгенерированные модулем мониторинга конфигурационные правила передаются в модуль конфигурации сети для непосредственной настройки элементов плоскости данных. Эти правила также могут использоваться для принятия решений по настройке сети, с учетом правил или знаний, полученных из плоскости знаний, и политик конфигурации, поступивших из прикладной плоскости.

Плоскость знаний

Плоскость знаний представляет собой логический уровень, отвечающий за генерацию, обработку и распространение знаний. Он помогает сетевым операторам (людям-администраторам, про-

граммным системам) эффективнее управлять поведением сети и точнее адаптироваться к изменяющимся условиям. В рамках данной плоскости используются декларативные (правила), процедурные (процессы) и контекстуальные знания (информация об окружающей среде и условиях работы сети) [10]. Они поступают из разных источников (сетевых устройств, сенсоров и т. д.), и анализируются для получения знаний и принятия интеллектуальных решений. Плоскость знаний помогает сетевым операторам заранее выявлять и устранять проблемы, оптимизировать поведение сети под изменяющиеся требования и внедрять усовершенствованные меры безопасности для защиты от угроз.

Плоскость знаний из-за постоянной обработки потока данных имеет наиболее сложную архитектуру среди всех плоскостей в KDN. Как показано на рисунке 6, плоскость знаний состоит из трех подуровней:

1) подуровень генерации знаний – генерирует описательные знания из исходных данных с использованием моделей генерации знаний;

2) подуровень композиции знаний – выявляет взаимосвязи между различными фрагментами знаний и создает комплексные знания с помощью редактора онтологий, которые после могут использоваться и для формирования правил, учитывающих, в том числе, и намерения пользователя;

3) подуровень управления и распределения знаний – отвечает за хранение информации, управление и обмен необработанными данными в пределах плоскости знаний, а также сформированными и сгенерированными знаниями, правилами и управляющими сообщениями между другими плоскостями.

Подуровни плоскости знаний детально описаны в следующих подразделах.

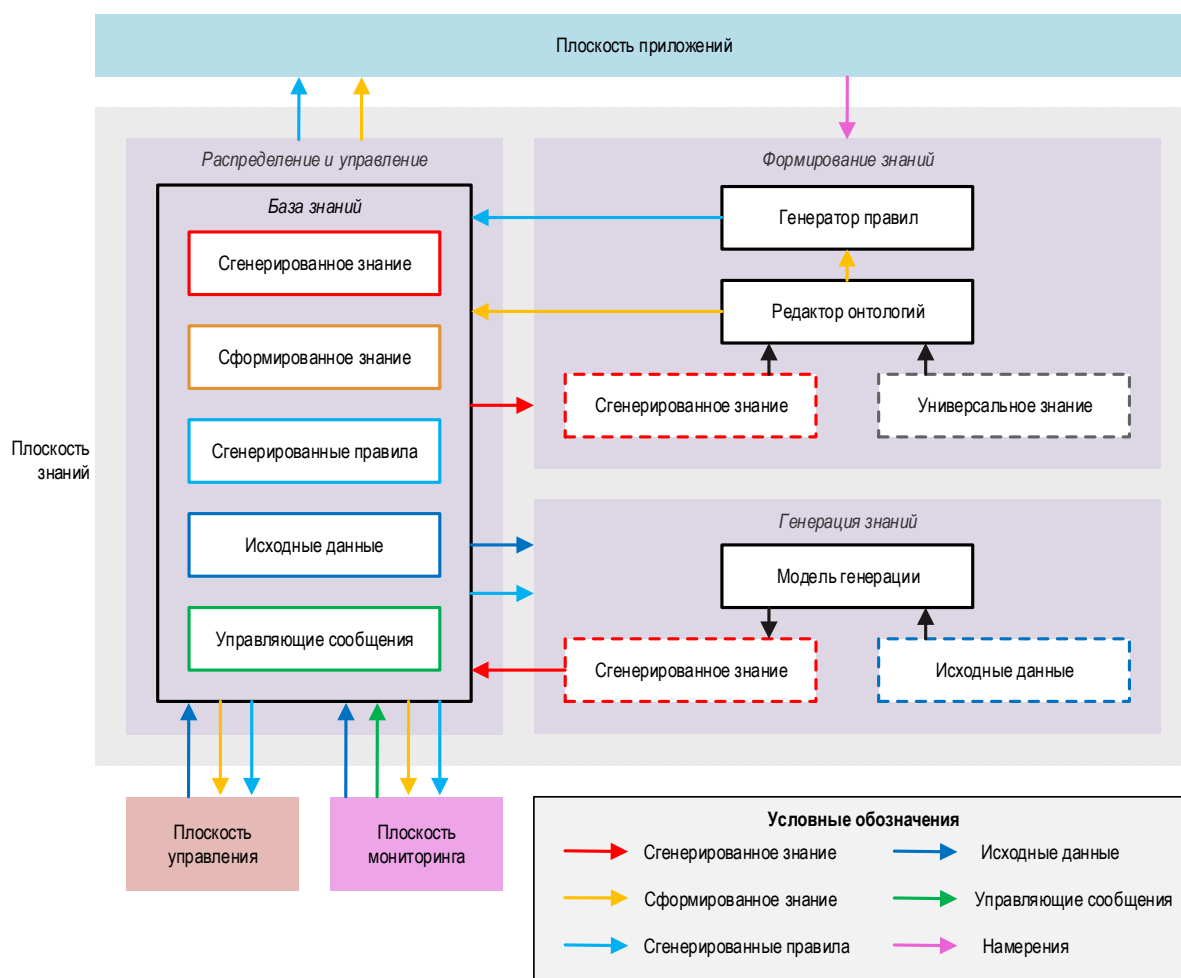


Рис. 6. Схематическая блок-диаграмма плоскости знаний KDN

Fig. 6. Schematic Block Diagram of the KDN Knowledge Plane

Подуровень генерации знаний

Подуровень генерации знаний выполняет извлечение описательных знаний из необработанных данных с использованием моделей генерации знаний. Этот процесс позволяет значительно сократить объем информации, передаваемой контроллеру, так как описательные знания являются более компактной и структурированной формой исходных данных. Исходные данные, собираемые на плоскостях управления и мониторинга, подаются в качестве входных данных для генераторов, в роли которых могут применяться как эвристические модели, так и методы ML.

Эвристические модели базируются на математических методах, позволяющих описывать исходные данные и выявлять внутренние корреляции между ними [11]. Они отличаются простотой реализации, низкой вычислительной сложностью и пригодны для задач, где требуется оперативная обработка информации в реальном времени. Однако возможности этих моделей ограничены, поскольку они плохо справляются с нелинейными и сложными зависимостями в данных, требуют высокой точности исходных данных и ручной настройки параметров. Такие модели не адаптированы для работы в условиях динамичных и изменчивых сетевых сред, но, несмотря на эти ограничения, эвристические методы могут быть полезны в рамках конвейерной обработки данных, где они дополняют более сложные методы ML.

Методы ML позволяют автономно анализировать данные и генерировать знания, что делает их особенно эффективными для сложных и динамичных сетевых сред. Для повышения точности алгоритмов, перед обучением таких моделей данные предварительно обрабатываются для выделения высокоуровневых признаков [12]. Основным преимуществом методов ML является их способность адаптироваться к изменениям в поведении сети, выявлять сложные закономерности и работать с большими объемами многомерных данных. Эти алгоритмы обеспечивают более высокую точность анализа и прогнозирования по сравнению с эвристическими подходами. В рамках генерации знаний могут применяться следующие типы ML.

1) *Методы с учителем (Supervised Learning)*. Этот подход заключается в обучении алгоритмов на размеченных данных, что позволяет делать прогнозирование на основе уже известных примеров [13]. Для достижения точных результатов входные данные должны быть четко размечены и отражать правильные ответы. Алгоритмы с учителем делятся на две основные категории: классификационные и регрессионные модели. Классификация используется для задач с дискретными выходными переменными, которые можно легко категоризо-

вать. Регрессия применяется для предсказания непрерывных числовых значений, основываясь на наблюдениях из прошлого. В контексте KDN наиболее распространенные примеры применения методов с учителем включают прогнозирование отказов соединений, определение требований к пропускной способности и оценку QoS.

2) *Методы без учителя (Unsupervised Learning)*. Этот подход обучает алгоритмы на неразмеченных данных, не обеспечивая четкими примерами для обучения. Основная идея состоит в том, чтобы предоставить системе большие объемы разнообразных данных и позволить ей самостоятельно извлекать полезные знания и выявлять скрытые закономерности [14]. Одним из наиболее популярных методов является кластеризация, которая группирует неразмеченные данные в схожие группы (кластеры). Другим примером является обнаружение аномалий – метод идентификации редких событий или наблюдений, которые значительно отличаются от большинства данных. В KDN эти методы применяются для кластеризации трафика и базовых станций, а также для мониторинга поведения устройств и сетей.

3) *Глубокое обучение (DL)* – раздел ML, основанный на использовании искусственных нейронных сетей с множеством скрытых слоев. Эти сети способны автоматически извлекать иерархические представления данных, что позволяет им решать комплексные и нелинейные задачи [15]. Алгоритмы DL обучаются путем настройки весов связей между нейронами через обратное распространение ошибки, таким образом DL может в процессе итераций самостоятельно вырабатывать условия для принятия корректных решений. Наиболее распространенные алгоритмы DL и области их применения представлены в таблице 1.

ТАБЛИЦА 1. Алгоритмы DL, применяемые в KDN

TABLE 1. Deep Learning Algorithms Used in KDN

Алгоритм	Ключевое применение
Долгая краткосрочная память (LSTM, аббр. от англ. Long Short-Term Memory)	Прогнозирование трафика, отслеживание мобильности, обнаружение неисправностей
Генеративно-состязательная сеть (GAN, аббр. от англ. Generative Adversarial Network)	Генерация синтетического трафика, противостояние угрозам безопасности, оптимизация QoS
Метод опорных векторов (SVM, аббр. от англ. Support Vector Machine)	Обнаружение аномалий, оптимизация маршрутизации, прогнозирование QoS

Тем не менее, описанные методы ML и DL зачастую требуют значительных вычислительных ресурсов и высококачественных размеченных наборов данных для обучения, которые в хаотичной сетевой системе могут быть недоступны. Кроме того, алгоритмы могут быть непрозрачными (проблема

«черного ящика», особенно характерная для DL), что усложняет интерпретацию процесса принятия решений.

Во многих случаях оптимальным решением может стать комбинирование обоих подходов. Эвристические модели могут выступать в качестве базовой или резервной системы, тогда как методы ML берет на себя более сложные задачи. Например, эвристическая модель может заниматься постоянной классификацией трафика, в то время как алгоритмы ML занимаются обнаружением аномалий или прогнозированием маршрутизации. Гибридные подходы объединяют сильные стороны обоих методов, что позволяет достичь баланса между эффективностью и адаптивностью.

Подуровень композиции знаний

В подуровне композиции знания, сгенерированные на подуровне генерации, проходят дальнейший анализ или объединяются для создания новых, комплексных знаний. Центральным элементом композиции знаний является редактор онтологий, который комбинирует универсальные, уже существующие, знания с множеством новых, сгенерированных знаний. Редактор онтологий представляет собой программное приложение, предназначенное для создания и поддержания онтологий, которые используются для определения свойств и иной информации с помощью онтологического языка.

Кроме того, подуровень композиции знаний отвечает на запросы и намерения плоскости приложений, сравнивая их с уже составленными правилами и политиками. Эта информация используется генератором правил для создания новых правил, которые применяются другими плоскостями сети.

Одним из динамичных методов генерации правил является обучение с подкреплением (Reinforcement Learning). Алгоритмы этого типа используют агентов (программные модули), которые самостоятельно обучаются выполнению задач через многократные пробные взаимодействия с динамичной средой. Основная цель агента – без явного программирования максимизировать вознаграждения и минимизировать штрафы, анализируя свои действия в конкретных ситуациях. Основные области применения этого метода в KDN включают маршрутизацию с учетом QoS, адаптивное распределение ресурсов, а также мониторинг и управление в сетях [16].

Подуровень управления и распределения знаний

Управление знаниями в контексте KDN представляет собой процесс сбора, хранения, обмена и использования знаний и правил в сетевой среде. Основная цель управления знаниями связана с по-

вышением производительности сети за счет оптимизации процесса обеспечения необходимых знаний и правил заинтересованным сторонам. Данная плоскость состоит из базы знаний, которая отвечает за хранение и распространение следующих элементов:

- знаний, сгенерированных на основе моделей генерации знаний;
- знаний, сформированных редакторами онтологий;
- правил, созданных генераторами правил;
- необработанных данных и управляющих сообщений, поступающих из сети.

Поскольку подуровень управления и распределения знаний собирает данные из всех частей сети, чтобы обеспечить совместимость различных систем KDN, необходимо наличие стандартов для их хранения, обмена и использования. Кроме того, в условиях ограниченных ресурсов сети должны быть установлены правила передачи трафика, связанного со знаниями (сгенерированные, сформированные знания или правила, созданные на их основе). Одной из задач подуровня распределения знаний является формирование правил передачи такого трафика как внутри плоскости знаний, так и его маршрутизация плоскостям управления, мониторинга и приложений.

Важно отметить, что плоскость знаний является центральным компонентом KDN, играющим ключевую роль во всей архитектуре. Учитывая ее критическую функцию, избыточность плоскости знаний имеет первостепенное значение для обеспечения устойчивости и надежности всей сети. Это подчеркивает преимущества гибридной архитектуры (см. рисунок 4с), где наличие избыточных узлов плоскости знаний на разных уровнях позволяет обеспечить плавное переключение и непрерывность работы даже в случае аппаратных или программных сбоев. Такой подход предотвращает возникновение единой точки отказа, которая могла бы нарушить критически важные функции, такие как маршрутизация трафика, распределение ресурсов и устранение неполадок в сети.

Плоскость приложений

Плоскость приложений является верхним уровнем в архитектуре KDN и отвечает за предоставление высокоуровневого представления сети, которое тесно связано с бизнес-целями и задачами пользователей, которые выражаются в форме намерений.

Плоскость приложений в KDN спроектирована для понимания требований приложений и динамической оптимизации их политик на основе знаний о состоянии сети. Это позволяет учитывать измене-

ния в сетевых условиях и эффективно адаптировать работу приложений к текущим потребностям и ограничениям. Кроме того, плоскость приложений предоставляет централизованную точку управления сетевыми политиками – администраторы могут задавать политики (которые относятся как к одному приложению, так и к их группе) и одновременно применять эти политики по всей сети.

Архитектура плоскости приложений и ее связь с другими элементами архитектуры KDN представлены на рисунке 7. Как видно из рисунка, плоскость приложений состоит из двух основных подуровней: *подуровня определения* намерений и политик, которая вложена в *подуровень обновления* намерений и политик.

Цель подуровня определения заключается в том, чтобы администраторы сети могли задавать намерения и политики вручную. После того как намерения и политики определены, их можно динамически развивать и изменять с использованием подуровня обновления намерений и политик. Для этого используются правила и знания, поступающие из плоскости знаний, конфигурации из плоскости мониторинга и сетевая информация (данные о топологии и статистика сети) полученная из плоскости управления.

Следует отметить, что в подуровне обновления намерений и политик имеется механизм правил, который обрабатывает данные из других плоскостей для принятия решений или создания инструкций по обновлению политик.

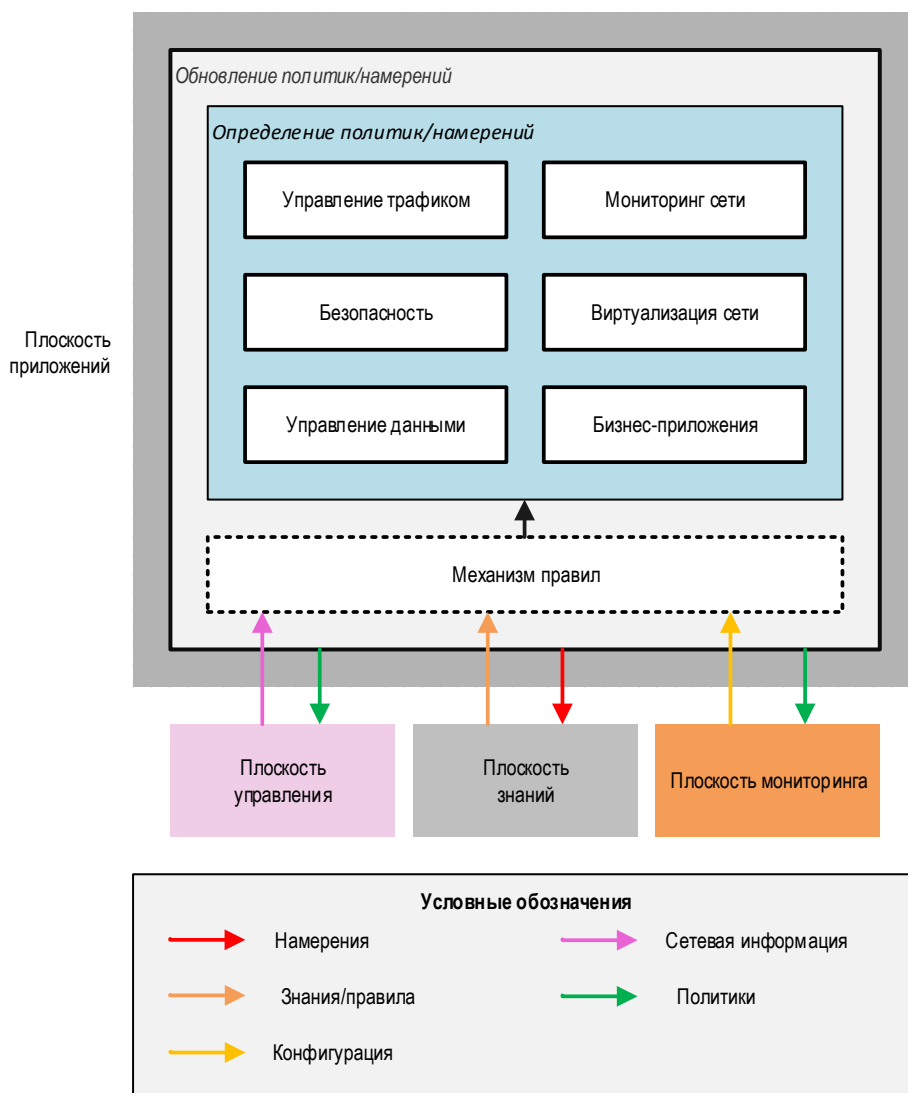


Рис. 7. Схематическая блок-диаграмма плоскости приложений KDN

Fig. 7. Schematic Block Diagram of the KDN Application Plane

В KDN возможно выделить шесть основных типов приложений.

1) *Управление трафиком*. Приложения отвечают за процесс оптимизации производительности и использования сетевых ресурсов для обеспечения эффективного потока данных и могут достигать этой цели различными способами [17], включая использование политик инженерии трафика, маршрутизацию, балансировку нагрузки и предоставление QoS.

2) *Мониторинг сети*. Приложения определяют политики для различных задач мониторинга, таких как управление отказами, управление мобильностью и управление энергопотреблением. Эти задачи реализуются в плоскости управления и для принятия решений в различных сценариях и учитывают и политики приложений, и актуальные данные о состоянии сети.

3) *Приложения безопасности*. Позволяют задавать политики для контроля доступа, шифрования, настройки межсетевых экранов, обнаружения и предотвращения атак, аномалий и вторжений. Такие приложения обеспечивают быстрое и эффективное выполнение политик безопасности, помогая гарантировать целостность, доступность и конфиденциальность сетевых ресурсов.

4) *Виртуализация сети*. Процесс заключается в создании нескольких логических сетей на физической инфраструктуре с использованием технологии разделения сети (Network Slicing) [18]. Это позволяет нескольким сетям сосуществовать и использовать одни и те же физические ресурсы, при этом каждая логическая сеть функционирует как отдельная сущность со своими политиками и конфигурациями. Виртуализация сети может осуществляться на уровне всей сети (Network-Level Slicing) или на уровне потоков (Flow-Level Slicing).

5) *Управление данными*. Эта категория включает приложения, работающие с объемами данных:

- *большие данные (Big Data)* работают с крупными и сложными наборами данных (по объему, скорости поступления и разнообразию), которые сложно обрабатывать традиционными методами [19]; KDN может решать проблемы с доставкой и обработкой данных в облачных центрах;

- *облачные вычисления* обеспечивают предоставление вычислительных ресурсов, включая серверы, хранилища, приложения и сервисы, по запросу через интернет; политики виртуализации сетевых функций позволяют эффективно использовать облачные ресурсы;

- *управление дата-центрами* (практика подключения множества серверов, хранилищ и других вычислительных ресурсов внутри одного центра обработки данных); такие приложения помогают минимизировать задержки, оптимально использовать ресурсы и повышать общую эффективность.

6) *Бизнес-приложения*. Приложения, ориентированные на получение прибыли и определяют политики, которые помогают организациям достигать бизнес-цели.

Дальнейшее развитие сетей управления на основе знаний

KDN на текущий момент находятся на начальном этапе своего развития, поэтому большая часть их концепции базируется на теоретических основах, а не на полностью реализованных практических системах. Это означает, что область KDN обладает значительным потенциалом для дальнейших исследований. Однако даже на данном этапе KDN представляют собой перспективную парадигму, вводящую в управление сетями знание-ориентированный подход. В отличие от традиционных сетей и SDN, KDN акцентирует внимание на интеллектуальности и автоматизации. Этот этап становления предоставляет исследователям уникальную возможность анализировать и формировать направление развития KDN, которое в перспективе станет неотъемлемой частью сетей следующего поколения.

По мере развития KDN возникают перспективы как для выявления преимуществ, так и для понимания возможных ограничений. Это позволяет уже на ранних этапах оценить потенциал технологии и повлиять на ее развитие. Так, например, *основными достоинствами KDN* являются следующие характеристики.

1) *Повышенная автоматизация и интеллектуальность*. KDN трансформирует управление сетями за счет внедрения ML и искусственного интеллекта (AI, аббр. от англ. Artificial Intelligence), что обеспечивает высокий уровень автоматизации и интеллектуальное принятие решений. Это позволяет в реальном времени адаптировать сеть, проводить предиктивный анализ и устранять неисправности, значительно снижая необходимость в ручном управлении. Используя знания, полученные из сетевых данных, системы KDN так же могут самоорганизовываться, адаптироваться к изменениям условий и внедрять эффективные решения, закладывая основу для полностью автономных сетей.

2) *Упрощенное управление сетью*. KDN упрощает управление сетью благодаря разделению плоскостей управления, данных и знаний, что создает единую модульную архитектуру для мониторинга, настройки и оптимизации сетевых операций. Такой подход позволяет операторам эффективно управлять сложными системами, минимизируя потребность в детальной ручной настройке. Плоскость знаний дополнительно автоматизирует рутинные задачи управления и динамически адаптируется к изменяющимся требованиям сети.

3) *Улучшение производительности сети и QoS.* Возможности KDN по оптимизации распределения ресурсов и управления трафиком напрямую влияют на повышение производительности сети. Используя знания в реальном времени и предиктивную аналитику, KDN обеспечивает снижение задержек, улучшение качества обслуживания и повышение надежности. Эти улучшения способствуют оптимизации пользовательского опыта, делая сеть более отзывчивой к требованиям конечных пользователей при сохранении высокой производительности.

4) *Снижение затрат на рабочую силу.* Автоматизация в KDN значительно снижает затраты на управление сетью, так как маршрутизация трафика, распределение ресурсов и обнаружение неисправностей выполняются автономно. Это уменьшает потребность в высококвалифицированных сетевых инженерах и позволяет организациям более эффективно распределять свои ресурсы.

5) *Повышение безопасности.* Интеграция AI и ML в KDN усиливает сетевую безопасность за счет способности в реальном времени обнаруживать и устранять опасности. Используя продвинутую аналитику, системы KDN могут проактивно выявлять уязвимости и прогнозировать потенциальные угрозы, создавая мощный механизм защиты. Централизованная плоскость знаний обеспечивает согласованность мер безопасности во всей сети, снижая риск изолированных уязвимостей.

С другой стороны, уже можно выделить и *недостатки KDN.*

1) *Проблемы с данными.* KDN в значительной степени зависит от больших объемов высококачественных данных для обучения моделей ML. Это создает серьезные сложности, связанные с хранением, обработкой и управлением информацией. Обеспечение однообразности и целостности знаний в сети может быть сложной задачей, особенно при работе с разнообразными источниками и форматами данных. Необходимость анализа данных в реальном времени, требуя устойчивой инфраструктуры для поддержки таких операций, так же усложняет процесс.

2) *Интеграция с существующими системами.* Переход от традиционных сетей или SDN к KDN создает проблемы совместимости с устаревшими системами. Во многих существующих сетях отсутствует инфраструктура, необходимая для интеграции плоскости знаний, что приводит к значительным затратам на модернизацию. Эти препятствия могут задерживать внедрение и увеличивать сложность развертывания решений KDN в реальных условиях.

3) *Интероперабельность и сложности внедрения.* Обеспечение интероперабельности устройств

и систем в KDN (как внутри одной сети, так и между двумя сетями) затруднено из-за многоплоскостной архитектуры и множества протоколов. С ростом сети поддержание согласованности и обмен данными становятся сложнее, а вычислительные требования ML-решений возрастают, что создает проблемы с масштабируемостью. Внедрение KDN требует значительных финансовых вложений в инфраструктуру, программное обеспечение и квалифицированный персонал. Постоянное обновление моделей ML и адаптация к технологиям повышают расходы, что затрудняет доступ для компаний с ограниченным бюджетом.

4) *Недостаток прозрачности.* Зависимость KDN от ML и AI часто приводит к недостаточной прозрачности процессов принятия решений. Операторам сети может быть сложно понять или проверить логику некоторых автоматических действий, особенно в случаях применения комплексных алгоритмов ML. Такая непрозрачность снижает уровень доверия и усложняет отладку или аудит, что создает серьезные препятствия для широкого внедрения.

Несмотря на это, траектория развития KDN предполагает действительно возможный переход к более адаптивному, предиктивному и контекстно-ориентированному сетевым средам. Развитие распределенного интеллекта в архитектурах KDN может обеспечить возможность совместной работы в реальном времени между периферийными и центральными системами, создавая сети, которые не только реагируют на текущие запросы, но и предвосхищают их. Кроме того, интеграция квантовых вычислений и более совершенных AI / ML-архитектур обещает значительно расширить возможности KDN, позволяя обрабатывать огромные объемы данных с беспрецедентной скоростью и точностью.

В развитии KDN выделяют *три ключевых направления.*

1) *Оптимизация существующих алгоритмов ML.* Развитие KDN напрямую зависит от совершенствования существующих алгоритмов ML, а также от их адаптации к специфическим требованиям динамичных сетевых сред. Индивидуально настроенные модели ML, приспособленные к различным сценариям, обеспечат более быстрое принятие решений и точное распределение ресурсов. Разработка новых алгоритмов с использованием таких передовых методов, как федеративное обучение [20] (децентрализованное / распределенное обучение на множестве физических платформ) или объяснимый AI [21] (системы и модели, способные объяснять свои действия), повысит адаптивность и прозрачность систем KDN, обеспечивая соответствие изменяющимся требованиям к производительности и этическим стандартам.

2) *Интеграция с существующими системами.* Критически важным направлением дальнейшего развития KDN является обеспечение плавной интеграции с устаревшими системами для минимизации сбоев и сокращения затрат на модернизацию. Инновации в гибридных архитектурах, связывающих традиционные сетевые устройства с плоскостью знаний KDN, могут сохранить существующую инфраструктуру при добавлении новых функций. Это позволит организациям внедрять KDN, сохраняя при этом непрерывность операций, обеспечивая экономичность модернизационных процессов.

4) *Оптимизация для самоорганизующихся сетей.* Будущее KDN так же тесно связано с инновациями в создании самоорганизующимися сетями, которые могут автономно организовываться, переконфигурироваться и оптимизироваться в ответ на изменения условий. Улучшение алгоритмов самообучения и использование принципов роевого интеллекта могут позволить сетям функционировать с минимальным участием человека, динамически адаптируясь к изменениям трафика, доступности ресурсов и потребностям пользователей для поддержания оптимальной производительности [22].

Заключение

В рамках проведенного анализа выявлено, что архитектура KDN предоставляет широкие возмож-

ности для автоматизации и повышения адаптивности сетей за счет интеграции плоскости знаний, способной в реальном времени генерировать и применять интеллектуальные стратегии управления. Рассмотрены преимущества использования KDN, такие как снижение операционных затрат, улучшение качества обслуживания, а также упрощение управления сетевой инфраструктурой в условиях растущей сложности сетевых сценариев.

Рассмотренная архитектура KDN, включающая плоскости знаний, мониторинга и управления, демонстрирует высокую степень автоматизации, гибкости и адаптивности. Эти свойства позволяют сетям, определяемым знаниями, не только повышать эффективность передачи данных, но и обеспечивать устойчивость и безопасность сетей в динамичных условиях. Интеграция методов ML и DL в рамках плоскости знаний открывает возможности для интеллектуального управления, автономного принятия решений и повышения QoS.

Дальнейшие исследования в этой области должны быть направлены на оптимизацию взаимодействия между различными плоскостями KDN, разработку методов интеграции с существующими сетевыми инфраструктурами и преодоление ограничений, связанных с необходимостью значительных вычислительных ресурсов. Реализация этих задач обеспечит более широкое внедрение KDN в практику построения сетей нового поколения.

Список источников

1. Zoraida B.S.E., Ganesan I. A Comparative Study on Software-Defined Network with Traditional Networks // TEM Journal. 2024. Vol. 13. Iss. 1. PP. 167–176. DOI:10.18421/TEM131-17
2. Hakiri A., Gokhale A., Berthou P., Schmidt D.C., Gayraud T. Software-defined Networking: Challenges and Research Opportunities for Future Internet // Computer Networks. 2014. Vol. 75. Part A. PP. 453–471. DOI:10.1016/j.comnet.2014.10.015
3. Ashtari S., Zhou I., Abolhasan M., Shariati N., Lipman J., Ni W. Knowledge-defined networking: Applications, challenges and future work // Array. 2022. Vol. 14. P. 100136. DOI:10.1016/j.array.2022.100136
4. Jarrahi M.H., Askay D., Eshraghi A., Smith P. Artificial intelligence and knowledge management: A partnership between human and AI // Business Horizons. 2023. Vol. 66. Iss. 1. PP. 87–99. DOI:10.1016/j.bushor.2022.03.002
5. Srinivas T.A., Donald A., Thippanna G., Kousar M., Priya A. From Control to Chaos: The Dynamic SDN Control Plane // International Journal of Advanced Research in Science, Communication and Technology. 2023. Vol. 3. Iss. 2. PP. 494–502. DOI:10.48175/IJARSC-8527
6. Mestres A., Rodriguez-Natal A., Carner J., Barlet-Ros P., Alarcón E., Solé-Simó M., et al. Knowledge-Defined Networking // ACM SIGCOMM Computer Communication Review. 2017. Vol. 47. Iss. 3. PP. 2–10. DOI:10.1145/3138808.3138810
7. Li W. Explore the Evolution of Computer Network Architecture // International Journal of Religion. 2024. Vol. 5. № 11. PP. 9034–9054. DOI:10.61707/ahr2da31
8. Das R., Pohrmen F., Maji A., Saha G. FT-SDN: A Fault-Tolerant Distributed Architecture for Software Defined Network // Wireless Personal Communications. 2020. Vol. 114. PP. 1045–1066. DOI:10.1007/s11277-020-07407-x
9. Wang T., Su Z., Hamdi M. Rethinking the Data Center Networking: Architecture, Network Protocols, and Resource Sharing // IEEE Access. 2014. Vol. 2. PP. 1481–1496. DOI:10.1109/ACCESS.2014.2383439
10. Agoulmine N. Chapter 1. Introduction to Autonomic Concepts Applied to Future Self-Managed Networks // In: Walsh S.M., Strano M.S. (ed.) Autonomic Network Management Principles. Academic Press, 2011. PP. 1–26. DOI:10.1016/B978-0-12-382190-4.00001-2
11. Silver E. An overview of heuristic solution methods // Journal of The Operational Research Society. 2004. Vol. 55. Iss. 9. PP. 936–956. DOI:10.1057/palgrave.jors.2601758
12. Oladipupo T. Machine Learning Overview // In: Zhang Y. (ed.) New Advances in Machine Learning. InTech, 2010. PP. 9–18. DOI:10.5772/9374
13. Nasteski V. An overview of the supervised machine learning methods // Horizons. 2017. Vol. 4. PP. 51–62. DOI:10.20544/HORIZONS.B.04.1.17.P05

14. Yazici İ., Shayea I., Din J. A survey of applications of artificial intelligence and machine learning in future mobile networks-enabled systems // *Engineering Science and Technology, an International Journal*. 2023. Vol. 44. P. 101455. DOI:10.1016/j.jestch.2023.101455
15. Zhang C., Patras P., Haddadi H. Deep Learning in Mobile and Wireless Networking: A Survey // *IEEE Communications Surveys & Tutorials*. 2019. Vol. 21. Iss. 3. PP. 2224–2287. DOI:10.1109/COMST.2019.2904897
16. Magadum A.A., Ranjan A., Narayan D.G. DeepQoS: A Deep Reinforcement Learning based QoS-Aware Routing for Software Defined Data Center Networks // *Proceedings of the 12th International Conference on Computing Communication and Networking Technologies (ICCCNT, Kharagpur, India, 06–08 July 2021)*. IEEE, 2021. DOI:10.1109/ICCCNT51525.2021.9579514
17. Volokyta A., Kogan A., Cherevatenko O., Korenko D., Oboznyi D., Kulakov Y. Traffic Engineering with Specified Quality of Service Parameters in Software-defined Networks // *International Journal of Computer Network and Information Security (IJCNIS)*. 2024. Vol. 16. Iss. 5. PP. 1–13. DOI:10.5815/ijcnis.2024.05.01
18. Sood M. 5G Network Slicing to Deliver Improved User Experience // *International Journal of Computer Trends and Technology*. 2023. Vol. 71. Iss. 9. PP. 59–68. DOI:10.14445/22312803/IJCTT-V71I9P107
19. Yu S., Liu M., Dou W., Liu X., Zhou S. Networking for Big Data: A Survey // *IEEE Communications Surveys & Tutorials*. 2017. Vol. 19. Iss. 1. PP. 531–549. DOI:10.1109/COMST.2016.2610963
20. Wen J., Zhang Z., Lan Y., Cui Z., Cai J., Zhang W. A survey on federated learning: challenges and applications // *International Journal of Machine Learning & Cybernetics*. 2023. Vol. 14. PP. 513–535. DOI:10.1007/s13042-022-01647-y
21. Strannegård C., Häggström O., Wessberg J., Balkenius C. Transparent Neural Networks: Integrating Concept Formation and Reasoning // *Proceedings of the 5th International Conference on Artificial General Intelligence (AGI, Oxford, UK, 8–11 December 2012)*. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2012. Vol. 7716. PP. 302–311. DOI:10.1007/978-3-642-35506-6_31
22. El Boudani B., Dagiuklas T., Iqbal M. SO-KDN: A Self-Organised Knowledge Defined Networks Architecture for Reliable Routing. 2021. PP. 160–166. DOI:10.1145/3459955.3460617

References

1. Zoraida B.S.E., Ganesan I. A Comparative Study on Software-Defined Network with Traditional Networks. *TEM Journal*. 2024;13(1):167–76. DOI:10.18421/TEM131-17
2. Hakiri A., Gokhale A., Berthou P., Schmidt D.C., Gayraud T. Software-defined Networking: Challenges and Research Opportunities for Future Internet. *Computer Networks*. 2014;75(A):453–71. DOI:10.1016/j.comnet.2014.10.015
3. Ashtari S., Zhou I., Abolhasan M., Shariati N., Lipman J., Ni W. Knowledge-defined networking: Applications, challenges and future work. *Array*. 2022;14:100136. DOI:10.1016/j.array.2022.100136
4. Jarrahi M.H., Askay D., Eshraghi A., Smith P. Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*. 2023;66(1):87–99. DOI:10.1016/j.bushor.2022.03.002
5. Srinivas T.A., Donald A., Thippanna G., Kousar M., Priya A. From Control to Chaos: The Dynamic SDN Control Plane. *International Journal of Advanced Research in Science, Communication and Technology*. 2023;3(2):494–502. DOI:10.48175/IJARSC-8527
6. Mestres A., Rodriguez-Natal A., Carner J., Barlet-Ros P., Alarcón E., Solé-Simó M., et al. Knowledge-Defined Networking. *ACM SIGCOMM Computer Communication Review*. 2017;47(3):2–10. DOI:10.1145/3138808.3138810
7. Li W. Explore the Evolution of Computer Network Architecture. *International Journal of Religion*. 2024;5(11):9034–9054. DOI:10.61707/ahr2da31
8. Das R., Pohrmen F., Maji A., Saha G. FT-SDN: A Fault-Tolerant Distributed Architecture for Software Defined Network. *Wireless Personal Communications*. 2020;114:1045–66. DOI:10.1007/s11277-020-07407-x
9. Wang T., Su Z., Hamdi M. Rethinking the Data Center Networking: Architecture, Network Protocols, and Resource Sharing. *IEEE Access*. 2014;2:1481–96. DOI:10.1109/ACCESS.2014.2383439
10. Agoulmine N. Chapter 1. Introduction to Autonomic Concepts Applied to Future Self-Managed Networks. In: *Walsh S.M., Strano M.S. (ed.) Autonomic Network Management Principles*. Academic Press; 2011. p.1–26. DOI:10.1016/B978-0-12-382190-4.00001-2
11. Silver E. An overview of heuristic solution methods. *Journal of the Operational Research Society*. 2004;55(9):936–56. DOI:10.1057/palgrave.jors.2601758
12. Oladipupo T. Machine Learning Overview. In: *Zhang Y. (ed.) New Advances in Machine Learning*. InTech; 2010. p.9–18. DOI:10.5772/9374
13. Nasteski V. An overview of the supervised machine learning methods. *Horizons*. 2017;4:51–62. DOI:10.20544/HORIZONS.B.04.1.17.P05
14. Yazici İ., Shayea I., Din J. A survey of applications of artificial intelligence and machine learning in future mobile networks-enabled systems. *Engineering Science and Technology, an International Journal*. 2023;44:101455. DOI:10.1016/j.jestch.2023.101455
15. Zhang C., Patras P., Haddadi H. Deep Learning in Mobile and Wireless Networking: A Survey. *IEEE Communications Surveys & Tutorials*. 2019;21(3):2224–87. DOI:10.1109/COMST.2019.2904897
16. Magadum A.A., Ranjan A., Narayan D.G. DeepQoS: A Deep Reinforcement Learning based QoS-Aware Routing for Software Defined Data Center Networks. *Proceedings of the 12th International Conference on Computing Communication and Networking Technologies, ICCNT, 06–08 July 2021, Kharagpur, India*. IEEE; 2021. DOI:10.1109/ICCCNT51525.2021.9579514

17. Volokyta A., Kogan A., Cherevatenko O., Korenko D., Oboznyi D., Kulakov Y. Traffic Engineering with Specified Quality of Service Parameters in Software-defined Networks. *International Journal of Computer Network and Information Security (IJCNIS)*. 2024;16(5):1–13. DOI:10.5815/ijcnis.2024.05.01
18. Sood M. 5G Network Slicing to Deliver Improved User Experience. *International Journal of Computer Trends and Technology*. 2023;71(9):59–68. DOI:10.14445/22312803/IJCTT-V71I9P107
19. Yu S., Liu M., Dou W., Liu X., Zhou S. Networking for Big Data: A Survey. *IEEE Communications Surveys & Tutorials*. 2017;19(1):531–49. DOI:10.1109/COMST.2016.2610963
20. Wen J., Zhang Z., Lan Y., Cui Z., Cai J., Zhang W. A survey on federated learning: challenges and applications. *International Journal of Machine Learning & Cybernetics*. 2023;14:513–35. DOI:10.1007/s13042-022-01647-y
21. Strannegård C., Häggström O., Wessberg J., Balkenius C. Transparent Neural Networks: Integrating Concept Formation and Reasoning. *Proceedings of the 5th International Conference on Artificial General Intelligence, AGI, 8–11 December 2012, Oxford, UK. Lecture Notes in Computer Science, vol.7716*. Berlin, Heidelberg: Springer; 2012. p.302–311. DOI:10.1007/978-3-642-35506-6_31
22. El Boudani B., Dagiuklas T., Iqbal M. SO-KDN: A Self-Organised Knowledge Defined Networks Architecture for Reliable Routing. 2021;160–6. DOI:10.1145/3459955.3460617

Статья поступила в редакцию 28.11.2024; одобрена после рецензирования 18.04.2025; принята к публикации 21.04.2025.

The article was submitted 28.11.2024; approved after reviewing 18.04.2025; accepted for publication 21.04.2025.

Информация об авторах:

БЛАН
Филип Сергеевич

аспирант кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0002-8760-1089>

ЕЛАГИН
Василий Сергеевич

кандидат технических наук, доцент, доцент кафедры Инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0003-4077-6869>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

**2.3.1 – Системный анализ,
управление и обработка
информации, статистика**

**2.3.6 – Методы и системы защиты
информации, информационная
безопасность**

Научная статья

УДК 519.872

<https://doi.org/10.31854/1813-324X-2025-11-2-84-100>

EDN:SIYZKF



Модель представления решающих функций для метрических методов распознавания образов

✉ Игорь Васильевич Александров, aleksandrov.iv@sut.ruРоман Михайлович Вивчарь, vivchar.rm@sut.ruРуслан Валентинович Киричек, kirichek@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

В настоящее время ключевую роль поддержки принятия решений в различных отраслях играет интеллектуальный анализ данных. Его важной составной частью является машинное обучение, **актуальная задача** которого на практике – осуществление классификации объектов в режиме реального времени. Она может быть достигнута за счет распараллеливания алгоритмов обработки данных как по входным данным, так и по данным решающих функций. Для повышения эффективности распараллеливания методов машинного обучения разработана унифицированная модель представления решающих функций. **Целью** настоящей статьи является представление унифицированной модели решающих функций для алгоритмов машинного обучения и функций, обеспечивающих ее распараллеливание как по входным данным, так и по данным решающих функций.

Сущность представленного подхода заключается в том, что на основе анализа особенностей работы метрических методов машинного обучения выявлены независимые данные для обработки, представленные в виде разных категорий анализируемого свойства, разработана модель представления решающих функций, описывающая выявленные характерные черты объектов входных данных и данных решающих функций в виде соответствующих наборов унифицированных элементов и включающая в себя функции, обеспечивающие их параллельную обработку на основе группового распараллеливания объектов. Предложенный подход базируется на использовании методов анализа алгоритмов и вычислительной сложности, математической статистики и методологии проектирования параллельных алгоритмов.

Эксперименты показали, что распараллеливание предложенной модели решающих функций на примере метода потенциальных функций позволяет повысить эффективность классификации как для одного объекта за счет возможности использования дополнительных вычислительных ресурсов, так и для группы объектов при наличии ограничений объема памяти компьютеров или горизонта планирования.

Научная новизна предложенного подхода заключается в том, что модель отличается от существующих единым способом формализации объектов и их свойств в виде унифицированных элементов для обучающих и классификационных данных и имеет структуру и функции, ориентированные на ее параллельную обработку методами распознавания образов на основе решающих функций в рамках группового распараллеливания объектов.

Теоретическая значимость: модель имеет унифицированный характер и может использоваться при распараллеливании других методов распознавания образов, которые могут быть описаны сходными параметрами, архитектурой и классификационными признаками.

Практическая значимость предложенного подхода заключается в том, что модель позволяет осуществить декомпозицию задачи классификации образов на отдельные подзадачи поиска закономерностей между входными данными и данными решающих функций.

Ключевые слова: распознавание образов, метод решающих функций, машинное обучение

Ссылка для цитирования: Александров И.В., Вивчарь Р.М., Киричек Р.В. Модель представления решающих функций для метрических методов распознавания образов // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 84–100. DOI:10.31854/1813-324X-2025-11-2-84-100. EDN:SIYZKF

Original article

<https://doi.org/10.31854/1813-324X-2025-11-2-84-100>

EDN:SIYZKF

Decision Functions Model for Metric Methods of Pattern Recognition

 **Igor V. Alexandrov** ✉, aleksandrov.iv@sut.ru

 **Roman M. Vivchar**, vivchar.rm@sut.ru

 **Ruslan V. Kirichkek**, kirichkek@sut.ru

The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Currently, data mining based on machine learning plays a key role in decision support in various industries. An important practical problem of machine learning is the implementation of object classification in real time, which can be achieved by parallelizing data processing algorithms both for input data and for decision function data. To improve the efficiency of parallelizing machine learning methods, a unified decision function model has been developed. **The Relevance** of this research is to present a unified decision function model in the framework of machine learning algorithms and functions for its parallelization both in terms of input data and decision function data.

The essence of the presented approach is that the features of the operation of metric methods of machine learning are analyzed, independent data for processing are presented using different categories of the analyzed property, developed decision function model describes the object features for input data and decision function data using standardized elements and including functions for their parallel processing based on group parallelization of objects. The proposed approach is based on the use of methods for analyzing algorithms and computational complexity, mathematical statistics and the methodology of designing parallel algorithms.

Experiments have shown that parallelization of the proposed decision function model for the potential function method allows increasing the classification efficiency for one object using additional computing resources, and for a group of objects within the limits of the computer's memory size or planning horizon.

The novelty of the proposed approach is that the model differs from existing ones in a method of formalizing objects and their features using unified elements for training and classification objects and has a structure and functions oriented towards its parallel processing by pattern recognition methods based on decision functions within the framework of group parallelization of objects.

Theoretical significance: the model is unified and can be used to parallelize other pattern recognition methods that can be described by similar parameters, architecture, and classification features.

The practical significance of the proposed approach is that the model allows decomposing the pattern classification problem into separate subtasks of finding regularities between input data and decision function data.

Keywords: pattern recognition, decision function method, machine learning

For citation: Aleksandrov I.V., Vivchar R.M., Kirichkek R.V. Decision Functions Model for Metric Methods of Pattern Recognition. *Proceedings of Telecommunication Universities*. 2025;11(2):84–100. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-84-100. EDN:SIYZKF

Введение

Наиболее перспективным направлением информационных технологий, используемым для организации поддержки принятия решений [1, 2] в таких областях, как экономика, образование, медицина и телекоммуникации, является интеллектуальный анализ данных [3, 4]. Одним из актуальных направлений исследований в области интел-

лектуального анализа данных являются вычислительные системы машинного обучения [5, 6]. Для решения задач машинного обучения применяется группа методов решающих функций [7, 8], основанных на алгоритмах метрической классификации и оценивания сходства объектов. Наличие обучающих данных является обязательным для использования методов решающих функций, по-

этому эти методы также называют алгоритмами, основанными на прецедентах [9]. Главным достоинством методов решающих функций является их высокая эффективность с точки зрения достоверности результата, а главным недостатком – вычислительно-емкие алгоритмы обработки данных. Особенно это актуально при работе с большими объемами данных.

В последнее десятилетие повышение производительности вычислительных узлов связано с развитием многоядерных процессоров и сети Интернет, что привело к распространению облачных систем. Однако современное программное обеспечение для решения задач машинного обучения значительно отстает от аппаратной части компьютерной техники и часто неэффективно использует предоставляемые вычислительные ресурсы [10]. Данная ситуация связана со следующими проблемами:

- 1) трудоемкость задачи распараллеливания вычислительных алгоритмов;
- 2) разнообразие типов данных решающих функций;
- 3) необходимость получения информации в режиме реального времени.

Способом решения рассмотренных проблем является разработка унифицированной модели представления решающих функций для метрических методов распознавания образов. В работах [11–13] представлена модель знаний для алгоритмов интеллектуального анализа данных, обеспечивающая распараллеливание вычислений только по входным данным. Предлагаемая модель отличается от вышеуказанной возможностью распараллеливания вычислений не только по входным данным, но и по данным решающих функций, что обеспечивает повышение эффективности классификации объекта в режиме реального времени.

Формальное описание данных

Входными данными для классификации образов являются множества однородных объектов $X = \{x_1, \dots, x_s\}$ – событий, явлений, сделок и др. – и меток их классов $Y = \{y_1, \dots, y_s\}$ – категорий, на которые делятся объекты. Каждый анализируемый объект описывается в виде набора из n атрибутов (длина, вес, возраст и др.). На практике наиболее часто используемой формой представления объекта является вещественный вектор числовых значений его атрибутов. Поэтому в работе ограничимся следующим способом представления объекта x :

$$x = (v_1^{Object}, \dots, v_i^{Object}, \dots, v_n^{Object}), \quad (1)$$

где $v_i^{Object} \in R$ – числовое значение i -го атрибута объекта x .

Набор допустимых классов U объекта x определим как конечное множество неповторяющихся меток из m элементов:

$$U = (u_1, \dots, u_m), \quad (2)$$

$$u_k \cap u_j = \emptyset \quad \forall k \neq j (k, j \in \{1..m\}),$$

где u_j – j -я метка из множества U .

Процесс определения класса u_i объекта x_i происходит на основе анализа его черт $\{q_1, \dots, q_w\}$, отражающих необходимые свойства данных для отделения объектов одного класса от другого в рамках применяемого метода распознавания образов. Проанализировав особенности работы рассматриваемых метрических методов распознавания образов, следует отметить, что каждый искомый параметр для объекта x_i также можно описать в виде вещественного вектора числовых значений анализируемых атрибутов:

$$q = (v_1^{Property}, \dots, v_i^{Property}, \dots, v_m^{Property}), \quad (3)$$

где $v_i^{Property} \in R$ – числовое значение i -го анализируемого атрибута особенности данных.

Дополнительно определим следующие технические и алгоритмические элементы данных.

1) Технический элемент *null*, обозначающий отсутствие какой-либо информации для рассматриваемого объекта данных.

2) Технический элемент *id*, обозначающий идентификационный номер объекта или класса в рамках всего множества данных. Например, идентификационным номером класса u является его порядковый номер id_u во множестве U .

3) Параметр p определяет особенность работы метода распознавания (количество искомых объектов, число итерации и др.) и представляет собой следующий кортеж:

$$p = \langle n_{Parameter}, v_{Parameter} \rangle, \quad (4)$$

где $n_{Parameter}$ – имя параметра; $v_{Parameter}$ – значение параметра.

4) Множество параметров:

$$P = \{p_1, \dots, p_g\}, \quad (5)$$

где p_i – i -й параметр работы алгоритма.

Когда множество параметров $\{p_1, \dots, p_g\}$ фиксировано и однозначно определено, множество P может быть упрощено до списка значений его параметров $P \equiv \{v_1^{Parameter}, \dots, v_g^{Parameter}\}$. В дальнейшем под множеством параметров P будет пониматься только набор их числовых значений.

Рассматриваемая задача распознавания образов состоит из следующих двух этапов [14].

Этап 1. Обучение

В рамках этого этапа определяются необходимые свойства данных для наборов обучающих объектов в рамках соответствующих решающих функций. Решающие функции могут содержать такие особенности обучающих данных, как усред-

ненный эталон для набора объектов одного класса, гиперплоскости, разделяющие объекты одного класса от других, подмножества объектов обучающей выборки.

Этап 2. Классификация

Целью этого этапа является нахождение искоемых параметров тестового объекта x_i^{Test} для определения его класса y_i^{Test} из множества U на основе данных решающих функций. В качестве искоемых параметров тестового объекта могут выступать такие параметры данных, как расстояние от тестового объекта до усредненного эталона группы объектов i -го класса, K ближайших соседей в обучающей выборке и др.

Обобщим способы представления объекта x , его класса y и черты q (1–3) для вышеуказанных этапов работы рассматриваемых методов распознавания образов в виде понятия унифицированного факта следующим образом:

$$w = \{id_{Fact}, y_{Fact}, V_{Fact}^{Data}, V_{Fact}^{Category}\}, \quad (6)$$

где id_{Fact} – идентификационный номер факта w в рамках множества объектов анализируемой характерной черты данных; y_{Fact} – класс описываемых фактом w данных; V_{Fact}^{Data} – описываемые фактом w данные (объект, разделяющая плоскость, расстояние, счетчик объектов и др.); $V_{Fact}^{Category}$ – набор соответствующих категорий для всех анализируемых свойств данных в рамках применяемого метода распознавания образов (более подробно будут рассмотрены в следующем разделе).

Обобщенная структура факта w реализуется на основе возможности назначения любому его аргументу вышеуказанного технического элемента *null*.

На основе сформулированного понятия факта (6) определим основные элементы данных.

1) Представим неделимый набор фактов в виде понятия параметра данных следующим образом:

$$k = \{w_1, \dots, w_d\}, \quad (7)$$

где w_j – j -й факт параметра k .

2) Независимое i -е множество объектов для анализа:

$$X_i = \{x_{i,1}, \dots, x_{i,a}\}, \quad (8)$$

где $x_{i,j}$ – j -й объект множества X_i .

3) Набор соответствующих классов для X_i :

$$K_i^{Class} = \{k_{i,1}^{Class}, \dots, k_{i,a}^{Class}\}, \quad (9)$$

где $k_{i,j}^{Class} = \{w_{i,j,1}^{Class}\}$ – класс j -го объекта $x_{i,j}$ множества X_i .

4) Набор выявленных однотипных черт для X_i :

$$K_i^{Property} = \{k_{i,1}^{Property}, \dots, k_{i,b}^{Property}\}, \quad (10)$$

где $k_{i,j}^{Property} = \{w_{i,j,1}^{Property}\}$ – j -я черта множества X_i .

Количество черт в $K_i^{Property}$ в обобщенном виде может не совпадать с числом объектов в X_i . Так, например, для каждого набора объектов одного класса в рамках метода эталонов существует только один усредненный эталон.

5) Множество объектов X_i , определенное на основе фактов:

$$K_i^{Object} = \{k_{i,1}^{Object}, \dots, k_{i,a}^{Object}\}, \quad (11)$$

где $k_{i,j}^{Object} = \{w_{i,j,1}^{Object}\}$ – j -й объект множеств X_i и K_i^{Object} .

В рамках разработки обобщенной модели представления решающих функции объекты K_i^{Object} множества X_i показаны в виде его свойства по аналогии с классами K_i^{Class} и параметрами $K_i^{Property}$.

6) Обобщенная особенность множества X_i :

$$K_i = \{k_{i,1}, \dots, k_{i,q}\} = \begin{cases} K_i^{Object}, & \text{если описываются объекты} \\ K_i^{Class}, & \text{если описываются классы} \\ K_i^{Property}, & \text{если описываются параметры} \end{cases}, \quad (12)$$

где $k_{i,j}$ – j -я особенность множества X_i .

Таким образом, под обрабатываемыми данными в дальнейшем будем понимать особенности (объекты, классы и параметры) анализируемых объектов.

7) Характеристика анализируемых объектов, описывающая множество однотипных особенностей:

$$c_i = \{K_{i,1}, \dots, K_{i,w}\}, \quad (13)$$

где $K_{i,j}$ – j -я особенность данных.

8) Набор характеристик, описывающих единую специфику данных, в виде характерной черты:

$$C_i = \{c_{i,1}, \dots, c_{i,e}\}, \quad (14)$$

где $c_{i,j}$ – j -я характеристика данных.

9) Множество взаимосвязанных характерных черт:

$$d_k = \{C_{k,1}, \dots, C_{k,t}\}, \quad (15)$$

где $C_{k,j}$ – j -я характерная черта данных.

10) Множество наборов характерных черт:

$$D_i = \{d_{i,1}, \dots, d_{i,r}\}, \quad (16)$$

где $d_{i,j}$ – j -й независимый набор характерных черт данных.

Формальное представление свойств независимых данных

В работе [11] была предложена концепция представления шагов работы алгоритмов интеллектуального анализа данных в виде функции поиска $f_{Characteristic}$ соответствующих закономерностей между входными и целевыми данными в рамках применяемого метода их обработки. Обобщим рассматриваемую концепцию для анализируемых метрических методов распознавания образов на основе вышеуказанных элементов данных (4–14):

$$C_{New} = f_{Characteristic}(C_{Input}, C_{Goal}, P_{Characteristic}), \quad (17)$$

где C_{New} – множество с искомыми особенностями анализируемых данных:

$$C_{New} = \{c_1^{New} = \{K_{1,1}^{New}, \dots, K_{1,c}^{New}\}\};$$

C_{Input} – исходное множество z особенностей входных данных в виде единого свойства анализируемой характерной черты данных:

$$C_{Input} = \{c_1^{Input} = \{K_{1,1}^{Input}, \dots, K_{1,z}^{Input}\}\};$$

C_{Goal} – исходное множество x особенностей целевых данных в виде единого свойства анализируемой характерной черты данных:

$$C_{Goal} = \{c_1^{Goal} = \{K_{1,1}^{Goal}, \dots, K_{1,x}^{Goal}\}\};$$

$P_{Characteristic}$ – параметры функции $f_{Characteristic}$.

Для рассматриваемых методов распознавания образов в качестве входных данных выступают особенности либо обучающей, либо тестовой выборки. Целевыми являются данные решающих функций в виде соответствующих анализируемых особенностей обучающих материалов. Под искомыми понимаются особенности входных данных для поиска. Одни и те же данные C_{Input} и C_{Goal} могут быть использованы разными алгоритмами обработки материалов для вышеуказанных этапов обучения и классификации тестовых объектов. Работы разных алгоритмов будут отличаться друг от друга типами обрабатываемых независимых характеристик (группами особенностей) для анализируемых особых черт входных и целевых данных. Поэтому необходимо разработать единый способ представления независимых характеристик для разных их видов в рамках множества особенностей анализируемых данных.

Для группирования соответствующих элементов по необходимым v категориям i -го свойства $v_{Feature.i}$ определяют следующие функции:

1) для группировки фактов по параметрам:

$$K_{Data.Parametr.i} = f_{Group.Parametr}(K_{Data.Parametr.Old.p}, v_{Feature.i}, flag), \quad (18)$$

где $K_{Data.Parametr.Old.p}$ – данные, у которых факты сгруппированы по категориям произвольного свойства $v_{Feature.p}$:

$$K_{Data.Parametr.Old.p} = \{k_1^{Data.Feature.Old.p}, \dots, k_s^{Data.Feature.Old.p}\};$$

$flag$ – параметр, принимающий значение True (истина), если факты необходимо группировать по b категориям свойства $v_{Feature.i}$, и False (ложь), когда необходимо объединить все факты в единое множество; $K_{Data.Parametr.i}$ – данные $K_{Data.Parametr.Old.p}$, у которых факты сгруппированы по категориям анализируемого свойства $v_{Feature.i}$:

$$K_{Data.Parametr.i} = \{k_1^{Data.Feature.i}, \dots, k_v^{Data.Feature.i}\};$$

2) для группировки особенностей:

$$C_{Data.Feature.i} = f_{Group.Property}(C_{Data.Feature.Old.p}, v_{Feature.i}, flag), \quad (19)$$

где $C_{Data.Feature.Old.p}$ – множество данных, у которых факты сгруппированы по категориям произвольного свойства $v_{Feature.p}$:

$$C_{Data.Feature.Old.p} = \{K_1^{Data.Feature.Old.p}, \dots, K_v^{Data.Feature.Old.p}\};$$

$C_{Data.Feature.i}$ – множество данных $C_{Data.Feature.Old.p}$, у которых факты сгруппированы по категориям анализируемого свойства $v_{Feature.i}$:

$$C_{Data.Feature.i} = \{K_1^{Data.Feature.i}, \dots, K_z^{Data.Feature.i}\};$$

3) для группировки характеристик:

$$C_{Data.Feature.i} = f_{Group.Characteristic}(C_{Data.Feature.Old.j}, v_{Feature.i}, flag), \quad (20)$$

где $C_{Data.Feature.Old.j}$ – характерная черта данных, у которой особенности сгруппированы по категориям произвольного свойства данных $v_{Feature.Old.j}$:

$$C_{Data.Feature.Old.j} = \{c_1^{Data.Feature.Old.j}, \dots, c_a^{Data.Feature.Old.j}\};$$

$C_{Data.Feature.i}$ – характерная черта данных $C_{Data.Feature.Old.j}$, у которой особенности сгруппированы по b категориям нового свойства данных $v_{Feature.i}$:

$$C_{Data.Feature.i} = \{c_1^{Data.Feature.i}, \dots, c_b^{Data.Feature.i}\};$$

4) для группировки характерных черт (15):

$$d_{Data.Feature.i} = f_{Group.Characteristics}(d_{Data.Feature.Old}, v_{Feature.i}, Flag), \quad (21)$$

где $d_{Data.Feature.Old}$ – множество из z произвольных характерных черт данных, у которых особенности сгруппированы по категориям соответствующих свойств данных $\{v_{Feature.Old.1}, \dots, v_{Feature.Old.z}\}$:

$$d_{Data.Feature.Old} = \{C_1^{Data.Feature.Old.1}, \dots, C_z^{Data.Feature.Old.z}\};$$

$d_{Data.Feature.i}$ – множество из z характерных черт данных, у которых особенности сгруппированы по b категориям анализируемого нового свойства $v_{Feature.i}$:

$$d_{Data.Feature.i} = \{C_1^{Data.Feature.i}, \dots, C_z^{Data.Feature.i}\};$$

$Flag$ – набор параметров для соответствующих z характерных черт $d_{Data.Feature.Old}$:

$$Flag = \{flag_1, \dots, flag_z\}.$$

В основе обработки вышеуказанных функций лежат логические операции:

- если $flag = True$, то элементы группируются по соответствующему k -му отличительному признаку (категории) некоторого анализируемого свойства $v_{Feature.i}$, которое применяемый метод распознавания образов учитывает при определении искомым данных в рамках работы функции $f_{Characteristic}$;

- если $flag = False$, то элементы объединяются в одно множество.

В качестве примера приведем следующие категории данных, позволяющие отличить друг от друга такие однотипные особенности, как источник данных (класс анализируемых объектов, порядковый номер объекта и др.), иерархические уровни, на которые делятся объекты и их свойства (уровни эталонов в методе дробящихся эталонов), порядковый номер этапа обработки данных (например, i -й случайный выбор объекта из обучающего множества).

Таким образом, для представления независимых характеристик анализируемой характерной черты данных C_{Data} в рамках i -го свойства $v_{Feature.i}$ необходимо:

- присвоить каждому p -му факту C_{Data} требуемую категорию $v_p^{Category.Feature.i}$ свойства $v_{Feature.i}$;
- сгруппировать особенности (наборы фактов) характерной черты C_{Data} по b категориям свойства $v_{Feature.i}$.

Важно отметить, что категории свойства присваиваются фактам, а группируются особенности (наборы фактов). Это возможно благодаря тому, что особенности являются минимальным неделимым набором данных для обработки (например, множество K ближайших соседей). Поэтому все факты должны иметь единую категорию при группировании по особенностям.

Представим множество категории свойства $v_{Feature.i}$ в виде набора неповторяющихся целых положительных чисел следующим образом:

$$V_{Values.Feature.i} = \{1, 2, 3, \dots, b\}. \quad (22)$$

Отметим, что произвольная особенность $K_{i,j}^{Data}$ характерной черты C_{Data} может повторяться в нескольких характеристиках разных категорий $\{C_1^{Data.Feature.i}, \dots, C_b^{Data.Feature.i}\}$ одного свойства $v_{Feature.i}$. Например, в рамках этапа обучения метода опорных векторов объекты i -го класса участвуют в построении гиперплоскостей решающих функций для всех комбинаций пар классов с участием i -го класса. Поэтому необходимо для каждого факта данных k_p устанавливать не одну категорию $v_p^{Category.Feature.i}$, а целый набор значений для определения принадлежности его каждой категории $\{1, 2, 3, \dots, b\}$ (22) свойства $v_{Feature.i}$. Это целесообразно делать следующим образом:

$$V_p^{Category.Feature.i} = \{v_{p,1}^{Category.Feature.i}, \dots, v_{p,b}^{Category.Feature.i}\}, \quad (23)$$

где $v_{p,j}^{Category.Feature.i}$ – значение принадлежности k_p факта j -й категории свойства $v_{Feature.i}$: 0 – если не принадлежит, 1 – если принадлежит.

Таким образом, любой метод распознавания образов должен иметь для всякой искомой характерной черты алгоритм, который будет определять для каждого p -го факта входных, целевых и искомым данных необходимые категории для группировки:

- по особенностям в рамках свойства $v_{Feature.j}$:

$$v_p^{Category.Feature.j};$$

- по категориям свойства $v_{Feature.i}$ (23):

$$V_p^{Category.Feature.i}.$$

Формальное представление взаимосвязей между данными

В качестве анализируемых метрических методов распознавания образов были выбраны и рассмотрены следующие шесть алгоритмов:

- метод K ближайших соседей;
- метод потенциальных функций;
- метод эталонов;
- метод коррекционных приращений;
- метод наименьших среднеквадратических ошибок;
- метод наивный Байесовский.

При реализации алгоритмов этапов обучения и классификации для вышеуказанных методов распознавания в рамках функции $f_{Characteristic}$ (17) одни и те же однотипные объекты и их характеристики могут группироваться и обрабатываться по-разному для искомым особенностей входных данных. Так, например, усредненные эталоны соответствующих q классов обрабатываются двумя способами:

1) отдельно друг от друга на этапе обучения:

$$C_1^{Etalon} = \{c_{1,1}^{Etalon} = \{K_{1,1,1}^{Etalon}\}, \dots, c_{1,q}^{Etalon} = \{K_{1,q,1}^{Etalon}\}\},$$

где C_1^{Etalon} – эталоны классов в виде независимых составляющих $\{c_{1,1,1}^{Etalon}, \dots, c_{1,1,q}^{Etalon}\}$ одной характерной черты C_1^{Etalon} ; $K_{1,i,1}^{Etalon} = \{k_{1,i,1,1}^{Etalon}\}$ – представление эталона i -го класса в виде особенности;

$$C_1^{Class} = \{c_{1,1}^{Class} = \{K_{1,1,1}^{Class}\}, \dots, c_{1,q}^{Class} = \{K_{1,q,1}^{Class}\}\},$$

где C_1^{Class} – соответствующие эталонам наборы обучающих объектов одного класса в виде независимых составляющих $\{c_{1,1,1}^{Class}, \dots, c_{1,1,q}^{Class}\}$ одной характерной черты C_1^{Class} ; $K_{1,i,1}^{Class} = \{k_{1,i,1,1}^{Class}, \dots, k_{1,i,1,x}^{Class}\}$ – представление набора обучающих объектов i -го класса в виде особенности;

2) в виде единого набора данных на этапе классификации тестового объекта в рамках поиска минимального расстояния до них от тестового объекта:

$$C_1^{Etalon} = \{c_{1,1}^{Etalon} = \{K_{1,1,1}^{Etalon}, \dots, K_{1,q,1}^{Etalon}\}\},$$

где C_1^{Etalon} – эталоны классов в виде одного свойства $c_{1,1}^{Etalon}$ характерной черты C_1^{Etalon} ;

$$C_1^{Test} = \{c_{1,1}^{Test} = \{K_{1,1,1}^{Test}, \dots, K_{1,w,1}^{Test}\}\},$$

где C_1^{Test} – тестовые объекты в виде одной характеристики $c_{1,1}^{Test}$ репрезентативной черты C_1^{Test} ; $K_{1,i,1}^{Test} = \{k_{1,i,1,1}^{Test}\}$ – представление i -го тестового объекта для классификации в виде особенности.

Чтобы учесть, какие входные данные являются независимыми, а какие – общими, необходимо представить работу алгоритма, основанного на функции $f_{Characteristic}$, как процесс обработки иерархических данных. Для этого представим входные, целевые и искомые закономерности в виде сгруппированных фактов по необходимым свойствам данных следующим образом:

$$d_{Input} = \{C_1^{Input}, \dots, C_z^{Input}\},$$

$$d_{Goal} = \{C_1^{Goal}, \dots, C_z^{Goal}\},$$

$$d_{New} = \{C_1^{New}, \dots, C_z^{Goal}\},$$

где d_{Input} , d_{Goal} , d_{New} – исходные множества анализируемых характерных черт входных, целевых и искоемых данных.

Для вышеуказанных методов распознавания свойственны следующие особенности представления данных d_{Input} , d_{Goal} и d_{New} :

1) множества характерных черт закономерностей d_{Input} , d_{Goal} и d_{New} сгруппированы по z категориям одного анализируемого свойства $v_{Feature.All.Feature}$ на основе функции $f_{Group.Characteristics}$ (21), в рамках которых определяются независимые данные для обработки;

2) множество элементов $\{c_{1,1}^{Input}, \dots, c_{1,a}^{Input}\}$ для каждой характерной черты C_i^{Input} данных d_{Input} сгруппировано по a категориям анализируемого свойства $v_{Feature.Input.Characteristic}$ в рамках функции $f_{Group.Characteristic}$ (20); с помощью числа a категорий свойства $v_{Feature.Input.Characteristic}$ определяется общий или индивидуальный характер искоемых данных относительно входных;

3) множество элементов $\{c_{1,1}^{Goal}, \dots, c_{1,b}^{Goal}\}$ для каждой характерной черты C_i^{Goal} данных d_{Goal} сгруппировано по b категориям анализируемого свойства $v_{Feature.Goal.Characteristic}$; аналогично предыдущей особенности общий или индивидуальный характер искоемых данных относительно целевых данных определяется по категориям свойства $v_{Feature.Goal.Characteristic}$;

4) для каждой характерной черты C_i^{New} данных d_{New} множество искоемых элементов $\{c_{i,1,1}^{New}, \dots, c_{i,1,b}^{Goal}, \dots, c_{i,a,1}^{Goal}, \dots, c_{i,a,b}^{Goal}\}$ сгруппировано по $a * b$ категориям анализируемых свойств $v_{Feature.Input.Characteristic}$ и $v_{Feature.Goal.Characteristic}$;

5) множества особенностей для каждой характеристики закономерностей d_{Input} , d_{Goal} и d_{New} сгруппированы по категориям анализируемого свойства $v_{Feature.Property}$ на основе функции $f_{Group.Property}$ (19) в рамках определения общих и независимых данных для обработки;

6) множества фактов для каждой особенности закономерностей d_{Input} , d_{Goal} и d_{New} сгруппированы по категориям анализируемого свойства $v_{Feature.Parametr}$ на основе функции $f_{Group.Parametr}$ (18) в рамках определения неделимых наборов данных для обработки.

Обозначим произвольную закономерность d в виде структуры дерева и определим для нее следующие обозначения:

– $e_i^{Feature} = d.C_i$ – i -я характерная черта;

– $e_{i,j}^{Characteristic} = d.C_i.C_j$ – j -я характеристика $d.C_i$;

– $e_{i,j,k}^{Property} = d.C_i.C_j.K_k$ – k -я особенность $d.C_i.C_j$;

– $e_{i,j,k,p}^{Parametr} = d.C_i.C_j.K_k.k_p$ – p -й параметр $d.C_i.C_j.K_k$.

Представим все элементы одного вида закономерности d в виде следующих объединений $n_{Parts} = q$ множеств, где в каждом i -м подмножестве $n_i^{Elements} = w$ элементов:

для *Feature*:

$$E_{Feature} = \{e_1^{Feature}, \dots, \cup, \dots, e_z^{Feature}\} = \{E_1^{Feature}, \dots, \cup, \dots, E_q^{Feature}\},$$

для *Characteristic*:

$$E_{Characteristic} = \{e_{1,1}^{Characteristic}, \dots, \cup, \dots, e_{z,s}^{Characteristic}\} = \{E_1^{Characteristic}, \dots, \cup, \dots, E_q^{Characteristic}\},$$

для *Property*:

$$E_{Property} = \{e_{1,1,1}^{Property}, \dots, \cup, \dots, e_{z,s,g}^{Property}\} = \{E_1^{Property}, \dots, \cup, \dots, E_q^{Property}\},$$

для *Parametr*:

$$E_{Parametr} = \{e_{1,1,1,1}^{Parametr}, \dots, \cup, \dots, e_{z,s,g,k}^{Parametr}\} = \{E_1^{Parametr}, \dots, \cup, \dots, E_q^{Parametr}\}.$$

Определим обобщенные функции копирования и разделения множества однотипных элементов для всех уровней иерархии произвольной закономерности d :

$$D = f_{Copy.Elements}(d, n_{Parts}),$$

$$D = f_{Split.Elements}(d, type, n_{Parts}, N_{Elements}), \quad (24)$$

где d – произвольная закономерность для обработки; $n_{Parts} = q$ – число копий закономерности d ; $type = \{Feature, Characteristic, Property, Parametr\}$ – виды элементов, которые функция может обрабатывать; $n_{Parts} = q$ – число частей, на которые разбивается единое множество элементов вида $type$ в рамках d ; $N_{Elements} = \{n_1^{Elements}, \dots, n_q^{Elements}\}$ – число элементов в соответствующих $n_{Parts} = q$ частях; $D = \{d_1, \dots, d_q\}$ – итоговое множество $n_{Parts} = q$ закономерностей, содержащее соответствующие части единого набора разбиваемых элементов.

Функция $f_{Split.Elements}$ разделяет закономерность d на множество $n_{Parts} = q$ неполных частей, которые содержат непересекающиеся данные вида $type$:

$$d_1 \cdot E_{Type}^1, \dots, \cup, \dots, d_q \cdot E_{Type}^q = d \cdot E_{Type}$$

$$d_1 \cdot E_{Type}^1, \dots, \cap, \dots, d_q \cdot E_{Type}^q = \emptyset. \quad (25)$$

Представим обобщенный алгоритм распараллеливания вычисления для вышеуказанных методов распознавания, основанный на использовании функции $f_{Characteristic}$, в виде последовательной обработки иерархических независимых данных следующим образом:

Этап 1. Обработка независимых характерных черт

Все характерные черты в рамках одной категории (входные, целевые и искомые данные) обрабатываются отдельно друг от друга, а характерные черты с одним индексом разных категорий – вместе (рисунок 1, где красная линия отображает направление потока данных для обработки в рамках i -й характерной черты). Разделим с помощью алгоритма, основанного на использовании функции $f_{Split.Elements}$ (24), данные входной, целевой и искомой закономерностей на независимые части вида D (16), состоящих из одной характерной черты, в рамках z категорий свойства $v_{Feature.All.Feature}$ следующим образом:

$$D_{Input-Feature} = \{d_1^{Input-Feature}, \dots, d_z^{Input-Feature}\},$$

где $D_{Input-Feature}$ – набор z частей характерных черт закономерности входных данных d_{Input} ;

$$D_{Goal-Feature} = \{d_1^{Goal-Feature}, \dots, d_z^{Goal-Feature}\},$$

где $D_{Goal-Feature}$ – набор z частей характерных черт закономерности целевых данных d_{Goal} ;

$$D_{New-Feature} = \{d_1^{New-Feature}, \dots, d_z^{New-Feature}\},$$

$D_{New-Feature}$ – набор z частей характерных черт закономерности искомых данных d_{New} .

Этап 2. Обработка независимых характеристик данных

Все характеристики входных и целевых данных в рамках своей категории обрабатываются отдельно друг от друга, а характеристики вида $(c_{i,j}^{Input}, \{c_{i,1}^{Goal}, \dots, c_{i,p}^{Goal}\})$ – вместе (см. рисунок 1, где зелеными линиями показана обработка каждой входной характеристики со всеми целевыми). Пара характеристик вида $(c_{i,j}^{Input}, c_{i,p}^{Goal})$ однозначно определяет искомую характеристику $c_{i,j,p}^{New}$, т.е. число искомых характеристик равно количеству входных, умноженному на число целевых характеристик.

Определим следующие данные для i -й характерной черты, полученные на предыдущем этапе – входные, целевые и искомые:

$$d_i^{Input-Feature} = \{c_{i,1}^{Input}, \dots, c_{i,a}^{Input}\},$$

$$d_i^{Goal-Feature} = \{c_{i,1}^{Goal}, \dots, c_{i,b}^{Goal}\},$$

$$d_i^{New-Feature} = \{c_{1,1,1}^{New}, \dots, c_{1,1,b}^{Goal}, \dots, c_{1,a,1}^{Goal}, \dots, c_{1,a,b}^{Goal}\}.$$

Разделим характеристики на примере входной, целевой и искомой закономерности ($d_i^{Input-Feature}$, $d_i^{Goal-Feature}$ и $d_i^{New-Feature}$) на независимые части данных с помощью $v_{Feature.Input.Characteristic}$ и $v_{Feature.Goal.Characteristic}$ следующим образом:

$$D_i^{Input-Characteristic} = \{d_1^{Input-Characteristic}, \dots, d_a^{Input-Characteristic}\},$$

где $D_i^{Input-Characteristic}$ – набор a закономерностей с одной характеристикой от $d_i^{Input-Feature}$;

$$D_i^{New-Characteristic} = \{d_{i,1}^{New-Characteristic.1}, \dots, d_{i,r}^{New-Characteristic.1}\},$$

где $D_i^{New-Characteristic}$ – набор a закономерностей, содержащих по b характеристик $d_i^{New-Feature}$;

$$D_i^{Goal-Characteristic} = \{d_1^{Goal-Characteristic}, \dots, d_b^{Goal-Characteristic}\},$$

где $D_i^{Goal-Characteristic}$ – набор b закономерностей с одной характеристикой от $d_i^{Goal-Feature}$;

$$D_{i,j}^{New-Characteristic} = \{d_{i,j,1}^{New-Characteristic}, \dots, d_{i,j,b}^{New-Characteristic}\},$$

где $D_{i,j}^{New-Characteristic}$ – набор b закономерностей с одной характеристикой от $d_{i,j}^{New-Characteristic.1}$.

В рамках рассматриваемого этапа независимыми данными для обработки будут закономерности следующего вида:

$$(d_{i,j}^{Input-Characteristic}, d_{i,k}^{Goal-Characteristic}, d_{i,j,k}^{New-Characteristic}).$$

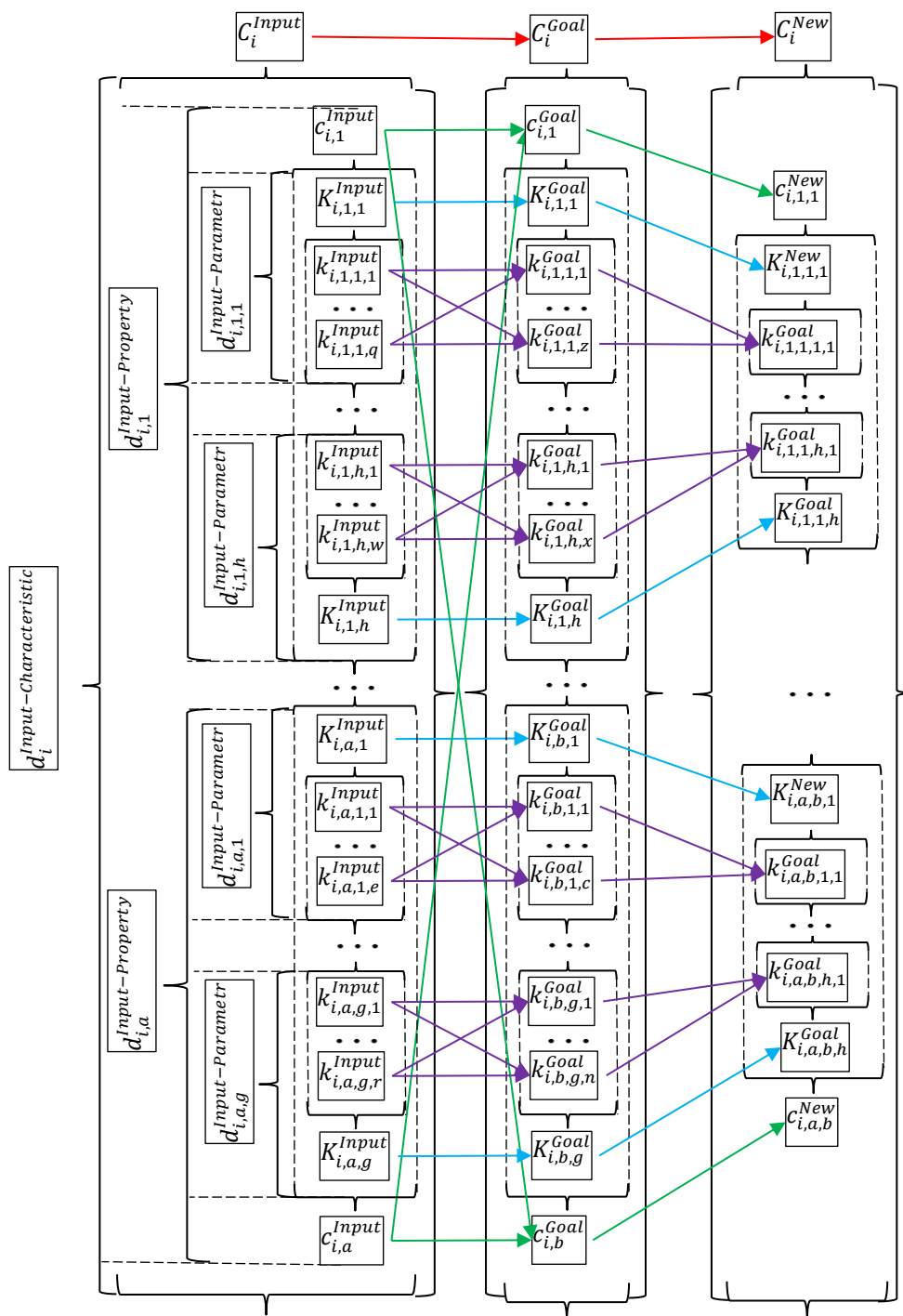


Рис. 1. Общий принцип обработки иерархических данных

Fig. 1. General Principle of Processing Hierarchical Data

Этап 3. Обработка независимых особенностей данных.

Особенности входных, целевых и искомых характеристик в рамках своей категории обрабатываются независимо друг от друга (см. рисунок 1, где синими линиями показаны отдельные обработки входной и целевой характеристик). Разделим данные входной, целевой и искомой закономерностей с помощью категорий свойства $v_{Feature.Property}$ на независимое количество v частей, содержащих по одной особенности, используя алгоритм (25):

$$D_{i,j}^{Input-Property} = \{d_{i,j,1}^{Input-Property}, \dots, d_{i,j,v}^{Input-Property}\},$$

где $D_{i,j}^{Input-Property}$ – набор z частей особенностей закономерности входных данных $d_{i,j}^{Input-Property}$;

$$D_{i,k}^{Goal-Property} = \{d_{i,k,1}^{Goal-Property}, \dots, d_{i,k,v}^{Goal-Property}\},$$

где $D_{i,k}^{Goal-Property}$ – набор z частей особенностей закономерности целевых данных $d_{i,k}^{Goal-Property}$;

$$D_{i,j,k}^{New-Property} = \{d_{i,j,k,1}^{New-Property}, \dots, d_{i,j,k,v}^{New-Property}\},$$

где $D_{i,j,k}^{New-Property}$ – набор z частей особенностей закономерности искомых данных $d_{i,j,k}^{New-Property}$.

Этап 4. Обработка независимых параметров данных.

Каждый параметр входных данных, как и на этапе 2, обрабатывается с целым набором параметров целевых данных в рамках категорий свойства $v_{Feature.Parametr}$ (см. рисунок 1, где фиолетовыми линиями показана общая обработка входных и целевых параметров). Параметры, в отличие от вышеуказанных этапов, необходимо разделять по заданному в них количеству объектов. Первые два этапа отвечают за учет общих и индивидуальных входных данных для целевых данных, принадлежащих разным классам (межклассовые взаимосвязи), а третий и четвертый этапы определяют общие и индивидуальные взаимосвязи между входными и целевыми данными внутри одного класса (внутриклассовые взаимосвязи).

Формальное представление модели решающих функций

В основе разработанной модели лежит принцип повторного использования входных данных для определения множества искомых закономерностей между ними и необходимыми для этого дан-

ными решающих функций. В рамках предлагаемого подхода определим модель решающих функций.

Модель m решающих функции хранит выявленные характерные черты обучающих данных d_{Goal} и искомые d_{New} входных данных в виде соответствующих наборов данных:

$$m = \langle d_{Goal}, d_{New} \rangle, \quad (26)$$

где $d_{Goal} = \{c_1^{Goal}, \dots, c_w^{Goal}\}$ – множество независимых наборов однотипных характеристик обучающей выборки, описывающих необходимые данные для определения заданного набора искомых характерных черт; $d_{New} = \{c_1^{New}, \dots, c_e^{New}\}$ – множество наборов искомых однотипных характеристик входных данных.

С помощью функции $f_{Characteristic}$ обобщим рассматриваемую концепцию для анализируемых метрических методов распознавания образов на основе предложенной модели (26) следующим образом:

$$m_e = f_{Features}(d_{Input}, m_0, P_{Features}), \quad (27)$$

где $m_e = (d_{Goal}, d_e^{New})$ – конечное состояние модели после определения всех e искомых характерных черт данных; $m_0 = (d_{Goal}, d_0^{New})$ – начальное состояние модели до определения всех o искомых характерных черт данных; d_{Input} – исходное множество обрабатываемых характерных черт входных данных:

$$d_{Input} = \{c_1^{Input}, \dots, c_i^{Input} = \{c_{i,1}^{Input} = \{K_{i,1}^{Input}, \dots, K_{i,z}^{Input}\}}, \dots, c_q^{Input}\};$$

d_{Goal} – исходное множество анализируемых характерных черт целевых данных:

$$d_{Goal} = \{c_1^{Goal}, \dots, c_i^{Goal} = \{c_{i,1}^{Goal} = \{K_{i,1}^{Goal}, \dots, K_{i,x}^{Goal}\}}, \dots, c_w^{Goal}\};$$

d_w^{New} – конечное состояние множества искомых характерных черт для входных данных:

$$d_w^{New} = \{c_{1,End}^{New}, \dots, c_{i,End}^{New} = \{c_{i,1}^{New} = \{K_{i,1}^{New}, \dots, K_{i,c}^{New}\}}, \dots, c_{e,End}^{New}\};$$

d_{New} – исходное множество искомых характерных черт для входных данных:

$$d_{New} = \{c_{1,0}^{New}, \dots, c_{i,0}^{New} = \{c_{i,1}^{New} = \{K_{i,1}^{New}, \dots, K_{i,c}^{New}\}}, \dots, c_{e,0}^{New}\};$$

$f_{Features}$ – соответствующая функция поиска искомых закономерностей (свойств, классов, параметров данных) m_w между d_{Input} и m_0 ; $P_{Features}$ – параметры работы функции $f_{Features}$.

Так как искомые характерные черты $\{c_1^{New}, \dots, c_e^{New}\}$ входных данных определяются последовательно, то процесс реализации алгоритма, использующего функцию $f_{Features}$ (27), представим в виде набора шагов по определению соответствующей h -й искомой характерной черты C_h^{New} :

$$m_h = f_h^{Feature}(d_{Input}, m_{h-1}, P_{Features}), \quad (28)$$

где $f_h^{Feature}$ – соответствующая функция поиска искомой особенности C_h^{New} между входными данными d_{Input} и текущими выявленными характерными чертами входных и обучающих данных m_{h-1} .

Выделим и рассмотрим следующие особенности этого алгоритма (28).

Во-первых, множества d_{Input} и d_{Goal} описывают достаточное количество данных для определения всех искомым характерных черт d_{New} . Однако для поиска h -й искомой особенности C_h^{New} не всегда требуется анализировать все характерные черты входных d_{Input} и целевых d_{Goal} данных.

Во-вторых, для определения h -й искомой характерной черты C_h^{New} функция $f_h^{Feature}$ также может использовать полученные на предыдущих этапах искомые особенности $\{c_1^{New}, \dots, c_{h-1}^{New}\}$ в качестве целевых данных.

Анализ вышеуказанных особенностей показывает, что для корректной работы алгоритма необ-

ходимо разработать способ представления данных, который позволит определять, какие характерные черты входных и целевых данных участвуют в получении искомой особенности C_h^{New} , а также – выявлять общие характерные черты для особенности C_h^{New} , которые необходимо обрабатывать целиком, без разбиения.

В интересах решения рассматриваемой проблемы дополнительно введем в набор значений категорий каждого свойства следующие технические индексы:

- индекс 0, отвечающий за отсутствие какой-либо связи между заданной характерной чертой и анализируемым свойством данных, т. е. характерная черта не участвует в обработке функции $f_h^{Features}$ для получения особенности C_h^{New} ;
- индекс All для группы однотипных особенностей, показывающий, что набор данных является общим для всех категорий свойства, и его необходимо обрабатывать целиком без разбиения.

Пусть метод распознавания образов в рамках модели m анализирует e различных искомым характерных черт данных с помощью функций вида $f_h^{Features}$. Сформируем в обобщенном виде для каждого p -го факта данных k_p собственный набор значений категорий соответствующих свойств (29) для всех e функции поиска искомым данных (30).

$$V_{p,h}^{Category.Feature.All} = \{V_{p,h}^{Category.Feature.All.Feature}, V_{p,h}^{Category.Feature.Input.Characteristic}, V_{p,h}^{Category.Feature.Goal.Characteristic}, V_{p,h}^{Category.Feature.Property}, V_{p,h}^{Category.Feature.Parametr}\}. \quad (29)$$

$$V_p^{Category.All.Feature.All} = \{V_{p,1}^{Category.Feature.All}, \dots, V_{p,e}^{Category.Feature.All}\}. \quad (30)$$

Операции над элементами модели знаний для параллельной обработки

Основным подходом к распараллеливанию вычислений для вышеуказанных методов распознавания образов является разделение данных на части по заданному в них количеству объектов. Тем не менее, как было показано выше, характерная черта может состоять как из общих, так и независимых частей данных. Таким образом, при распараллеливании по объектам необходимо учитывать принадлежность каждого объекта своей группе. Например, при формировании независимой части из $n * t$ объектов одного класса в рамках этапа обучения метода эталонов на вычислительный узел необходимо передать только усредненный эталон соответствующего класса. А для случая, когда формируется n частей разных классов по t объектов, необходимо передать n усредненных эталонов анализируемых классов. Таким образом,

объемы вычислений для двух рассматриваемых случаев равны, а объемы передаваемых данных – нет.

Также отметим, что на объем передаваемых данных влияет вышеуказанная возможность дублирования одной особенности для нескольких характеристик. Поэтому необходимо обеспечить распараллеливание вычислений по объектам без учета их зависимостей между собой, влияющих на объем передаваемых данных. Для решения этой задачи предлагается использовать подход, заключающийся в следующем:

- 1) представить все данные каждой категории (входной, целевой и искомой) на всех уровнях в виде непересекающихся подмножеств данных;
- 2) распараллеливать вычисления по независимым группам особенностей с целью получения только общих входных и целевых данных.

При этом, если отсутствует возможность разбиения данных на непересекающиеся части, особенности должны быть представлены в виде единого набора данных.

Так, например, в рамках операции построения гиперплоскостей:

$$C_{Goal} = \{c_{1,1,2}^{Hyper}, \dots, c_{1,n-1,n}^{Test}\}$$

для всех комбинации пар n классов метода опорных векторов множество наборов сгруппированных по классам обучающих объектов:

$$C_{Input} = \{K_{1,1}^{Class}, \dots, c_{1,n}^{Class}\}$$

невозможно разделить на непересекающиеся части относительно гиперплоскостей соответствующих решающих функций (рисунок 2а). Поэтому данные одного из множеств C_{Input} или C_{Goal} должны быть представлены в виде одной характеристики (рисунок 2б).

На практике количество целевых данных меньше, чем входных, поэтому целесообразно ограничиваться передачей только целевых данных, а для входных и искомых использовать распараллеливание, т. е. необходимо исключить учет категорий свойств $v_1^{Feature.Goal.Property}$ и $v_1^{Feature.Characteristic}$ из распараллеливания.

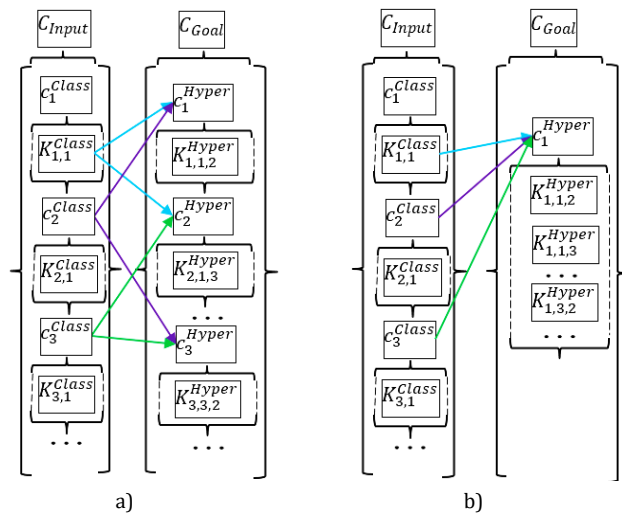


Рис. 2. Виды зависимостей между входными и целевыми данными метода опорных векторов: а) пересекающиеся; б) непересекающиеся

Fig. 2. Types of Dependencies between Input and Target Data of the Support Vector Machine: a) Intersecting; b) Non-Intersecting

Выделим два этапа предлагаемого подхода к распараллеливанию вычислений по объектам без учета их зависимостей между собой.

Этап 1. Распараллеливание вычислений на основе разделения входных, целевых и искомых данных на независимые части в рамках предложенного разбиения для последовательной обработки.

Этап 2. Распараллеливание вычислений в рамках разделения входных и целевых данных на части по заданному в них количеству объектов.

Рассмотрим выделенные этапы предлагаемого подхода к распараллеливанию вычислений на примере метода потенциальных функций более подробно (рисунок 3).

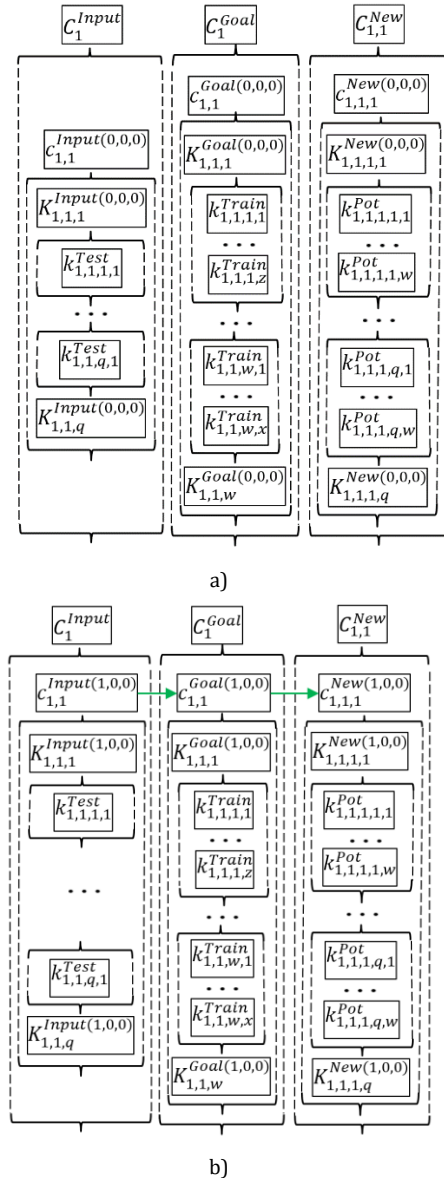


Рис. 3. Этапы распараллеливания метода потенциальных функций: а) этап инициализации данных; б) этап распараллеливания по свойству $v_1^{Feature.All.Feature}$

Fig. 3. Parallelization Stages of the Potential Function Method: a) Initialization Stage; b) Parallelization Stage by Property $v_1^{Feature.All.Feature}$

Входными данными для метода потенциальных функций являются следующие элементы (см. рисунок 3а):

$$d_{Input} = \left\{ C_1^{Input} = \left\{ c_{1,1}^{Input} = \left\{ K_{1,1,1}^{Input} = \right. \right. \right. \\ \left. \left. \left. = \left\{ k_{1,1,1,1}^{Test} \right\}, \dots, K_{1,1,q}^{Input} = \left\{ k_{1,1,q,1}^{Test} \right\} \right\} \right\} \right\},$$

d_{Input} – исходное множество q объектов для классификации, где $k_{1,1,i,1}^{Test}$ – i -й тестовый объект;

$$m_0 = (d_{Goal}, d_0^{New}),$$

где m_0 – начальное состояние модели до определения всех q классов тестовых объектов d_{Input} ;

$$d_{Goal} = \left\{ C_1^{Goal} = \left\{ c_{1,1}^{Goal} = \left\{ K_{1,1,1}^{Goal} = \right. \right. \right. \\ \left. \left. \left. = \left\{ k_{1,1,1,1}^{Train}, \dots, k_{1,1,1,z}^{Train} \right\}, \dots, K_{1,1,w}^{Goal} = \left\{ k_{1,1,w,1}^{Train}, \dots, k_{1,1,w,x}^{Train} \right\} \right\} \right\} \right\},$$

где d_{Goal} – исходное множество наборов обучающих объектов одного класса для w соответствующих классов;

$$d_0^{New} = \left\{ C_{1,1}^{New} = \left\{ c_{1,1,1}^{New} = \left\{ K_{1,1,1,1}^{New}, \dots, K_{1,1,1,i}^{New} = \right. \right. \right. \\ \left. \left. \left. = \left\{ k_{1,1,1,i,1}^{Pot}, \dots, k_{1,1,1,i,w}^{Pot} \right\}, \dots, K_{1,1,1,q}^{New} \right\} \right\} \right\},$$

где d_0^{New} – исходное множество наборов значений потенциалов w классов для соответствующих q тестовых объектов.

В рамках метода потенциальных функции входные, целевые и искомые данные содержат только по одной характерной черте данных, поэтому функция вида $f_h^{Feature}$ тоже одна: $f_1^{Feature}$. Как следствие, данные будут группироваться в рамках независимых множеств элементов только по следующим пяти свойствам $V_{p,h}^{Category.Feature.All}$ (29):

- $v_1^{Feature.All.Feature}$;
- $v_1^{Feature.Input.Characteristic}$;
- $v_1^{Feature.Goal.Characteristic}$;
- $v_1^{Feature.Property}$;
- $v_1^{Feature.Parametr}$.

Для простоты изложения последующего материала добавим индексы первых трех свойств к элементам обрабатываемых данных:

$$K_i^{(IG,IN,GN)} c_i^{(IG,IN,GN)} \text{ и } C_i^{(IG,IN,GN)},$$

где IG – индекс свойства $v_1^{Feature.All.Feature}$; IN – индекс свойства $v_1^{Feature.Input.Characteristic}$; GN – индекс свойства $v_1^{Feature.Goal.Characteristic}$.

Этап обработки независимых характерных черт для распараллеливания отличается от вышеуказанного этапа последовательной обработки тем, что необходимо представлять характерные черты в виде непересекающихся групп данных. Если это невозможно осуществить, целевые и искомые данные передаются целиком, а входные распараллеливаются. А в случае возможности представления данных в виде непересекающихся

групп допускается передавать искомые данные целиком. Тогда входные и целевые данные становятся общими относительно друг друга и, как следствие, можно сразу перейти к распараллеливанию по объектам.

Метод потенциальных функций имеет всего одну характерную черту в каждом наборе данных. Таким образом, выделим следующие особенности распараллеливания по $v_1^{Feature.All.Feature}$ (рисунок 3b).

Во-первых, тестовые объекты (входные данные) в рамках метода потенциальных функций являются общими для обучающих объектов (целевые данные), сгруппированным по классам, поэтому существует только одна индивидуальная категория для свойства $v_1^{Feature.All.Feature} = [1]$, и у всех особенностей обучающих данных первый индекс в скобках равен 1.

Во-вторых, аналогично входным данным у всех особенностей целевых данных первый индекс в скобках тоже равен 1.

В-третьих, аналогично входным и целевым данным у всех особенностей искомых данных первый индекс в скобках тоже равен 1.

Этап обработки независимых характеристик данных по категориям свойств $v_1^{Feature.Input.Characteristic}$ и $v_1^{Feature.Goal.Characteristic}$ отличается от соответствующего этапа последовательной обработки той же спецификой, что и этап обработки характерных черт.

Выделим следующие особенности распараллеливания по $v_1^{Feature.Input.Characteristic}$ (рисунок 4a).

Во-первых, входные данные содержат независимые друг от друга характеристики относительно искомых, поэтому существует q категорий свойства $v_1^{Feature.Input.Characteristic}$. Таким образом, входные данные разделяются на q групп, и второй индекс равен номеру соответствующей группы.

Во-вторых, целевые данные не участвуют в группировке по свойству $v_1^{Feature.Input.Characteristic}$, поэтому в рамках обобщения для всех целевых особенностей второй индекс равен All .

В-третьих, вторым индексам искомых данных аналогично входным также было присвоено соответствующее q значение.

Сформулируем особенности распараллеливания по $v_1^{Feature.Goal.Characteristic}$ (рисунок 4b).

Во-первых, входные данные не участвуют в группировке по свойству $v_1^{Feature.Goal.Characteristic}$, поэтому в рамках обобщения третий индекс всех входных особенностей равен техническому индексу All .

Во-вторых, целевые данные содержат независимые друг от друга множества обучающих объек-

тов для w классов ($v_1^{Feature.Goal.Characteristic} = \{1, \dots, w\}$). Таким образом, целевые данные разделяются на w групп и второй индекс равен номеру соответствующей группы.

Кроме того, третьим индексам искомым данных аналогично целевым также было присвоено соответствующее w значение.

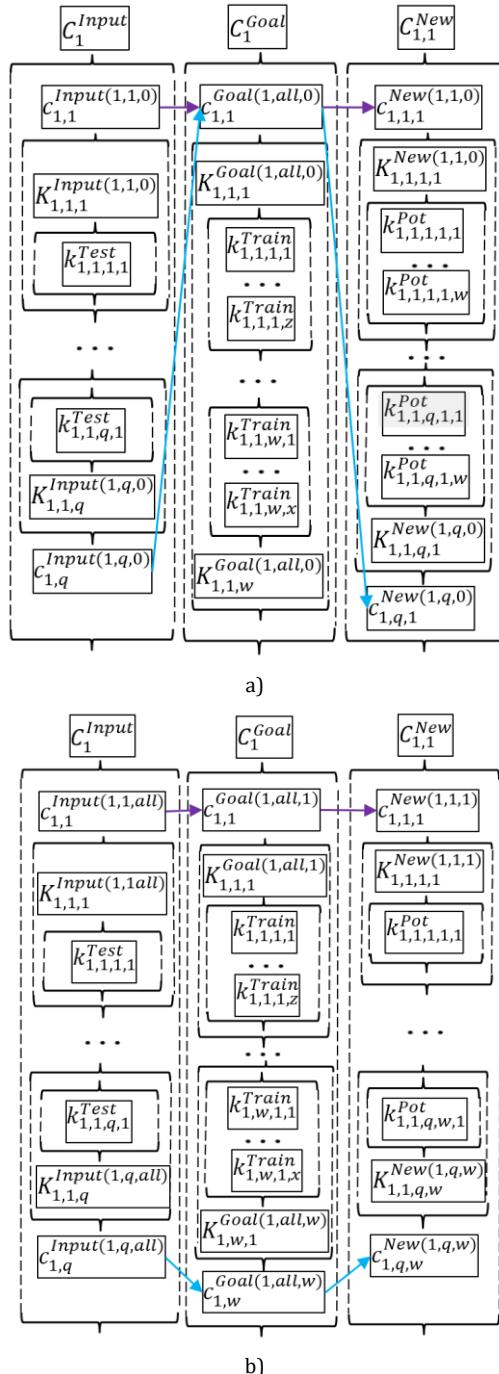


Рис. 4. Этапы распараллеливания метода потенциальных функций: а) по свойству $v_1^{Feature.Input.Characteristic}$; б) по свойству $v_1^{Feature.Goal.Characteristic}$

Fig. 4. Parallelization Stages of the Potential Function Method: а) by Property $v_1^{Feature.Input.Characteristic}$; б) by property $v_1^{Feature.Goal.Characteristic}$

Этап обработки независимых особенностей данных: в рамках вышеуказанных ограничений на непересекаемость групп данных между собой их независимые особенности могут быть либо общими друг для друга, либо индивидуальными в рамках одинакового количества особенностей в характеристиках. Для метода потенциальных функций рассматриваемый этап отсутствует, так как в каждой характеристике находится только одна особенность (свойство $v_1^{Feature.Property}$).

Этап обработки независимых параметров данных с помощью $v_1^{Feature.Parametr}$: в силу неделимости фактов некоторых параметров получить точное число элементов в каждой части не всегда является возможным. Например, в рамках метода K ближайших соседей число фактов в каждой независимой части целевых данных в обобщенном виде всегда будет кратно числу k .

Проанализировав особенности распараллеливания вычислений в рамках этапов последовательной обработки независимых данных, определим на основе ранее разработанных функций обработки закономерностей следующие обобщенные функции:

1) функция копирования модели:

$$M = f_{Copy.Models}(m, n_{parts}) = \{m_1, \dots, m_q\}, \quad (31)$$

где $n_{parts} = q$ – число копий модели m ; $M = \{m_1, \dots, m_q\}$ – выходное множество моделей;

2) функция разделения модели:

$$M = f_{Split.Models}(m, type, n_{parts}, N_{Elements}) = \{m_1 \cdot E_{type}^1, \dots, m_q \cdot E_{type}^q\}, \quad (32)$$

$type = \{Feature, Characteristic, Property, Parametr\}$ – виды элементов, которые функция может обрабатывать; $n_{parts} = q$ – число частей, на которые разбивается единое множество элементов вида $type$ в рамках m ; $N_{Elements} = \{n_1^{Elements}, \dots, n_q^{Elements}\}$ – число элементов в соответствующих $n_{parts} = q$ частях;

3) функция агрегирования набора моделей в единую модель:

$$m = f_{Union.Models}(M, type) = \{m \cdot \{d_1 \cdot E_{type}^1, \dots, \oplus, \dots, d_q \cdot E_{type}^q\}\}, \quad (33)$$

где $\oplus$ – операция агрегирования для соответствующих элементов вида $type$.

Разработанные функции (31–33) обеспечивают распараллеливание вычислений процесса классификации объекта как по входным данным, так и по данным решающих функций.

Результаты эксперимента по распараллеливанию метода потенциальных функций

Алгоритм распараллеливания как по входным данным, так и по данным предложенной модели решающих функций метода потенциальных функций был протестирован на различных частях данных задачи *Customer Churn DataSet* [15]. Задача заключается в разделении множества тестируемых клиентов (объектов) на $z = 2$ класса (останется или уйдет клиент), где каждый классифицируемый объект имеет $e = 10$ атрибутов (столбцов) и идентификационный номер. Три атрибута имеют строковый тип с ограниченным числом принимаемых значений вида $\{a_1, a_2, \dots, a_q\}$ и поэтому были приведены к числовому типу на основе следующего принципа: $a_1 = 1, \dots, a_q = q$.

Определим следующие параметры:

c – количество компьютеров (клиентов);

TN_i – время передачи i -ых входных данных D_i и модели решающих функций M_i по сети между сервером и i -ым клиентом;

TP_i – время построения i -й модели M_i с учетом входных данных D_i на i -м клиенте.

В качестве критерия эффективности выступает усредненный коэффициент полезного действия (КПД) построения одной части модели решающих функций на c компьютерах с учетом времени передачи данных по сети:

$$\text{КПД} = \frac{1}{c} * \frac{\sum_{i=1}^c TP_i}{\sum_{i=1}^c (TP_i + TN_i)}. \quad (34)$$

В качестве вычислительного ресурса выступает Grid-система с переменным количеством клиентов (от 1 до 10) и одним сервером, которые имеют следующие одинаковые аппаратные характеристики: процессор *Intel(R) Core (TM) i3-4130 CPU 3.40 GHz*, ОЗУ 4 Гб.

Входные параметры данных, алгоритма, конфигурации Grid-системы, на основе которых тестировалось программное обеспечение.

1) Число c клиентов в Grid-системе менялось следующим образом: 1, 3, 5, 7, 10.

2) Выборка обучающих объектов оставалась неизменной и содержала $w = 3000$ объектов (строк), каждый ее элемент имел по $e = 10$ атрибутов (столбцов) плюс один столбец класса.

3) Данные модели решающих функций метода потенциальных функций в виде набора обучающих объектов из w элементов распараллеливались по соответствующему числу используемых клиентов на следующие количества частей: 1 ($w = 3000$); 3 ($w = 1000$); 5 ($w = 600$); 7 ($w = 428$); 10 ($w = 300$).

4) Количество тестируемых объектов для распознавания в рамках распараллеливания по вход-

ным данным изменялось следующим образом: 1000, 2000, 3000, 4000, 5000.

5) Число классов равняется двум, что не оказывает существенного влияния на время передачи и обработки данных, и поэтому в рамках эксперимента будут учитываться только зависимости времени обработки и передачи данных от количества обучающих и тестируемых объектов.

Результаты эксперимента представлены в таблице 1, а графическое представление ее данных – на рисунке 5.

Обобщенные выводы

Во-первых, наличие только одного клиента ($c = 1$) отвечает за случай, когда распараллеливаются входные данные, а не модель решающих функций.

Во-вторых, распараллеливание предложенной модели решающих функций позволяет создавать подзадачи меньшего размера и задействовать дополнительные вычислительные ресурсы с целью минимизации времени обработки одного объекта относительно случая, когда распараллеливаются вычисления только по входным данным.

В-третьих, уменьшение размера модели решающих функций приводит к уменьшению КПД обработки, так как время передачи данных по сети прямо пропорционально количеству передаваемых обучающих и тестовых объектов, а время построения модели зависит от произведения числа обучающих объектов и количества тестовых объектов (поиск расстояний производится от каждого из них до всех обучающих). Таким образом, существует нелинейная зависимость КПД обработки от количества обучающих и тестовых объектов.

В-четвертых, в случаях наличия ограничений объема памяти компьютеров для хранения данных решающих функций или времени горизонта планирования возможность распараллеливания предложенной модели решающих функций позволяет более эффективно использовать вычислительные ресурсы за счет учета вышеуказанной нелинейной зависимости КПД от размеров обрабатываемых данных.

ТАБЛИЦА 1. Результаты КПД построения одной модели решающих функций для разного числа клиентов

TABLE 1. Efficiency Results on Building one Decision Functions Model for a Different Number of Clients

	$c = 1,$ $w = 3000$	$c = 3,$ $w = 1000$	$c = 5,$ $w = 600$	$c = 7,$ $w = 428$	$c = 10,$ $w = 300$
$q = 1000$	88,2	83,3	78,9	74,9	69,7
$q = 2000$	92,3	86,9	82,1	77,9	72,2
$q = 3000$	93,7	88,2	83,3	78,9	73,1
$q = 4000$	94,4	88,8	83,9	79,4	73,6
$q = 5000$	94,9	89,2	84,2	79,7	73,8

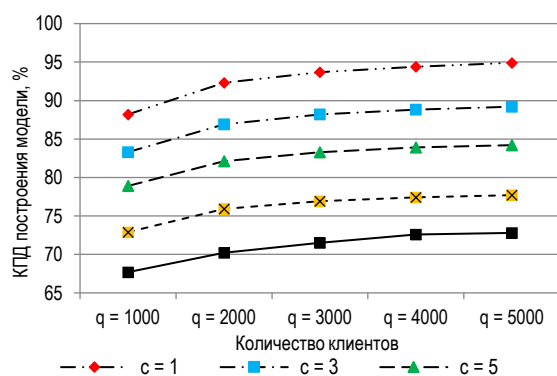


Рис. 5. Результаты КПД построения одной модели решающих функций для разного числа клиентов

Fig. 5. Efficiency Results on Building one Decision Functions Model for a Different Number of Clients

Заключение

В результате проведенных исследований была разработана унифицированная модель решающих функций для алгоритмов машинного обучения и функции, обеспечивающей ее распараллеливание как по входным данным, так и по данным решаю-

щих функций. Приведены примеры, показывающие, что предложенный алгоритм распараллеливания вычислений на основе вышеуказанной модели повышает эффективность классификации как для одного объекта, так и для группы объектов.

Научная новизна предложенного подхода заключается в том, что модель описывает объекты и их свойства единым способом в виде унифицированных элементов как для обучающих, так и классификационных данных. Кроме того, модель имеет структуру и функции, ориентированные на ее параллельную обработку методами распознавания образов на основе решающих функций в рамках группового распараллеливания объектов. Это позволяет осуществить классификацию объекта в режиме реального времени.

Предложенная модель носит также унифицированный характер и применима при распараллеливании других методов распознавания образов, которые могут быть описаны сходными параметрами, архитектурой и классификационными признаками.

Список источников

1. Балашова И.В., Терещенко Т.А. Системы поддержки принятия решений // The Scientific Heritage. 2021. № 79-4(79). С. 3–7. DOI:10.24412/9215-0365-2021-79-4-3-7. EDN:VBGSEN
2. Конев К.А. Машинное обучение для поддержки принятия решений в сфере качества на промышленном предприятии // Экономика. Информатика. 2023. Т. 50. № 3. С. 689–703. DOI:10.52575/2687-0932-2023-50-3-689-703. EDN:EOVNCU
3. Макшанов А.В., Журавлев А.Е. Технологии интеллектуального анализа данных: учебное пособие. СПб.: Лань, 2019. 212 с.
4. Замятин А.В. Интеллектуальный анализ данных: учебное пособие. Томск: Национальный исследовательский Томский государственный университет, 2020. 196 с. EDN:RODEPA
5. Платонов А.В. Машинное обучение: учебное пособие для вузов. М.: Издательство Юрайт, 2025. 89 с.
6. Лимановская О.В., Алферьева Т.И. Основы машинного обучения: учебное пособие. Екатеринбург: Уральский федеральный университет имени первого Президента России Б.Н. Ельцина, 2020. 88 с. EDN:YLJQGT
7. Фомин В.В., Миклуш В.А. Интеллектуальные информационные системы: учебное пособие. СПб.: Российский государственный гидрометеорологический университет, 2013. 150 с. EDN:UPSHTP
8. Кочубей Р.И., Бычковский М.М., Зайкин Н.Н., Фатьянова Е.В., Свидло А.В., Чуприков О.В. Обоснование метода распознавания для решения задачи контроля содержимого данных о состоянии техники связи специального назначения // Известия Тульского государственного университета. Технические науки. 2023. № 12. С. 533–535. DOI:10.24412/2071-6168-2023-12-533-534. EDN:HVOKOQ
9. Дюк В.А., Малыгин И.Г., Прицкер В.И. Распознавание транспортных средств по силуэтам – трехкаскадный метод машинного обучения в системах технического зрения // Морские интеллектуальные технологии. 2022. № 2-1(56). С. 162–167. DOI:10.37220/MIT.2022.56.2.022. EDN:WQMNAQ
10. Зайнидинов Х.Н., Каршиев З.А. Особенности параллельного выполнения алгоритмов интеллектуального анализа данных // Автоматика и программная инженерия. 2020. № 1(31). С. 83–91. EDN:YRWQDE
11. Холод И.И. Модели и методы построения параллельных алгоритмов анализа распределенных данных. Дис. ... докт. техн. наук. СПб.: Санкт-Петербургский государственный электротехнический университет "ЛЭТИ" им. В.И. Ульянова (Ленина), 2018. 351 с. EDN:HKSFMU
12. Kholod I.I., Shorov A.V. Unification of mining model for parallel processing // Proceedings of the Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus, St. Petersburg and Moscow, Russian Federation, 01–03 February 2017). IEEE, 2017. PP. 450–455. DOI:10.1109/EIConRus.2017.7910588. EDN:XMZHNX
13. Холод И.И. Метод определения возможностей параллельного выполнения функций алгоритмов анализа данных // Программные продукты и системы. 2018. № 2. С. 268–274. DOI:10.15827/0236-235X.122.268-274. EDN:UYSUBG
14. Bishop C. Pattern Recognition and Machine Learning. Berlin: Springer, 2006. 738 p.
15. Customer Churn DataSet // Kaggle. URL: <https://www.kaggle.com/datasets/muhammadshahidazeem/customer-churn-dataset> (Accessed 10.04.2025)

References

1. Balashova I., Tereshchenko T. Decision Support Systems. *The Scientific Heritage*. 2021;79-4(79):3–7. (in Russ.) DOI:10.24412/9215-0365-2021-79-4-3-7. EDN:VBGSEN
2. Konev K.A. Machine Learning for Quality Decision Support in an Industrial Enterprise. *Economics. Information technologies*. 2023;50(3):689–703. (in Russ.) DOI:10.52575/2687-0932-2023-50-3-689-703. EDN:EOVNCU
3. Makshanov A.V., Zhuravlev A.E. *Technologies of Intellectual Data Analysis*. St. Petersburg: Lan' Publ.; 2019. 212 p. (in Russ.)
4. Zamyatin A.V. *Intelligent Data Analysis*. Tomsk: National Research Tomsk State University Publ.; 2020. 196 p. (in Russ.) EDN:RODEPA
5. Platonov A.V. *Machine Learning*. Moscow: Yurajt Publ.; 2025. 89 p. (in Russ.)
6. Limanovskaya O.V., Alferyeva T.I. *Basics of Machine Learning*. Ekaterinburg: Ural Federal University Publ.; 2020. 88 p. (in Russ.) EDN:YLJQGT
7. Fomin V.V., Miklush V.A. *Intelligent Information Systems*. St. Petersburg: Russian State Hydrometeorological University Publ.; 2013. 150 p. (in Russ.) EDN:UPSHTP
8. Kochubey R.I., Bychkovsky M.M., Zaikin N.N., Fatyanova E.V., Svidlo A.V., Chuprikov O.V. Justification of the Recognition Method for Solving the Problem of Monitoring the Content of Status Data Special Purpose Communication Technicians. *Izvestiya Tula State University*. 2023;12:533–535. (in Russ.) DOI:10.24412/2071-6168-2023-12-533-534. EDN:HVOKOQ
9. Duke V.A., Malygin I.G., Pritsker V.I. Vehicle Recognition by Silhouettes – A Three-Stage Machine Learning method in Computer Vision Systems. *Marine Intelligent Technologies*. 2022;2-1(56):162–167. (in Russ.) DOI:10.37220/MIT.2022.56.2.022. EDN:WQMNAQ
10. Zaynidinov X.N., Karshiev Z.A. Features of Parallel Execution of Data Mining Algorithms. *Automatics & Software Enginery*. 2020;1(31):83–91. (in Russ.) EDN:YRWQDE
11. Kholod I.I. *Models and Methods of Building Parallel Algorithms for Analyzing Distributed Data*. D.Sc Thesis. St. Petersburg: Saint Petersburg Electrotechnical University "LETI" Publ.; 2018. 351 p. (in Russ.) EDN:HKSFMU
12. Kholod I.I., Shorov A.V. Unification of mining model for parallel processing. *Proceedings of the Conference of Russian Young Researchers in Electrical and Electronic Engineering, EIConRus, 01–03 February 2017, St. Petersburg and Moscow, Russian Federation*. IEEE; 2017. p.450–455. DOI:10.1109/EIConRus.2017.7910588. EDN:XMZHNX
13. Kholod I.I. A Method for Determining the Capabilities of Parallel Execution of Data Mining Algorithm Functions. *Software & Systems*. 2018;2:268–274. (in Russ.) DOI:10.15827/0236-235X.122.268-274. EDN:UYSUBG
14. Bishop C. *Pattern Recognition and Machine Learning*. Berlin: Springer, 2006. 738 p.
15. Kaggle. Customer Churn DataSet. URL: <https://www.kaggle.com/datasets/muhammadshahidazeem/customer-churn-dataset> [Accessed 10.04.2025]

Статья поступила в редакцию 20.02.2025; одобрена после рецензирования 03.04.2025; принята к публикации 10.04.2025.

The article was submitted 20.02.2025; approved after reviewing 03.04.2025; accepted for publication 10.04.2025.

Информация об авторах:

**АЛЕКСАНДРОВ
Игорь Васильевич**

аспирант кафедры программной инженерии и вычислительной техники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-6258-8191>

**ВИВЧАРЬ
Роман Михайлович**

кандидат технических наук, доцент кафедры программной инженерии и вычислительной техники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0003-3865-9102>

**КИРИЧЕК
Руслан Валентинович**

доктор технических наук, профессор, ректор Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-8781-6840>

Киричек Р.В. является главным редактором журнала «Труды учебных заведений связи» с 2023 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.

Kirichek R.V. has been an editor-in-chief of the journal "Proceedings of Telecommunication Universities" since 2023, but has nothing to do with the decision to publish this article. The article has passed the review procedure accepted in the journal. The authors have not declared any other conflicts of interest.

Научная статья

УДК 004.056

<https://doi.org/10.31854/1813-324X-2025-11-2-101-108>

EDN:DWRJHM



Реализация электронной подписи ECC в ограниченных средах

Наурап Хуссейн Сабри, nawrashussein@mail.ru

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация

Аннотация

Обеспечение безопасных операций криптографии в средах с ограниченными ресурсами представляет собой сложную задачу из-за ограниченной вычислительной мощности и памяти. В условиях стремительного роста беспилотных транспортных систем возрастает потребность в эффективных и безопасных криптографических решениях. Оптимизация криптографических алгоритмов для таких систем становится особенно актуальной с учетом их ограниченных вычислительных ресурсов и высоких требований к безопасности.

Целью данного исследования является оптимизация операций электронной подписи на основе эллиптической кривой (ECC) для систем с ограниченными ресурсами, в частности для беспилотных транспортных систем. Исследование направлено на повышение вычислительной эффективности и снижение использования памяти, делая механизмы безопасности на основе ECC более подходящими для встроенных приложений.

Новизна данного исследования заключается в интеграции множества методов оптимизации. Улучшается скалярное умножение точки, используя свойства циклической группы, противоположного числа, а также усовершенствованный оконный метод умножения. Кроме того, вводится детерминированный метод генерации одноразового используемого числа (попс), вдохновленный EdDSA, для дальнейшего повышения эффективности цифровой подписи. Эти оптимизации в совокупности способствуют более эффективному криптографическому процессу, подходящему для сред с ограниченными ресурсами.

Теоретическая значимость заключается в разработке нового математического аппарата, позволяющего оптимизировать операции электронной подписи.

Практическая значимость данного исследования заключается в его применимости в маломощных встраиваемых системах, где вычислительные ресурсы и память крайне ограничены. Оптимизируя операции ECC, это исследование повышает безопасность и производительность криптографических реализаций в беспилотных транспортных системах и аналогичных встраиваемых приложениях, обеспечивая безопасную связь без превышения аппаратных ограничений.

Реализация предложенного метода была осуществлена на микроконтроллере ATmega 2560, **полученные результаты показывают** сокращения количества циклов на 54,1 % и уменьшения использования SRAM на 72,6 % при генерации ключей, а также значительного повышения производительности в процессах подписи и проверки. Экспериментальные результаты подтверждают его эффективность в оптимизации операций ECC для ограниченных устройств беспилотных транспортных систем.

Ключевые слова: электронная подпись, умножение точек эллиптической кривой, устройства с ограниченными ресурсами, беспилотные транспортные системы, микроконтроллер

Финансирование. Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации № 075-00003-24-01 от 08.02.2024 (проект FSEE-2024-0003).

Ссылка для цитирования: Сабри Н.Х. Реализация электронной подписи ECC в ограниченных средах // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 101–108. DOI:10.31854/1813-324X-2025-11-2-101-108. EDN:DWRJHM

Original research

<https://doi.org/10.31854/1813-324X-2025-11-2-101-108>

EDN:DWRJHM

Implementation of an ECC Digital Signature Technique in Constrained Environments

 Nawras H. Sabbry, nawrashussein@mail.ru

Saint Petersburg Electrotechnical University,
St. Petersburg, 197022, Russian Federation

Annotation

Relevance. Ensuring secure cryptographic operations in resource-constrained environments presents challenges due to limited computational power and memory. With the rapid growth of Unmanned Vehicle Systems, the need for efficient and secure cryptographic solutions is increasing. Optimizing cryptographic algorithms for such systems becomes especially relevant given their limited computational resources and high security demands.

The **purpose** of this study is to optimize Elliptic Curve Cryptography (ECC) digital signature operations for resource-limited systems, particularly in unmanned vehicle systems. The research aims to enhance computational efficiency and reduce memory usage, making ECC-based security mechanisms more feasible for embedded applications.

The **novelty** of this study lies in its integration of multiple optimization techniques. It improves scalar point multiplication by leveraging cyclic group properties, additive inverses, and an enhanced windowed multiplication method. Additionally, it introduces a deterministic nonce generation approach inspired by EdDSA to further refine digital signature efficiency. These innovations collectively contribute to a more efficient cryptographic process suitable for constrained environments.

The **theoretical significance** lies in the development of a new mathematical apparatus that makes it possible to optimize electronic signature operations.

The **practical significance** of this study is its applicability in low-power embedded systems, where computational and memory resources are highly limited. By optimizing ECC operations, this research enhances the security and performance of cryptographic implementations in unmanned vehicle systems and similar embedded applications, ensuring secure communications without exceeding hardware constraints.

The **proposed method** was implemented on the Arduino Atmega 2560 R3, achieves up to a 54,1 % **results are showing** reduction in cycle count and a 72,6 % decrease in SRAM usage for key generation, alongside significant performance improvements in signing and verification processes. Experimental results confirm its effectiveness in optimizing ECC operations for constrained devices of unmanned vehicle systems.

Keywords: digital signature, elliptic curve point multiplication, resource-constrained devices, unmanned vehicle systems (UVS), microcontroller

Funding. This research was funded by the Ministry of Science and Higher Education of the Russian Science Foundation (No. 075-00003-24-02, FSEE-2024-0003).

For citation: Sabbry N.H. Implementation of an ECC Digital Signature Technique in Constrained Environments. *Proceedings of Telecommunication Universities*. 2025;11(2):101–108. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-101-108. EDN:DWRJHM

Введение

В современном мире обеспечение безопасности в средах с ограниченными ресурсами (ограниченных средах) является критически важной задачей. Криптография на эллиптических кривых (ECC,

аббр. от англ. Elliptic Curve Cryptography) обеспечивает высокий уровень безопасности при небольших размерах ключей, что делает ее идеальной для таких устройств, как микроконтроллеры в беспилотных транспортных системах. Среди применений ECC – схемы электронной подписи, такие как

ECDSA [1] и EdDSA [2], они гарантируют аутентификацию и целостность данных. Однако вычислительные требования для операции скалярного умножения точки, используемых в ЕСС, создают сложности для устройств с ограниченной памятью и вычислительной мощностью.

В данном исследовании предлагается оптимизированный метод электронной подписи ЕСС для сред с ограниченными ресурсами. Оптимизация включает два ключевых этапа: (1) улучшение базовой операции ЕСС скалярного умножения точки и (2) повышение эффективности электронной подписи за счет детерминированной генерации закрытых однократно используемых чисел (nonce).

Первый этап расширяет предшествующее исследование [3], используя свойства циклической группы и противоположных чисел для оптимизации умножения точек. В частности, он снижает сложность удвоения точки в расширенных скрученных кривых Эдвардса ($\mathcal{E}^e$) [4] и интегрирует эти свойства в оконный метод умножения точек на Edwards25519.

Второй этап основывается на [5], в данной работе представлена идея оптимизации алгоритма электронной подписи за счет детерминированной генерации закрытых однократно используемых чисел, путем комбинирования сообщения с хешированным закрытым ключом (x), аналогично алгоритму, используемому в EdDSA.

Эти оптимизации были реализованы на микроконтроллере ATmega 2560 R3 для оценки их эффективности в средах с ограниченными ресурсами. Результаты демонстрируют значительное сокращение времени выполнения, использование памяти и числа тактов при реализации на низкоуровневом C и ассемблера, подчеркивая возможность применения предложенных разработок в реальных криптографических приложениях.

Структура данной работы следующая: раздел 2 описывает предлагаемую методологию, раздел 3 представляет оценку производительности, раздел 4 содержит ключевые выводы и направления будущих исследований.

Предложенная методология

Предлагаемая методология основана на реализации оптимизированных техник в рамках кривой Edwards25519, используя $\mathcal{E}^e$ и оконный метод скалярного умножения точки.

Оптимизированная стратегия скалярного умножения точки для устройств с ограниченными ресурсами

Свойства циклической группы и противоположного числа могут быть интегрированы с проверенными методами умножения точек [3]. В данном исследовании эти свойства интегрируются с окон-

ным методом скалярного умножения точки при размере окна $\omega = 4$ на кривой Edwards25519, используя $\mathcal{E}^e$.

В деталях методологию данного подхода можно представить следующим образом.

1) Свойство циклической группы эллиптической кривой.

Точки на эллиптической кривой образуют циклическую группу по сложению, т. е. существует базовая точка P такая, что любая точка может быть выражена как $Q = scalar \cdot P$. Порядок группы $\#E$ определяет момент повторения последовательности, удовлетворяя $\#E \cdot P = O$ (точка на бесконечности), $(\#E + 1) \cdot P = P$, $(\#E + 2) \cdot P = 2P$ и так далее циклически [6]. Например, вместо вычисления $(\#E + 2)P$ через несколько умножений оно эффективно выводится из $2P$, минимизируя общее количество вычислений.

2) Свойство противоположного числа в группах эллиптической кривой.

В теории групп свойство обратного элемента означает, что для каждой точки P существует обратная точка $-P$, такая, что $P + (-P) = O$, где O – единичный элемент. Для уменьшенных скаляров $\tilde{s}$, превышающих половину порядка эллиптической кривой $\#E$, свойство противоположного числа в теории групп обеспечивает дополнительную оптимизацию, используя обратные точки на эллиптической кривой.

3) Исключение параметров в $\mathcal{E}^e$.

Как обсуждается в [4], $\mathcal{E}^e$ упрощает сложение, исключая параметры d и a , если $a = -1$. Однако при удвоении точки a остается как:

$$D \leftarrow a \cdot A,$$

где $a = -1(\bmod p)$; p – простой модуль, определяющий конечное поле, $p = 2^{255} - 19$.

В этом контексте a хранится как 32-байтовый массив в формате little-endian в 8-разрядных микроконтроллерах: $a = \{0xEC, 0xFF, \dots, 0xFF, 0x7F\}$.

Вместо вычисления $D = -A$ через умножение $D = a \cdot A$ предложенный метод использует противоположные числа в конечных полях, заменяя умножение вычитанием: $D = p - A$.

Эта оптимизированная техника инверсии упрощает $D = -A$, снижая вычислительную сложность. Поскольку вычитание более эффективно, чем умножение, этот подход повышает производительность в устройствах с ограниченными ресурсами, особенно в оконном методе скалярного умножения точки, где частые операции удвоения выигрывают за счет уменьшенного числа вычислительных циклов.

4) Интеграция с существующими методами.

Интеграция предложенного подхода с оконным методом умножения точек эллиптической кривой осуществляется в два этапа.

Этап 1. Необходимо проверить, что скаляр s находится в пределах порядка эллиптической кривой $[0, \#E - 1]$ в зависимости от свойства циклической группы эллиптической кривой:

Пока $s \geq \#E$: Сокращенный скаляр $\tilde{s} = s - \#E$.

Этап 2. Построение оптимизированной функции оконного умножения точек.

Входные данные:

- $\tilde{s}$ или оптимизированный скаляр $\tilde{s}$;
- базовая точка G ;
- $\#E$;
- p ;
- размер окна ω (определяет длину окна в оконном методе);
- предварительно вычисленные точки $(P, 2P, 3P, \dots, P_{2^{\omega-1}})$.

Выходные данные: аффинные координаты (x, y) результирующей точки $(Q = s \cdot G)$.

Описание алгоритма на псевдокоде:

- 1) Если $(\tilde{s} > \frac{\#E}{2})$, использовать свойство противоположного числа в группах точек эллиптической кривой:
 - оптимизированный скаляр $\tilde{s} = \tilde{s} - \frac{\#E}{2}$;
 - вызвать «Оптимизированная функция оконного умножения точек»;
 - Q_x (х-аффинная координата) = p (инверсная точка (х-координата));
 - Q_y (у-аффинная координата) = p (инверсная точка (у-координата));
 - осуществить возврат.
- 2) Q = «Предвычисленная точка» (наиболее значимое окно).
- 3) Главный цикл (итерация по окнам).

Внутренний цикл (итерация ω раз):

$$Q = 2Q // \text{Удвоение,}$$

$Q = Q + \text{Предвычисленные точки [текущее значение окна].}$

Классический оконный метод включает удвоение точки ω раз, а затем операцию сложения с предварительно вычисленной точкой, если значение окна не равно нулю. В противном случае сложение пропускается [7], что вызывает несоответствия в выполнении и делает реализацию уязвимой к атакам по времени [8, 9]. Чтобы обеспечить последовательное выполнение, в метод добавляется дополнительная предварительно вычисленная единичная точка OG , представленная как $(0 : 1 : 0 : 1)$ в расширенной проективной системе координат кривой E^e . Так как сложение любой точки с OG со-

храняет исходную точку, это изменение поддерживает целостность операций. Размер массива предварительно вычисленных точек увеличивается следующим образом: $\omega = 4: 17$ точек (16 стандартных + OG).

С добавлением OG сложение выполняется равномерно, обеспечивая реализацию с постоянным временем выполнения. Поскольку безопасность ЕСС зависит от половины битовой длины порядка группы эллиптической кривой (из-за того, что Роллард алгоритм Полларда решает ЕCDLP за $O(\sqrt{\#E})$ [10], сокращение количества операций умножения на основе порядка кривой не ослабляет безопасность скалярного умножения точки.

Оптимизированный алгоритм электронной подписи на основе эллиптических кривых для устройств с ограниченными ресурсами

Оптимизация электронной подписи включает:

- использование аналогичного подхода к генерации закрытых ключей, как в алгоритме электронной подписи (EdDSA), для генерации детерминированных попис; в данном исследовании предлагается генерировать одноразовое число путем комбинирования сообщения с хешированным значением закрытого ключа (x) ;
- исключение открытого одноразового числа (R) из расчета вызова $c = H_{sig}(X, R, m)$.

Алгоритм генерации и проверки подписи представляет собой последовательность действий.

Во-первых, чтобы подписать сообщение (m) , подписант выполняет следующие действия.

- 1) Выбирает закрытый ключ (x) такой, что:

$$0 < x < p.$$

- 2) Открытый ключ $(X): X = G^x$.

- 3) Вычисляет $h = H(\text{prefix})$.

- 4) Вычисляет закрытое одноразовое число:

$$r = H(h || m).$$

- 5) Вычисляет открытое одноразовое число:

$$R = G^r.$$

- 6) Вычисляет $c = H(X, m)$.

- 7) Вычисляет $s \equiv (r + cx)$.

Подпись представляет собой пару (R, s) . Таким образом, подписант отправляет сообщение (m) , открытый ключ (X) , (R) и (s) .

В-вторых, проверяющий считает подпись действительной, только если:

$$G^s = RX^c.$$

Доказательство корректности алгоритма:

$$G^s = RX^c,$$

т. к. $R = G^r$ и $X = G^x$, то $G^s = G^r(G^x)^c = G^{r+xc} = G^s$, где $s = r + xc$.

Реализация предложенного метода на микроконтроллере ATmega2560

Алгоритмы электронной подписи обычно включают в себя генерацию ключей, подписание и проверку [2].

Процесс генерации ключей электронной подписи состоит из трех этапов.

Этап 1. Генерация случайных чисел (RNG, аббр. от англ. Random Number Generator)

Аппаратный источник энтропии (аналоговый вход на контакте A0) фиксирует шум окружающей среды в качестве исходного значения. Поточковый шифр ChaCha20 обрабатывает эту энтропию в криптографически безопасный 32-байтовый случайный выходной сигнал [11]. Криптографические шифры и хеш-функции обеспечивают надежную случайность для получения префикса и закрытого ключа, требуя 117,618 тактов.

Этап 2. Хеширование

Выходные данные RNG хешируются с использованием 512-битной хеш-функции, формируя 64-байтовое значение, разделенное на две 32-байтовые части [2]:

- нижние 32 байта формируют закрытый ключ;
- верхние 32 байта используются как префикс для генерации закрытого поппс, комбинируемого с сообщением при подписании.

Этот этап требует 277 575 тактов.

Этап 3. Вычисление открытого ключа

Закрытый ключ умножается на базовую точку для получения открытого ключа с использованием метода, представленного в [3], при $w = 4$, требующего 15 105 854 тактов, используя 30 518 байтов (12 %) из доступных 253 952 байтов памяти программ (флеш-памяти) и потребляя 429 байтов (5 %) статической памяти с произвольным доступом

(SRAM, аббр. от англ. Static Random Access Memory), оставляя 7 763 байта из максимальных 8 192 байтов.

Для оценки эффективности предлагаемого метода генерации ключей в электронной подписи сравнения должны проводиться в идентичных средах и с использованием одной и той же кривой, как в [12]. Сравнения с исследованиями, использующими значительно отличающиеся аппаратные условия, могут исказить точность усовершенствований предложенного метода. В таблице 1 и на рисунке 1 представлено это сравнение, демонстрирующее результаты проведенного авторами исследования и [12], оба из которых используют микроконтроллер ATmega 2560 и кривую Edwards25519. Для реализации предлагаемого метода был использован язык C и ассемблер.

При сравнении с высокоскоростной реализацией (High-Speed) [12] предлагаемый метод сокращает количество необходимых тактов на 31,1 % и уменьшает использование SRAM на 72,6 %. Аналогично, по сравнению с реализацией, ориентированной на минимальное использование ресурсов (Low-Area), предлагаемый метод обеспечивает сокращение тактов на 54,1 % и уменьшение использования SRAM на 66,5 %.

ТАБЛИЦА 1. Сравнение генерации ключей: предложенный метод vs исследование в сопоставимых условиях [12]

TABLE 1. Key Generation Comparison: Proposed Method vs Prior Research under Comparable Conditions [12]

Реализация	Требуемые такты	SRAM, байт
crypto_sign_keypair (High-Speed) [16] ($w = 4$)	21924771	1566
crypto_sign_keypair (Low-Area) [16] ($w = 2$)	32937,940	1282
Предложенный метод ($w = 4$)	15105854	429

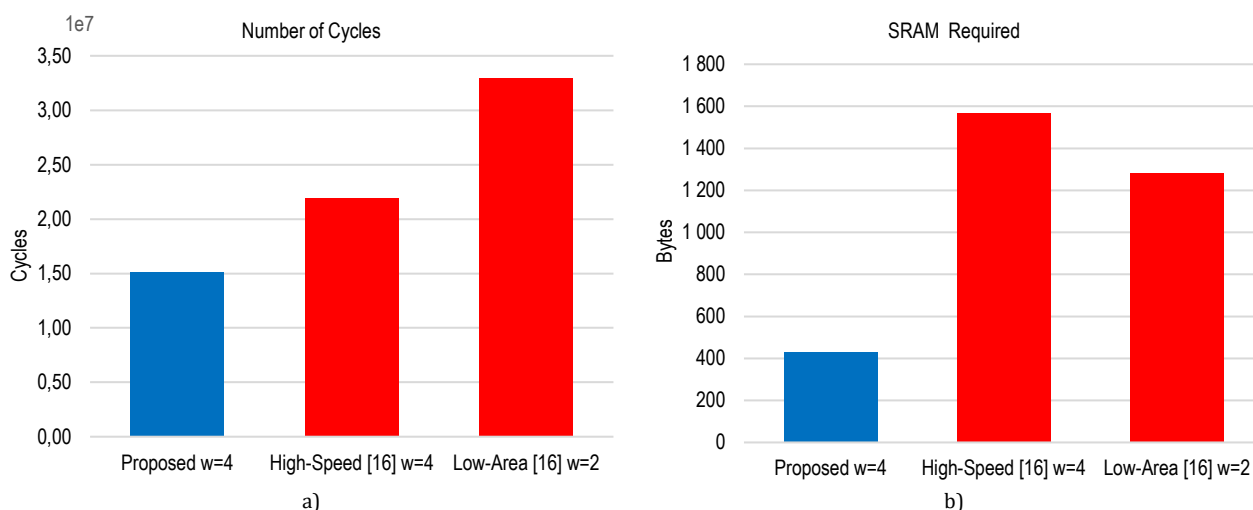


Рис. 1. Сравнение генерации ключей: предложенный метод (а) с предыдущими исследованиями (б) в сопоставимых условиях [12]

Fig. 1. Key Generation Comparison: Proposed Method (a) vs Prior Research (b) under Comparable Conditions [12]

В процессе подписания производительность предлагаемого метода требует 15 435 539 тактов, используя 25 652 байта (10 %) памяти программ (флеш-памяти) и 639 байтов (7 %) SRAM. Для оценки эффективности предлагаемого метода подписания проводятся сравнения с предыдущими исследованиями в аналогичных средах с ограниченными ресурсами. В таблице 2 и на рисунке 2 представлено сравнение количества тактов из данного исследования с результатами из crypto_sign (High-Speed) и crypto_sign (Low-Area) [12, 13], выполненным в сопоставимых условиях. Эти исследования были выбраны для обеспечения справедливой и точной оценки, так как сравнение с исследованиями, использующими существенно отличающиеся аппаратные условия, могло бы снизить четкость и надежность оценки производительности предлагаемого метода.

ТАБЛИЦА 2. Сравнение количества циклов: предложенный метод с предыдущими исследованиями в сопоставимых условиях [12, 13]

TABLE 2. Number of Cycles Comparison: Proposed Method vs Prior Researches under Comparable Conditions [12, 13]

Метод	Реализация	Количество циклов
Предложенный метод	$\omega = 4$	15435539
Ed25519 [13]	–	22688583
crypto_sign [12]	High-Speed, $\omega = 4$ Low-Area, $\omega = 2$	23211611 34342230

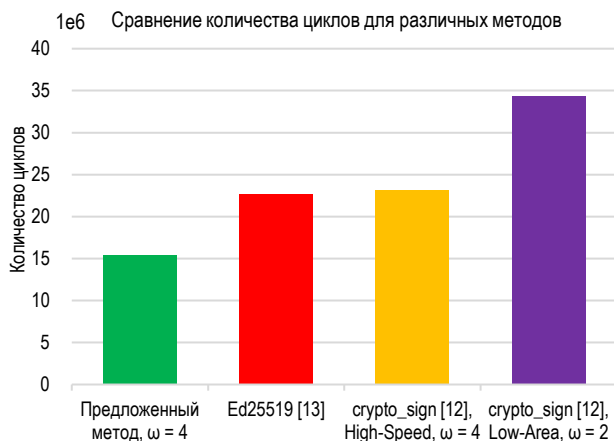


Рис. 2. Сравнение количества циклов: предложенный метод с исследованиями в сопоставимых условиях [12, 13]

Fig. 2. Number of Cycles Comparison: Proposed Method vs Prior Researches under Comparable Conditions [12, 13]

Предлагаемый метод требует меньше тактов: на 31,97 %, чем у Ed25519 [13], и на 33,50 %, чем у высокоскоростной реализации crypto_sign [12]. При сравнении с реализацией crypto_sign, ориентированной на минимальное использование ресурсов [12], предлагаемый метод достигает сокращения числа тактов на 55,05 %.

В процессе проверки производительность предлагаемого метода составляет 30 521 674 такта, используя 28 624 байта (11 %) памяти программ (флеш-памяти) и 1 187 байтов (14 %) SRAM. В таблице 3 и на рисунке 3 представлено сравнение между предлагаемым методом и подходами High-Speed и Low-Area, описанными в [12], выполненным в аналогичных средах с ограниченными ресурсами.

ТАБЛИЦА 3. Сравнение результатов этапа проверки: предложенный метод с предыдущими исследованиями в сопоставимых условиях

TABLE 3. Comparison of Verification Step Results: Proposed Method vs Prior Research under Comparable Conditions

Реализация	ω	Количество циклов	Требуемая SRAM, байт
Предложенный метод	4	30521674	1187
High-Speed [12]	4	32619197	1317
Low-Area [12]	2	40093186	1349

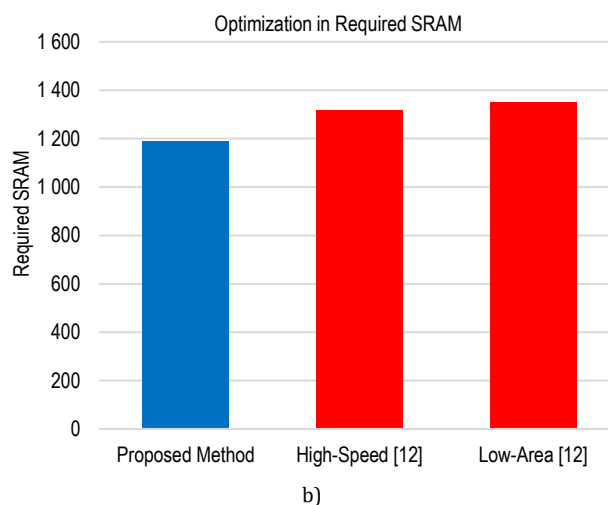
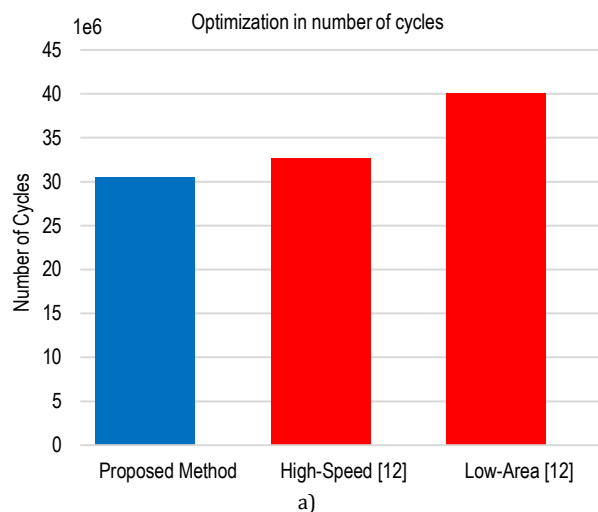


Рис. 3. Сравнение результатов этапа проверки: предложенный метод (а) с предыдущими исследованиями (б) в сопоставимых условиях

Fig. 3. Comparison of Verification Step Results: Proposed Method (a) vs Prior Research (b) Under Comparable Conditions

При сравнении с High-Speed [12] реализацией (размер окна – 4), предложенный метод снижает количество циклов на 6,43 % и использование SRAM – на 9,87 %. По сравнению с Low-Area реализацией (размер окна 2), предложенный метод достигает сокращения числа циклов на 23,87 % и использования SRAM – на 12,01 %. Эти результаты подчеркивают превосходную производительность и эффективность предложенного метода на этапе проверки электронной подписи при сопоставимых условиях.

Заключение

Данное исследование представляет оптимизированную технику построения электронной подписи ЕСС, разработанную для ограниченных сред, в частности для беспилотных транспортных систем.

Вследствие повышения эффективности скалярного умножения точки за счет свойств циклической группы, противоположных чисел и модифицированного метода оконного умножения, предложенный подход значительно снижает вычислительные затраты. Кроме того, интеграция метода детерминированной генерации однократно используемого числа повышает безопасность при электронной подписи, сохраняя ее эффективность. Экспериментальные результаты на микроконтроллере ATmega2560 R3 демонстрируют значительное сокращение времени выполнения, потребления памяти и количества циклов по сравнению с существующими методами. Эти результаты подтверждают эффективность предложенных оптимизаций и повышение криптографической безопасности для встраиваемых систем с ограниченными ресурсами.

Список источников

1. Johnson D., Menezes A., Vanstone S. The Elliptic Curve Digital Signature Algorithm (ECDSA) // *International Journal of Information Security*. 2001. Vol. 1. PP. 36–63. DOI:10.1007/s102070100002
2. Josefsson S., Liusvaara I. Edwards-Curve Digital Signature Algorithm (EdDSA). 2017. DOI:10.17487/RFC8032
3. Sabbry N.H., Levina A.B. An Optimized Point Multiplication Strategy in Elliptic Curve Cryptography for Resource-Constrained Devices // *Mathematics*. 2024. Vol. 12. Iss. 6. P. 881. DOI:10.3390/math12060881. EDN:JZKADF
4. Hisil H., Wong K.K., Carter G., Dawson E. Twisted Edwards Curves Revisited // *Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT 2008, Melbourne, Australia, 7–11 December 2008)*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2008. Vol. 5350. PP. 326–343. DOI:10.1007/978-3-540-89255-7_20
5. Sabbry N.H., Levina A. Nonce generation techniques in Schnorr multi-signatures: Exploring EdDSA-inspired approaches // *AIMS Mathematics*. 2024. Vol. 9. Iss. 8. PP. 20304–20325. DOI:10.3934/math.2024988. EDN:YSJYIC
6. Paar C., Pelzl J. *Understanding Cryptography: A Textbook for Students and Practitioners*. Berlin, Heidelberg: Springer, 2010. DOI:10.1007/978-3-642-04101-3
7. Hankerson D., Menezes A. *Elliptic Curve Cryptography* // In: Jajodia S., Samarati P., Yung M. (eds.) *Encyclopedia of Cryptography, Security and Privacy*. Berlin, Heidelberg: Springer, 2021. PP. 1–2. DOI:10.1007/978-3-642-27739-9_245-2
8. Izu T., Möller B., Takagi T. Improved Elliptic Curve Multiplication Methods Resistant Against Side Channel Attacks // *Proceedings of the Third International Conference on Cryptology in India (INDOCRYPT 2002, Hyderabad, India, 16–18 December 2002)*. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2002. Vol. 2551. PP. 296–313. DOI:10.1007/3-540-36231-2_24
9. Shenets N.N., Petushkov A.S. New Regular Sliding Window Algorithms for Elliptic Curve Scalar Point Multiplication // *Automatic Control and Computer Sciences*. 2021. Vol. 55. PP. 1029–1038. DOI:10.3103/S0146411621080289. EDN:FWONTD
10. Cheon J.H., Hong J., Kim M. Speeding Up the Pollard Rho Method on Prime Fields // *Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT 2008, Melbourne, Australia, 7–11 December 2008)*. Berlin, Heidelberg: Springer, 2008. PP. 471–488. DOI:10.1007/978-3-540-89255-7_29
11. Alyas H.H., Abdullah A.A. Enhancement the ChaCha20 Encryption Algorithm Based on Chaotic Maps // In: Kumar R., Mishra B.K., Pattnaik P.K. *Next Generation of Internet of Things: Proceedings of ICNIIoT 2021*. Lecture Notes in Networks and Systems. Singapore: Springer, 2021. Vol. 201. PP. 91–107. DOI:10.1007/978-981-16-0666-3_10. EDN:ZPPOQO
12. Hutter M., Schwabe P. NaCl on 8-Bit AVR Microcontrollers // *Proceedings of the 6th International Conference on Cryptology in Africa «Progress in Cryptology» (AFRICACRYPT 2013, Cairo, Egypt, 22–24 June 2013)*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013. Vol. 7918. PP. 156–172. DOI:10.1007/978-3-642-38553-7_9
13. Nouma S.E., Yavuz A.A. Lightweight and Resilient Signatures for Cloud-Assisted Embedded IoT Systems // *arXiv preprint arXiv:2409.13937*. 2024. DOI:10.48550/arXiv.2409.13937

References

1. Johnson D., Menezes A., Vanstone S. The Elliptic Curve Digital Signature Algorithm (ECDSA). *International Journal of Information Security*. 2001;1:36–63. DOI:10.1007/s102070100002
2. Josefsson S., Liusvaara I. *Edwards-Curve Digital Signature Algorithm (EdDSA)*. 2017. DOI:10.17487/RFC8032

3. Sabbry N.H., Levina A.B. An Optimized Point Multiplication Strategy in Elliptic Curve Cryptography for Resource-Constrained Devices. *Mathematics*. 2024;12(6):881. DOI:10.3390/math12060881. EDN:JZKADF
4. Hisil H., Wong K.K., Carter G., Dawson E. Twisted Edwards Curves Revisited. *Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security, ASIACRYPT 2008, 7–11 December 2008, Melbourne, Australia. Lecture Notes in Computer Science, vol.5350*. Berlin, Heidelberg: Springer; 2008. p.326–343. DOI:10.1007/978-3-540-89255-7_20
5. Sabbry N.H., Levina A. Nonce generation techniques in Schnorr multi-signatures: Exploring EdDSA-inspired approaches. *AIMS Mathematics*. 2024;9(8):20304–20325. DOI:10.3934/math.2024988. EDN:YSJYIC
6. Paar C., Pelzl J. *Understanding Cryptography: A Textbook for Students and Practitioners*. Berlin, Heidelberg: Springer; 2010. DOI:10.1007/978-3-642-04101-3
7. Hankerson D., Menezes A. Elliptic Curve Cryptography. In: Jajodia S., Samarati P., Yung M. (eds.) *Encyclopedia of Cryptography, Security and Privacy*. Berlin, Heidelberg: Springer; 2021. p.1–2. DOI:10.1007/978-3-642-27739-9_245-2
8. Izu T., Möller B., Takagi T. Improved Elliptic Curve Multiplication Methods Resistant Against Side Channel Attacks. *Proceedings of the Third International Conference on Cryptology in India, INDOCRYPT 2002, 16–18 December 2002, Hyderabad, India. Lecture Notes in Computer Science, vol.2551*. Berlin Heidelberg: Springer; 2002. p.296–313. DOI:10.1007/3-540-36231-2_24
9. Shenets N.N., Petushkov A.S. New Regular Sliding Window Algorithms for Elliptic Curve Scalar Point Multiplication. *Automatic Control and Computer Sciences*. 2021;55:1029–1038. DOI:10.3103/S0146411621080289. EDN:FWONTD
10. Cheon J.H., Hong J., Kim M. Speeding Up the Pollard Rho Method on Prime Fields. *Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security, ASIACRYPT 2008, 7–11 December 2008, Melbourne, Australia*. Berlin, Heidelberg: Springer; 2008. p.471–488. DOI:10.1007/978-3-540-89255-7_29
11. Alyas H.H., Abdullah A.A. Enhancement the ChaCha20 Encryption Algorithm Based on Chaotic Maps. In: Kumar R., Mishra B.K., Pattnaik P.K. *Next Generation of Internet of Things: Proceedings of ICNGIoT 2021. Lecture Notes in Networks and Systems, vol.201*. Singapore: Springer; 2021. p.91–107. DOI:10.1007/978-981-16-0666-3_10. EDN:ZPPOQO
12. Hutter M., Schwabe P. NaCl on 8-Bit AVR Microcontrollers. *Proceedings of the 6th International Conference on Cryptology in Africa «Progress in Cryptology», AFRICACRYPT 2013, 22–24 June 2013, Cairo, Egypt. Lecture Notes in Computer Science, vol.7918*. Berlin, Heidelberg: Springer; 2013. p.156–172. DOI:10.1007/978-3-642-38553-7_9
13. Nouma S.E., Yavuz A.A. Lightweight and Resilient Signatures for Cloud-Assisted Embedded IoT Systems. *arXiv preprint arXiv:2409.13937*. 2024. DOI:10.48550/arXiv.2409.13937

Статья поступила в редакцию 05.03.2025; одобрена после рецензирования 25.03.2025; принята к публикации 07.04.2025.

The article was submitted 05.03.2025; approved after reviewing 25.03.2025; accepted for publication 07.04.2025.

Информация об авторе:

**САБРИ
НАУРАС ХУССЕЙН**

аспирант кафедры факультет информационно-измерительных и биотехнических систем Санкт-Петербургского государственного электротехнического университет «ЛЭТИ» имени В.И. Ульянова (Ленина)

 <https://orcid.org/0009-0003-2429-5122>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.

Научная статья

УДК 519.725, 512.623, 004.312

<https://doi.org/10.31854/1813-324X-2025-11-2-109-120>

EDN:VKAHMV



Метод вычисления полевой свертки на основе разложения многозначного расширенного поля Галуа

Илья Владимирович Ульянов, lopi2.lll@mail.ru

Академия Федеральной службы охраны Российской Федерации,
Орел, 302020, Российская Федерация

Аннотация

Актуальность. Одной из нерешенных проблем теории помехоустойчивого кодирования остается проблема построения декодеров длинных кодов с низкой вычислительной сложностью. С точки зрения алгебраической теории кодирования краеугольным камнем для этого является операция умножения двух многочленов a и b над полем $GF(q^k)$ по модулю третьего многочлена d . С возрастанием q и k применение методов вычисления полевой свертки на основе операций логарифмирования и антилогарифмирования становится малоэффективным ввиду задействования большого объема памяти для построения таблиц. Упрощенные реализации полевой свертки, использующие несимметричность сопровождающей матрицы, и аналитические (не табличные) методы логарифмирования и антилогарифмирования, использующие полиномы Жегалкина, разработаны только для $q = 2$. Умножители на основе регистров сдвига обладают значительно меньшим быстродействием при больших q и k .

Целью исследования является поиск вариантов снижения вычислительной сложности операции полевой свертки в многозначных расширенных полях Галуа при ее синтезе в логическом базисе «И»–«ИЛИ»–«НЕ».

Методы. Проведен анализ одноктактных методов умножения элементов многозначного расширенного поля Галуа, заданных в векторном или полиномиальном виде для различных степенных базисов. Приведены примеры вычисления полевых свертки в многозначных полях Галуа различными методами. Изучена структура рассматриваемого типа полей.

Решение. Показано, что операции сложения и умножения в поле $GF(q)$, синтезированные на элементах логического базиса «И»–«ИЛИ»–«НЕ», вносят основной вклад в сложность итоговой логической схемы. Выявлено, что использование свойства разложения поля $GF(q^k)$ на подмножества по степени примитивного элемента поля $GF(q)$ позволяет сократить число операций умножения. Предложен метод полевой свертки на основе матричного метода и преобразования Ганкеля – Теплица, учитывающий структуру поля, что позволяет сократить общее число логических элементов и повысить быстродействие проектируемого схемотехнического решения, а именно уменьшить цену по Квайну и ранг схемы. Дана сравнительная оценка разработанного метода.

Новизна: впервые предложен метод полевой свертки двух векторов в поле $GF(q^k)$, один из которых представлен в индикаторном виде.

Теоретическая значимость. Предложен новый метод вычисления полевой свертки на основе разложения многозначного расширенного поля Галуа. Доказано сокращение общего числа логических операций.

Практическая значимость. Предложенное решение может быть использовано при синтезе кодирующих-декодирующих устройств многозначных (символьных) кодов на элементах двоичной логики.

Ключевые слова: помехоустойчивое кодирование, полевая свертка, поле Галуа, матричный метод, преобразование Гаккеля – Теплица, цена по Квайну

Ссылка для цитирования: Ульянов И.В. Метод вычисления полевой свертки на основе разложения многозначного расширенного поля Галуа // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 109–120. DOI:10.31854/1813-324X-2025-11-2-109-120. EDN:VKAHMV

Original article

<https://doi.org/10.31854/1813-324X-2025-11-2-109-120>

EDN:VKAHMY

Method for Calculating Field Convolution Based on the Decomposition of a Multi-Valued Extended Galois Field

 Ilya V. Ulyanov, lopi2.lll@mail.ru

Academy of the Federal Security Service of the Russian Federation,
Orel, 302020, Russian Federation

Annotation

Relevance. One of the unsolved problems of the theory of error-correcting coding is the problem of constructing decoders of long codes with low computational complexity. From the point of view of algebraic coding theory, the cornerstone for this is the operation of multiplying two polynomials a and b over the field $GF(q^k)$ modulo the third polynomial g . As q and k increase, the use of methods for calculating the field convolution based on the logarithm and antilogarithm operations becomes ineffective due to the use of a large amount of memory for constructing tables. Simplified implementations of the field convolution using the asymmetry of the accompanying matrix and analytical (non-tabular) methods of logarithm and antilogarithm using Zhegalkin polynomials have been developed only for $q = 2$. Multipliers based on shift registers have a significantly lower speed for large q and k .

The aim of the study is to find options for reducing the computational complexity of the field convolution operation in multivalued extended Galois fields during its synthesis in the logical basis "AND"–"OR"–"NOT".

Methods. An analysis of single-cycle methods for multiplying elements of a multivalued extended Galois field specified in vector or polynomial form for various power bases is carried out. Examples of calculating field convolutions in multivalued Galois fields by various methods are given. The structure of the considered type of fields is studied.

Results. It is shown that addition and multiplication operations in the field $GF(q)$ synthesized on the elements of the logical basis "AND"–"OR"–"NOT" make the main contribution to the complexity of the resulting logical circuit. It is revealed that using the property of decomposition of the field $GF(q^k)$ into subsets by the power of the primitive element of the field $GF(q)$ allows to reduce the number of multiplication operations. A field convolution method based on the matrix method and the Hankel – Toeplitz transform is proposed, taking into account the field structure, which allows to reduce the total number of logical elements and increase the performance of the designed circuit solution, namely, to reduce the Quine price and the circuit rank. A comparative assessment of the developed method is given.

Scientific novelty. For the first time, a method of field convolution of two vectors in the field $GF(q^k)$ is proposed, one of which is presented in the indicator form.

Theoretical / Practical significance. A new method for calculating the field convolution based on the decomposition of a multivalued extended Galois field is proposed. Reduction of the total number of logical operations is proved. The proposed solution can be used in the synthesis of encoding and decoding devices for multi-valued (symbolic) codes on binary logic elements.

Keywords: error-correcting coding, field convolution, Galois field, matrix method, Gakkel – Toeplitz transforms, Quine price

For citation: Ulyanov I.V. Method for Calculating Field Convolution Based on the Decomposition of a Multi-Valued Extended Galois Field. *Proceedings of Telecommunication Universities*. 2025;11(2):109–120. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-109-120. EDN:VKAHMY

Введение

Одной из фундаментальных проблем теории кодирования, препятствующих построению длин-

ных кодов, является трудоемкость операции умножения двух многочленов a и b над полем $GF(q^k)$ по модулю третьего многочлена g . Такая

операция называется сверткой многочленов. Особый интерес с теоретической и практической точки зрения представляет полевая свертка, когда многочлен g неприводим над полем $GF(q^k)$. Полевая свертка в качестве бинарной операции умножения в совокупности с бинарной операцией сложения являются основой для формирования поля $GF(q^k)$, состоящего из множества всевозможных многочленов-остатков по модулю неприводимого многочлена g . Очевидно, что умножение самих элементов (многочленов-остатков) также выполняется с помощью полевой свертки [1–3]. Другие виды сверток, а именно линейная и циклическая, в работе не рассматриваются. Применимость предлагаемого метода для их реализации является предметом дальнейших исследований автора.

В теории и практике кодирования информации существует ряд подходов к вычислению полевой свертки. В общем случае аналитическое выражение для прямого вычисления полевой свертки достаточно громоздкое. Поэтому для синтеза аппаратной реализации на этапе проектирования по заданному многочлену-константе g итерационно просчитывают рекуррентную формулу и получают компактную прямую формулу, на основе которой разрабатывают устройство [1].

При малых q и k с точки зрения аппаратной реализации предпочтительным является синтез устройства на основе компактной прямой формулы. В случае же увеличения q и k возрастает вычислительная сложность такого подхода, поэтому находят свое применение таблицы логарифмов и антилогарифмов. Это кардинально снижает вычислительную сложность, однако требует значительных объемов памяти и вносит задержки на поиск в соответствующих таблицах. Дальнейшее увеличение q и k приводит к невозможности использования таблиц логарифмов и антилогарифмов в виду их значительного объема. Это требует разработки эффективных с вычислительной точки зрения и с точки зрения используемой памяти алгоритмов расчета полевой свертки, рассчитываемой за один такт для больших q и k . Стоит отметить, что существуют многотактные умножители произвольных элементов поля $GF(q^k)$, которые синтезируются на основе регистров сдвига [4, 5]. В данной работе они не рассматриваются по причине своей нерегулярности.

В исследованных автором источниках [1–21], кроме, пожалуй [2, 6], рассматриваются вопросы практической реализации операций в полях $GF(q^k)$, где $q = 2$. Предлагаемый далее метод предполагает $q > 2$, т. е. полевая свертка рассматривается в многозначных (символьных) полях Галуа. При $q = 2$ предлагаемый метод вырождается в общеизвестный матричный метод вычисления полевой свертки [2].

Стоит отметить, что по тексту статьи при указании поля Галуа в общем виде – $GF(q^k)$ подразумевается, что оно образовано соответствующим неприводимым полиномом $P(x)$. Если же поле Галуа указано с конкретными значениями q и k , например $GF(3^3)$, то это значит, что поле образовано полиномом, указанным в соответствующем примере, ссылка на который дается по тексту.

Логарифмирование–антилогарифмирование

Наиболее простым с вычислительной точки зрения является полевая свертка на основе операций логарифмирования–антилогарифмирования (рисунок 1). Под логарифмом в расширенном поле $GF(q^k)$ понимают степень i , в которую необходимо возвести примитивный элемент ϵ , чтобы получить рассматриваемый элемент поля ϵ^i : $i = \log_{\epsilon} \epsilon^i$ [1]. Множество логарифмов $\{0, 1, \dots, q^k - 2\}$ поля $GF(q^k)$ с определенными на нем бинарными операциями сложения и умножения образует коммутативное кольцо по модулю $q^k - 1$.

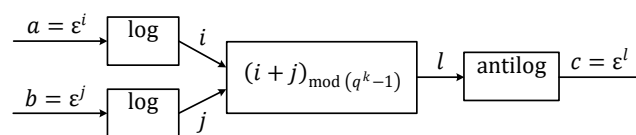


Рис. 1. Структурная схема умножения в поле $GF(q^k)$ на основе операций логарифмирования–антилогарифмирования

Fig. 1. Block Diagram of Multiplication in the Field $GF(q^k)$ Based on Logarithm–Antilogarithm Operations

Пример 1. Вычисление полевой свертки с помощью логарифмирования–антилогарифмирования

Рассмотрим свертку элементов поля $GF(3^3)$, построенного на основе неприводимого многочлена $P(x) = x^3 + 2x + 1$ (основание кода $q = 3$, порядок – $k = 3$). В таблице 1 приведены элементы поля для левого степенного базиса $(1, \epsilon, \epsilon^2)$:

ТАБЛИЦА 1. Элементы поля $GF(3^3)$ по основанию $P(x) = x^3 + 2x + 1$ для левого степенного базиса

TABLE 1. Elements of the Field $GF(3^3)$ to the Base $P(x) = x^3 + 2x + 1$ for the Left Power Basis

Десятичный номер	Степень ϵ^i	Логарифм i	Вектор
1	ϵ^0	0	(1 0 0)
2	ϵ^1	1	(0 1 0)
3	ϵ^2	2	(0 0 1)
4	ϵ^3	3	(2 1 0)
5	ϵ^4	4	(0 2 1)
6	ϵ^5	5	(2 1 2)
7	ϵ^6	6	(1 1 1)
8	ϵ^7	7	(2 2 1)
9	ϵ^8	8	(2 0 2)
10	ϵ^9	9	(1 1 0)
11	ϵ^{10}	10	(0 1 1)

Десятичный номер	Степень ε^i	Логарифм i	Вектор
12	ε^{11}	11	(2 1 1)
13	ε^{12}	12	(2 0 1)
14	ε^{13}	13	(2 0 0)
15	ε^{14}	14	(0 2 0)
16	ε^{15}	15	(0 0 2)
17	ε^{16}	16	(1 2 0)
18	ε^{17}	17	(0 1 2)
19	ε^{18}	18	(1 2 1)
20	ε^{19}	19	(2 2 2)
21	ε^{20}	20	(1 1 2)
22	ε^{21}	21	(1 0 1)
23	ε^{22}	22	(2 2 0)
24	ε^{23}	23	(0 2 2)
25	ε^{24}	24	(1 2 2)
26	ε^{25}	25	(1 0 2)

Множители:

$$a = (a_0 \ a_1 \ a_2) = \varepsilon^8 = (2 \ 0 \ 2),$$

$$b = (b_0 \ b_1 \ b_2) = \varepsilon^{13} = (2 \ 0 \ 0).$$

Результат:

$$c = (c_0 \ c_1 \ c_2) = a \times b = \varepsilon^8 \times \varepsilon^{13} = \varepsilon^{21} = (1 \ 0 \ 1).$$

Матричный способ

В ряде случаев вычислять полевою свертку удобнее на основе прямых выражений. Одним из способов получения таких выражений является матричный способ [2] (рисунок 2).

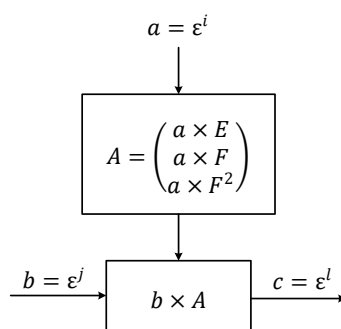


Рис. 2. Блок-схема матричного способа умножения элементов поля $GF(q^k)$

Fig. 2. Block Diagram of the Matrix Method for Multiplying Elements of the Field $GF(q^k)$

Пример 2. Вычисление полевой свертки матричным способом

Синтезируем прямые выражения для вычисления полевой свертки произвольных элементов поля, рассмотренного в примере 1:

1) определим сопровождающую матрицу и вычислим ее степени:

$$F = \begin{pmatrix} \varepsilon^1 \\ \varepsilon^2 \\ \varepsilon^3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}, F^2 = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix};$$

2) вычислим составную матрицу:

$$A = \begin{pmatrix} a \times E \\ a \times F \\ a \times F^2 \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & a_2 \\ 2a_2 & a_0 + a_2 & a_1 \\ 2a_1 & a_1 + 2a_2 & a_0 + a_2 \end{pmatrix},$$

$$\text{где } E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

3) синтезируем прямые выражения для вычисления результата $c = b \times A = (c_0 \ c_1 \ c_2)$:

$$c_0 = a_0 b_0 + 2a_1 b_2 + 2a_2 b_1, \quad (1)$$

$$c_1 = a_0 b_1 + a_1 b_0 + a_1 b_2 + a_2 b_1 + 2a_2 b_2, \quad (2)$$

$$c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0 + a_2 b_2; \quad (3)$$

4) примеры вычисления свертки для произвольных элементов поля:

$$a = \varepsilon^8 = (2 \ 0 \ 2), b = \varepsilon^{13} = (2 \ 0 \ 0);$$

$$c_0 = 2 \cdot 2 + 2 \cdot 0 \cdot 0 + 2 \cdot 2 \cdot 0 = 1,$$

$$c_1 = 2 \cdot 0 + 0 \cdot 2 + 0 \cdot 0 + 2 \cdot 0 + 2 \cdot 2 \cdot 0 = 0,$$

$$c_2 = 2 \cdot 0 + 0 \cdot 0 + 2 \cdot 2 + 2 \cdot 0 = 1;$$

$$c = \varepsilon^{21} = (1 \ 0 \ 1);$$

$$a = \varepsilon^{11} = (2 \ 1 \ 1), b = \varepsilon^{20} = (1 \ 1 \ 2);$$

$$c_0 = 2 \cdot 1 + 2 \cdot 1 \cdot 2 + 2 \cdot 1 \cdot 1 = 2;$$

$$c_1 = 2 \cdot 1 + 1 \cdot 1 + 1 \cdot 2 + 1 \cdot 1 + 2 \cdot 1 \cdot 2 = 1;$$

$$c_2 = 2 \cdot 2 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 2 = 2;$$

$$c = \varepsilon^5 = (2 \ 1 \ 2).$$

Преобразование Ганкеля – Теплица

Другим подходом, позволяющим при вычислениях над полями Галуа обходиться без операций логарифмирования и антилогарифмирования, является применения матриц Ганкеля (4) и Теплица (5) [3]. Выражение для полевой свертки элементов поля $GF(2^k)$, заданных в векторном или полиномиальном виде, может быть реализовано с помощью специальных векторно-матричных конструкций (рисунок 3). Элементами таких матриц являются модифицированные логарифмы поля $GF(2^k)$ (десятичные номера из таблиц логарифмирования-антилогарифмирования), которые затем заменяются на вектора соответствующих элементов поля.

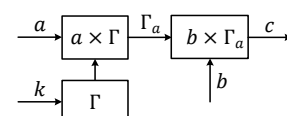


Рис. 3. Блок-схема умножения элементов поля $GF(2^k)$ на основе преобразования Ганкеля – Теплица

Fig. 3. A Block Diagram for Multiplying Elements of the $GF(2^k)$ Field Based on the Hankel – Toeplitz Transformation

$$\Gamma = \begin{pmatrix} 1 & 2 & \dots & k \\ 2 & 3 & \dots & (k+1) \\ \vdots & \vdots & \ddots & \vdots \\ k & (k+1) & \dots & (2k-1) \end{pmatrix}, \quad (4)$$

$$T = \begin{pmatrix} k & \dots & 2 & 1 \\ (k+1) & \dots & 3 & 2 \\ \vdots & \ddots & \vdots & \vdots \\ (2k-1) & \dots & (k+1) & k \end{pmatrix}. \quad (5)$$

Очевидно, что матрицы Γ и T связаны с помощью операции перестановки столбцов: $\Gamma = T$.

При замене элементов матрицы Γ соответствующими векторами, например, для поля $GF(2^3)$ с $(x) = x^3 + x + 1$ и правым степенным базисом $(\varepsilon^2, \varepsilon, 1)$, получим следующее векторное представление матрицы Γ :

$$\Gamma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} \varepsilon^0 & \varepsilon^1 & \varepsilon^2 \\ \varepsilon^1 & \varepsilon^2 & \varepsilon^3 \\ \varepsilon^2 & \varepsilon^3 & \varepsilon^4 \end{pmatrix} = \begin{pmatrix} (001) & (010) & (100) \\ (010) & (100) & (011) \\ (100) & (011) & (110) \end{pmatrix}.$$

Вычисление полевой свертки в полях $GF(2^k)$ производится в соответствии со схемой (см. рисунок 3), которая может быть записана в $\Sigma\Pi$ -форме следующим образом [3]:

$$c = \sum_{i=0}^{k-1} a_i \sum_{j=0}^{k-1} b_j \varepsilon^{i+j}, \quad (6)$$

где $c = (c_k \dots c_1 c_0)$ – вектор-результат; $a = (a_k \dots a_1 a_0)$ и $b = (b_k \dots b_1 b_0)$ – векторы-множители.

Пример 3. Вычисление полевой свертки с помощью преобразования Ганкеля – Теплица

Предложенный для полевой свертки в полях $GF(2^k)$ [3] подход может быть использован для полей $GF(q^k)$. Рассмотрим поле $GF(3^3)$, построенное в примере 1, с оговоркой для правого степенного базиса (векторное представление элементов поля будет взято с операцией перестановки столбцов по отношению к данным, указанным в таблице 1).

Раскроем вначале правую и затем левую суммы в выражении (6):

$$\begin{aligned} c &= \sum_{i=0}^2 a_i (b_0 \varepsilon^{i+0} + b_1 \varepsilon^{i+1} + b_2 \varepsilon^{i+2}) = \\ &= a_0 b_0 \varepsilon^0 + a_0 b_1 \varepsilon^1 + a_0 b_2 \varepsilon^2 + \\ &+ a_1 b_0 \varepsilon^1 + a_1 b_1 \varepsilon^2 + a_1 b_2 \varepsilon^3 + a_2 b_0 \varepsilon^2 + a_2 b_1 \varepsilon^3 + \\ &+ a_2 b_2 \varepsilon^4 = a_0 b_0 (0 \ 0 \ 1) + a_0 b_1 (0 \ 1 \ 0) + \\ &+ a_0 b_2 (1 \ 0 \ 0) + a_1 b_0 (0 \ 1 \ 0) + a_1 b_1 (1 \ 0 \ 0) + \\ &+ a_1 b_2 (0 \ 1 \ 2) + a_2 b_0 (1 \ 0 \ 0) + \\ &+ a_2 b_1 (0 \ 1 \ 2) + a_2 b_2 (1 \ 2 \ 0). \end{aligned}$$

Таким образом, компоненты вектора c могут быть вычислены следующим образом:

$$c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0 + a_2 b_2, \quad (7)$$

$$c_1 = a_0 b_1 + a_1 b_0 + a_1 b_2 + a_2 b_1 + 2a_2 b_2, \quad (8)$$

$$c_0 = a_0 b_0 + 2a_1 b_2 + 2a_2 b_1. \quad (9)$$

Очевидно, что выражения (7–9) идентичны выражениям (1–3). Это следует из того, что по своей сути преобразование Ганкеля – Теплица (ГТ-преобразование) является произведением исходного вектора a (или его перестановки $\acute{a}$) на примитивный элемент поля $GF(q^k)$ α [3]:

$$a\Gamma = [(\acute{a})(\acute{a}\alpha) \dots (\acute{a}\alpha^{k-1})]$$

и

$$\acute{a}\Gamma = [(a)(a\alpha) \dots (a\alpha^{k-1})].$$

На рисунке 4 показана структурная схема полевой свертки произвольных элементов поля $GF(3^3)$, определенного в примере 1, демонстрирующая вычисления векторов, заданных в примере 2. Схема синтезирована на основе выражений (1–3). Сложение и умножение выполняются в поле $GF(3^3)$.

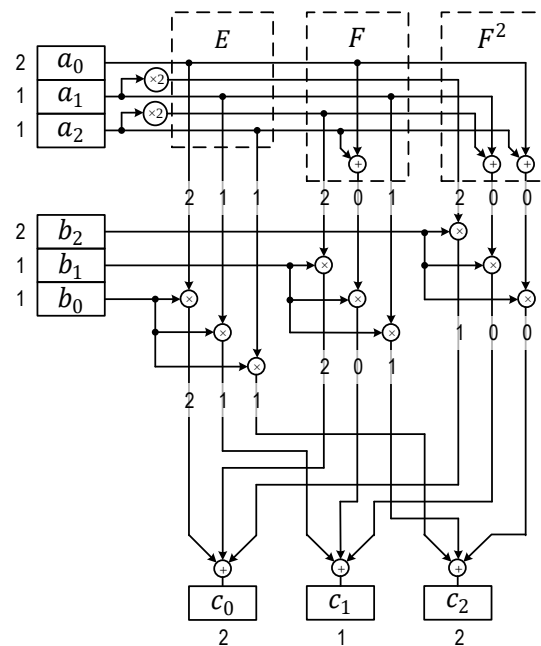


Рис. 4. Функциональная схема матричного метода умножения произвольных элементов поля $GF(3^3)$

Fig. 4. Functional Diagram for the Matrix Multiplication Method for Arbitrary Elements of the Field $GF(3^3)$

Структурная схема (см. рисунок 4) проста с теоретической точки зрения. Однако ее синтез на двоичных логических элементах в программируемых логических интегральных схемах (ПЛИС) или больших интегральных схемах (БИС) для больших q и k приводит к резкому росту цены (стоимости) по Квайну и снижению быстродействия (повышению ранга). Это связано с необходимостью синтеза сумматоров и умножителей в полях $GF(q^k)$ на ос-

нове соответствующих двоичных операций из поля $GF(2^k)$.

Пример 4. Синтез логических операций в поле $GF(3^3)$ на основе элементов логического базиса «И»–«ИЛИ»–«НЕ»

Синтезируем совершенные дизъюнктивные нормальные формы (СДНФ) для частичных булевых функций в поле $GF(3^3)$ из примера 1, реализующих соответствующие операции (см. рисунок 4). Для этого построим таблицу 2.

ТАБЛИЦА 2. Таблица истинности для логических операций в поле $GF(3^3)$

TABLE 2. Truth Table for Logical Operations in the Field $GF(3^3)$

a_i	a_i^1	a_i^0	b_i	b_i^1	b_i^0	$(a_i \oplus b_i)_{\text{mod}3}$			$(a_i \otimes b_i)_{\text{mod}3}$			$(a_i \otimes 2)_{\text{mod}3}$		
						c_i	c_i^1	c_i^0	c_i	c_i^1	c_i^0	c_i	c_i^1	c_i^0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	1	0	1	0	0	0	2	1	0
2	1	0	0	0	0	2	1	0	0	0	0	1	0	1
0	0	0	1	0	1	1	0	1	0	0	0	–	–	–
1	0	1	1	0	1	2	1	0	1	0	1	–	–	–
2	1	0	1	0	1	0	0	0	2	1	0	–	–	–
0	0	0	2	1	0	2	1	0	0	0	0	–	–	–
1	0	1	2	1	0	0	0	0	2	1	0	–	–	–
2	1	0	2	1	0	1	0	1	1	0	1	–	–	–

Получены следующие СДНФ:

– для операции $(a_i \oplus b_i)_{\text{mod}3}$:

$$c_i^1 = \overline{a_i^1 a_i^0 b_i^1 b_i^0} \vee \overline{a_i^1 a_i^0 b_i^1 b_i^0} \vee \overline{a_i^1 a_i^0 b_i^1 b_i^0}, \quad (10)$$

$$c_i^0 = \overline{a_i^1 a_i^0 b_i^1 b_i^0} \vee \overline{a_i^1 a_i^0 b_i^1 b_i^0} \vee \overline{a_i^1 a_i^0 b_i^1 b_i^0}, \quad (11)$$

– для операции $(a_i \otimes b_i)_{\text{mod}3}$:

$$c_i^1 = \overline{a_i^1 a_i^0 b_i^1 b_i^0} \vee \overline{a_i^1 a_i^0 b_i^1 b_i^0}, \quad (12)$$

$$c_i^0 = \overline{a_i^1 a_i^0 b_i^1 b_i^0} \vee \overline{a_i^1 a_i^0 b_i^1 b_i^0}, \quad (13)$$

– для операции $(a_i \times 2)_{\text{mod}3}$:

$$c_i^1 = \overline{a_i^1 a_i^0}, \quad (14)$$

$$c_i^0 = \overline{a_i^1 a_i^0}. \quad (15)$$

Очевидно, что функции (10–13) требуется минимизировать, а функции (14, 15) – не нужно. Для синтеза минимальных дизъюнктивных нормальных форм (МДНФ) построим соответствующие карты Карно (таблицы 3, 4). Карты дополнены неопределенными значениями («–») там, где значения СДНФ (см. таблицу 2) не определены. Карта Карно – это альтернативная форма представления таблицы истинности, которая позволяет механизировать способ минимизации логических функций без применения алгебраических средств [22].

Стоит отметить, что существуют и другие методы минимизации логических функций, например, метод непосредственных преобразований, метод неопределенных коэффициентов, различные аналитические методы, диаграммы Вейча и др.

Метод карт Карно основывается на табличном представлении логических функций. Он используется для ручной минимизации функций с числом переменных, не превышающих шести. Данный метод выбран ввиду своей простоты и наглядности.

ТАБЛИЦА 3. Карта Карно для операции сложения в поле $GF(3^3)$

TABLE 3. Karnaugh Map for the Addition Operation in the Field $GF(3^3)$

$(a_i \oplus b_i)_{\text{mod}3}$												
c_i^1	b_i	0	1	3	2		c_i^0	b_i	0	1	3	2
a_i	$b_i^1 b_i^0, a_i^1 a_i^0$	00	01	11	10		a_i	$b_i^1 b_i^0, a_i^1 a_i^0$	00	01	11	10
0	00	0	0	–	1		0	00	0	1	–	0
1	01	0	1	–	0		1	01	1	0	–	0
3	11	–	–	–	–		3	11	–	–	–	–
2	10	1	0	–	0		2	10	0	0	–	1

ТАБЛИЦА 4. Карта Карно для операции умножения в поле $GF(3^3)$

TABLE 4. Karnaugh Map for the Multiplication Operation in the Field $GF(3^3)$

$(a_i \otimes b_i)_{\text{mod}3}$												
c_i^1	b_i	0	1	3	2		c_i^0	b_i	0	1	3	2
a_i	$b_i^1 b_i^0, a_i^1 a_i^0$	00	01	11	10		a_i	$b_i^1 b_i^0, a_i^1 a_i^0$	00	01	11	10
0	00	0	0	–	0		0	00	0	0	–	0
1	01	0	0	–	1		1	01	0	1	–	0
3	11	–	–	–	–		3	11	–	–	–	–
2	10	0	1	–	0		2	10	0	0	–	1

Сумма полученных произведений образует МДНФ:

$$c_i^1 = a_i^0 b_i^0 \vee \overline{a_i^1 a_i^0 b_i^1} \vee a_i^1 \overline{b_i^1 b_i^0}, \quad (16)$$

$$c_i^0 = a_i^1 b_i^1 \vee \overline{a_i^1 a_i^0 b_i^0} \vee a_i^0 \overline{b_i^1 b_i^0}, \quad (17)$$

$$c_i^1 = a_i^0 b_i^1 \vee a_i^1 b_i^0, \quad (18)$$

$$c_i^0 = a_i^0 b_i^0 \vee a_i^1 b_i^1. \quad (19)$$

На основе выражений (14–19) получим следующие схемы для булева базиса «И»–«ИЛИ»–«НЕ» (рисунки 5–7).

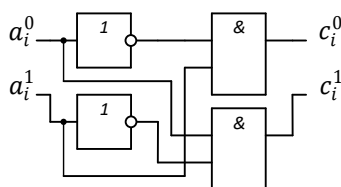


Рис. 5. Функциональная схема умножения произвольных элементов поля $GF(3^3)$ на 2

Fig. 5. Functional Diagram for Multiplying Arbitrary Elements of the Field $GF(3^3)$ by 2

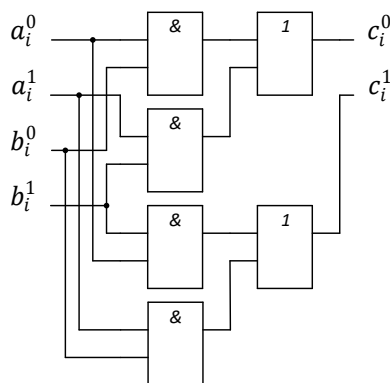


Рис. 6. Структурная схема умножения произвольных элементов поля $GF(3^3)$

Fig. 6. Block Diagram of Multiplication of Arbitrary Elements of the Field $GF(3^3)$

В таблице 5 приведены значения цены по Квайну и ранга схем, показанных на рисунках 5–7 [23]. За сложность по Квайну принято суммарное количество входов логических элементов. Рангом схемы обозначено максимальное количество логических элементов, через которые проходит сигнал от выхода ко входу.

ТАБЛИЦА 5. Значения цены по Квайну и ранга для логических схем на рисунках 5–7

TABLE 5. Quine Price and Rank Values for the Logic Circuits in Figures 5–7

Выражение	Цена по Квайну	Ранг схемы
$(a_i \otimes 2)_{\text{mod } 3}$	6	2
$(a_i \otimes b_i)_{\text{mod } 3}$	12	2
$(a_i \oplus b_i)_{\text{mod } 3}$	26	3

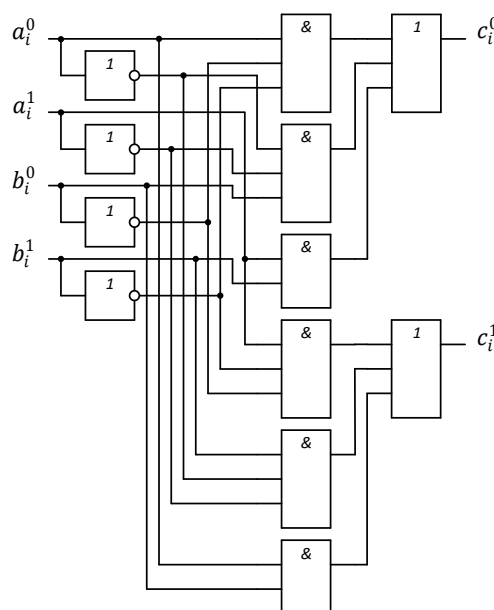


Рис. 7. Структурная схема сложения произвольных элементов поля $GF(3^3)$

Fig. 7. Block Diagram of the Addition of Arbitrary Elements of the Field $GF(3^3)$

Метод на основе разложения многозначного расширенного поля Галуа

Анализ схем (рисунки 2 и 4) позволяет сделать вывод, что этап умножения вектора b_i на составную матрицу A может быть оптимизирован. Неприводимые многочлены $P(x)$ в полях $GF(q^k)$ для $q \geq 3$ задают M -последовательности, обладающие свойством деления на подпоследовательности длины $(q^k - 1)/(q - 1)$, связанные друг с другом коэффициентом пропорциональности λ (λ – первообразный элемент поля $GF(q)$) [2, 24].

Например, из таблицы 1 видно, что для $GF(3^3)$ по основанию полинома $P(x) = x^3 + 2x + 1$ элементы $\varepsilon^{13}, \varepsilon^{14}, \dots, \varepsilon^{25}$ могут быть получены из элементов $\varepsilon^0, \varepsilon^1, \dots, \varepsilon^{12}$ умножением их на коэффициент $\lambda = 2$.

Таким образом, поле $GF(q^k)$ может рассматриваться как совокупность $(q - 1)$ подмножеств $\{M_0, M_1, \dots, M_{q-2}\}$ размера $(q^k - 1)/(q - 1)$ каждое (20), где: $M_0, M_1, \dots, M_r, \dots, M_{q-2}$ – подмножества, состоящие из $(q^k - 1)/(q - 1)$ подряд следующих элементов поля $GF(q^k)$, с номерами $r, r + 1, \dots, r + (q^k - 1)/(q - 1)$ ($r = 0, 1, \dots, q - 2$).

Анализ матричного способа умножения произвольных элементов $a = (a_0 \dots a_{k-1}) = \varepsilon^i$ и $b = (b_0 \dots b_{k-1}) = \varepsilon^j$ поля $GF(q^k)$ для левого степенного базиса позволяет сформулировать [2] выражение (21), соответствующее функциональной схеме умножения произвольных элементов поля $GF(q^k)$, которая показана в качестве примера на рисунке 4. Из (21) следует выражение (22), где слагаемые c_i^h элементов c_i вектора c являются произведениями скаляров b_h и $a_i \varepsilon^h$ ($l, h = 0, \dots, k - 1$) в поле $GF(q)$.

$$GF(q^k) = \left\{ \begin{matrix} \varepsilon^0 \\ \varepsilon^1 \\ \vdots \\ \varepsilon^{q^k-2} \end{matrix} \right\} = \left\{ \begin{matrix} \left\{ \begin{matrix} \lambda^0 \varepsilon^0 \\ \lambda^0 \varepsilon^1 \\ \vdots \\ \lambda^0 \varepsilon^{\frac{q^k-1}{q-1}-1} \end{matrix} \right\} \\ \left\{ \begin{matrix} \lambda^1 \varepsilon^0 \\ \lambda^1 \varepsilon^1 \\ \vdots \\ \lambda^1 \varepsilon^{\frac{q^k-1}{q-1}-1} \end{matrix} \right\} \\ \vdots \\ \left\{ \begin{matrix} \lambda^{q-2} \varepsilon^0 \\ \lambda^{q-2} \varepsilon^1 \\ \vdots \\ \lambda^{q-2} \varepsilon^{\frac{q^k-1}{q-1}-1} \end{matrix} \right\} \end{matrix} \right\} = \left\{ \begin{matrix} M_0 \\ M_1 \\ \vdots \\ M_{q-2} \end{matrix} \right\} = \left\{ \begin{matrix} \lambda^0 M_0 \\ \lambda^1 M_0 \\ \vdots \\ \lambda^{q-2} M_0 \end{matrix} \right\}. \quad (20)$$

$$\begin{aligned} c &= (c_0 \quad \dots \quad c_{k-1}) = \varepsilon^{i+j} = \{ab\}_{GF(q^k)} = \{aF^j\}_{GF(q)} = \\ &= \{(a_0 \quad \dots \quad a_{k-1})(b_0 F^0 + b_1 F^1 + \dots + b_{k-1} F^{k-1})\}_{GF(q)} = \\ &= \{b_0(a_0 \varepsilon^0 \quad \dots \quad a_{k-1} \varepsilon^{k-1}) + \dots + b_{k-1}(a_0 \varepsilon^{k-1} \quad \dots \quad a_{k-1} \varepsilon^{2k-2})\}_{GF(q)}. \end{aligned} \quad (21)$$

$$\begin{aligned} c_0 &= \{b_0 a_0 \varepsilon^0 + b_1 a_0 \varepsilon^1 + \dots + b_{k-1} a_0 \varepsilon^{k-1}\}_{GF(q)}, \quad c_1 = \{b_0 a_1 \varepsilon^1 + b_1 a_1 \varepsilon^2 + \dots + b_{k-1} a_1 \varepsilon^k\}_{GF(q)}, \dots \\ \dots, c_{k-1} &= \{b_0 a_{k-1} \varepsilon^{k-1} + b_1 a_{k-1} \varepsilon^k + \dots + b_{k-1} a_{k-1} \varepsilon^{2k-2}\}_{GF(q)}. \end{aligned} \quad (22)$$

Учитывая свойство (20), слагаемое c_l^h может быть представлено следующим образом:

$$\begin{aligned} c_l^h &= i(b_h, \lambda) \begin{pmatrix} a_l \{\varepsilon^{h_0}\}_{M_0} \\ a_l \{\varepsilon^{h_1}\}_{M_1} \\ \vdots \\ a_l \{\varepsilon^{h_{q-2}}\}_{M_{q-2}} \end{pmatrix} = \\ &= i(b_h, \lambda) \begin{pmatrix} \lambda^0 a_l \{\varepsilon^h\}_{M_0} \\ \lambda^1 a_l \{\varepsilon^h\}_{M_0} \\ \vdots \\ \lambda^{q-2} a_l \{\varepsilon^h\}_{M_0} \end{pmatrix}. \end{aligned} \quad (23)$$

где $l, h, h_0, h_1, \dots, h_{q-2} = 0, \dots, k-1, i(b_h, \lambda) = (i_0, i_1, \dots, i_{q-2})$ – строка-индикатор, элементы которой определяются в соответствии со следующим индикаторным преобразованием:

$$\begin{aligned} (i_0, \dots, i_j, \dots, i_{q-2}) &= \\ &= \begin{cases} i_0 = i_1 = \dots = i_{q-2} = 0, \text{ если } b_h = 0, \\ i_j = \begin{cases} 1, \text{ если } b_h = \lambda^j, \\ 0, \text{ если } b_h \neq \lambda^j, \end{cases} \text{ если } b_h \neq 0. \end{cases} \end{aligned}$$

Пример 5. Индикаторное преобразование вектора в поле $GF(5^3)$

Поле $GF(5^3)$ по основанию любого примитивного полинома в соответствии со свойством (20) может рассматриваться как совокупность четырех подмножеств. Одним из примитивных элементов поля $GF(5)$ является число $\lambda = 3$ ($\lambda^0 = 1, \lambda^1 = 3, \lambda^2 = 4, \lambda^3 = 2$).

Следовательно, если поле $GF(5^3)$ по основанию соответствующего полинома делится на подмножества по элементу $\lambda = 3$, то значениям элементов b_h вектора b будут соответствовать следующие строки-индикаторы $i(b_h, \lambda)$:

$$\begin{aligned} i(0,5) &= (0,0,0,0), i(1,5) = (1,0,0,0), \\ i(2,5) &= (0,0,0,1), i(3,5) = (0,1,0,0), i(4,5) = (0,0,1,0). \end{aligned}$$

Другим примитивным элементом поля $GF(5)$ является число $\lambda = 2$. Для случая разбиения поля $GF(5^3)$ на подмножества по элементу $\lambda = 2$ индикаторные вектора вычисляются соответствующим образом.

Таким образом, предлагаемый метод (20, 23) позволяет при вычислении полевой свертки перейти от операции умножения произвольных элементов по модулю q к операции индикаторного выбора, которая при синтезе логической схемы может быть реализована с помощью элемента «И». Такой переход требует реализации в синтезируемой схеме, кроме матриц $E = F^0 = \lambda^0 F^0, F^1 = \lambda^0 F^1, F^2 = \lambda^0 F^2, \dots, F^{k-1} = \lambda^0 F^{k-1}$, еще и матриц $\lambda^1 F^0, \lambda^1 F^1, \dots, \lambda^1 F^{k-1}, \lambda^2 F^0, \lambda^2 F^1, \dots, \lambda^2 F^{k-1}, \dots, \lambda^{q-2} F^0, \lambda^{q-2} F^1, \dots, \lambda^{q-2} F^{k-1}$. Стоит отметить, что такой подход при должной оптимизации приводит лишь к $q-1$ кратному росту числа умножителей элементов поля $GF(q)$ на константу. Пример функциональной схемы умножения произвольных элементов поля $GF(3^3)$ показан на рисунке 8. Исходные данные взяты из примеров 1 и 2. Примитивный элемент поля $GF(3)$ – $\lambda = 2$. Строки-индикаторы $i(0,2) = (0,0), i(1,2) = (1,0), i(2,2) = (0,1)$ получены в соответствии с выражением (23).

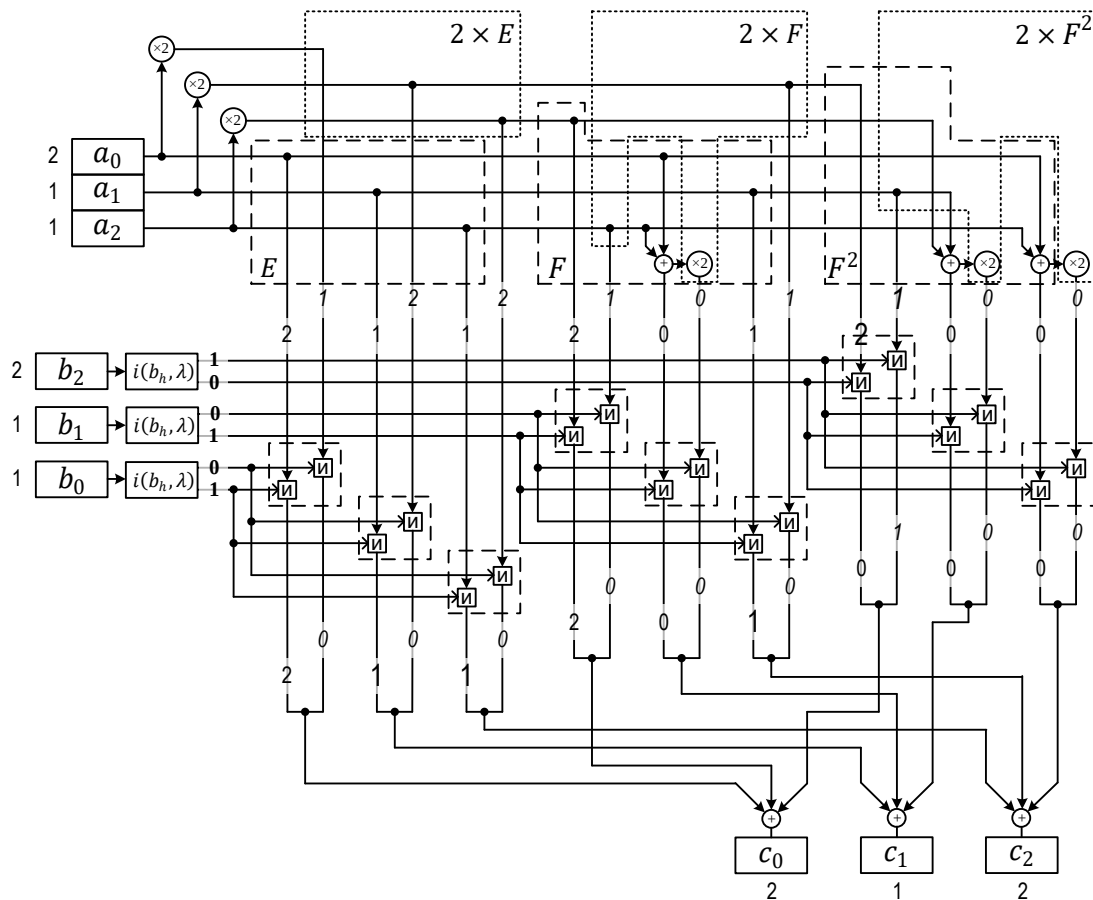


Рис. 8. Функциональная схема предлагаемого метода умножения произвольных элементов поля $GF(3^3)$

Fig. 8. Functional Diagram for the Proposed Method of Multiplying Arbitrary Elements from the Field $GF(3^3)$

Предлагаемая схема полевой свертки (см. рисунок 8) построена по аналогии со схемой, реализующей матричный метод (см. рисунок 4), и работает следующим образом. Элемент поля $GF(3^3)$ $a = (a_0 \ a_1 \ a_2) = \varepsilon^{11} = (2 \ 1 \ 1)$ умножается на матрицы $E, F, F^2, 2 \times E, 2 \times F, 2 \times F^2$.

Из векторов:

$$\begin{aligned} a \times E &= (2 \ 1 \ 1), \quad a \times 2 \times E = (1 \ 2 \ 2), \\ a \times F &= (2 \ 0 \ 1), \quad a \times 2 \times F = (1 \ 0 \ 1), \\ a \times F^2 &= (2 \ 0 \ 0), \quad a \times 2 \times F^2 = (1 \ 0 \ 0) \end{aligned}$$

с помощью строк-индикаторов:

$$\begin{aligned} i(b_0 = 1, 2) &= (1, 0), \quad i(b_0 = 1, 2) = (1, 0), \\ i(b_0 = 2, 2) &= (0, 1), \end{aligned}$$

полученных из вектора $b = \varepsilon^{20} = (1 \ 1 \ 2)$, выбираются вектора:

$$(2 \ 1 \ 1), (2 \ 0 \ 1) \text{ и } (1 \ 0 \ 0),$$

которые затем поступают на соответствующие поразрядные сумматоры. На выходе сумматоров формируется искомый вектор $c = \varepsilon^5 = (2 \ 1 \ 2)$.

Представленные на рисунках 4 и 8 схемы являются функциональными. При построении на их

основе логических схем (БИС или ПЛИС) предполагается, что элементы $\{a_i\}_{GF(q)}$ вектора a будут заменяться двоичными векторами ($GF(2)$) в соответствии с таблицами истинности, построенными по аналогии с таблицей 2. Операции в поле $GF(q)$ (умножение произвольных элементов, умножение произвольного элемента на константу и сложение произвольных элементов) реализуются в поле $GF(2)$ (в логическом базисе «И»–«ИЛИ»–«НЕ») на основе выражений, которые могут быть получены способом, аналогичным способу вывода выражений (10–15), и минимизированы соответствующим образом. В результате могут быть синтезированы соответствующие логические схемы по аналогии с рисунками 5–7.

Стоит отметить, что выражения, аналогичные выражениям (20–23), могут быть выведены на основании ГТ-преобразования. Блок-схема предлагаемого метода с их использованием показана на рисунке 9.

Ввиду аналогичности матричного метода и метода на основе ГТ-преобразования, разработанный метод целесообразно сравнивать с одним из них, например, с матричным методом. Для этого вна-

чае оценим требуемое количество операций в поле $GF(q)$ для элементов векторов в поле $GF(q^k)$. На основе обобщения функциональной схемы (см. рисунок 8) для любых q и k получена следующая сравнительная оценка количества логических операций в поле $GF(q)$ (таблица 6).

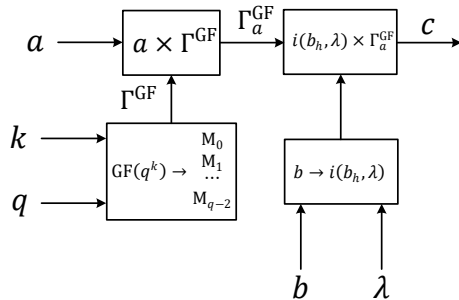


Рис. 9. Блок-схема предлагаемого метода умножения элементов поля $GF(q^k)$ на основе преобразования Ганкеля – Теплица

Fig. 9. Block Diagram for the Proposed Method of Multiplying Elements from the Field $GF(q^k)$ Based on the Hankel – Toeplitz Transform

ТАБЛИЦА 6. Сравнительная оценка количества логических операций в поле $GF(q^k)$ для матричного и предложенного методов

TABLE 6. Comparative Estimation for the Number of Logical Operations in the $GF(q^k)$ Field for the Matrix and Proposed Methods

Метод	$GF(q)$			$GF(2)$
	$a_i \otimes \text{const}$	$a_i \otimes b_i$	$a_i \oplus b_i$	«И» (двухходовый)
Матричный [2]	$\leq (q-2)k$	k^2	$\geq k^2$	–
Разработанный	$\leq 2(q-2)k$	–	$\geq k^2$	$(q-1)k^2$

Из таблицы 6 видно, что разработанный метод рационально использовать для полевой свертки векторов поля $GF(q^k)$ в случаях, когда $k \gg q$. Тогда выигрыш от сокращения числа умножителей превзойдет проигрыш от роста количества сумматоров по модулю q .

В таблице 7 представлены сравнительные оценки стоимости по Квайну и ранга функциональных схем для различных значений k при $q = 3$, полученных на основании следующих выражений:

– матричный метод:

$$6(q-2)k + 12k^2 + 26k^2;$$

– разработанный метод:

$$12(q-2)k + 26k^2 + 2(q-1)k^2.$$

Ранг схемы: матричный метод – $2 + 3 + 2 + 3$; разработанный метод – $2 + 3 + 1 + 3$. Для построения выражений оценки стоимости по Квайну количественные значения стоимости из таблицы 5 для соответствующей операции в поле $GF(q)$ умножены на количество таких операций из таблицы 6. Затем выполнена сумма по всем указан-

ным операциям для соответствующего метода полевой свертки. Ранг схемы вычислен по максимальному пути: вначале определена цепочка логических элементов, затем для каждого элемента подставлено значения ранга из таблицы 5, в конце – выполнена сумма.

ТАБЛИЦА 7. Сравнительная оценка цены по Квайну и ранга схемы операции полевой свертки в поле $GF(3^k)$ для матричного и предложенного методов

TABLE 7. Comparative Evaluation of the Quine Price and Rank of the Field Convolution Operation Scheme in the $GF(3^k)$ Field for the Matrix and Proposed Methods

k	Цена по Квайну			Ранг схемы		
	M_m	M_{GF}	$W, \%$	M_m	M_{GF}	$W, \%$
3	360	306	15,00	33	32	3,03
30	34 380	27 360	20,42			
300	3 421 800	2 703 600	20,99			
3 000	342 018 000	270 036 000	21,05			
30 000	34 200 180 000	27 000 360 000	21,05			
300 000	3 420 001 800 000	2 700 003 600 000	21,05			

Усл. обозначения: M_m – матричный метод; M_{GF} – разработанный метод; W – выигрыш

Из таблицы 7 видно, что для больших k формируется устойчивое соотношение цены по Квайну, разработанное к матричному методу, равное 0,7895. Таким образом для $q = 3$ и $k \gg q$ выигрыш составляет $\approx 21,05 \%$. Из анализа таблицы 6 можно предположить, что с ростом q выигрыш снизится. Анализ зависимости выигрыша от q требует дополнительного изучения, а именно синтеза выражений и функциональных схем для операций в соответствующих полях Галуа.

Заключение

Проведенный анализ однократных методов умножения элементов расширенного поля Галуа позволил выявить целесообразность поиска оптимальных методов вычисления полевых свертки. Логарифмирование–антилогарифмирование имеет ограниченную область применения для больших значений q и k .

Разработан метод вычисления полевой свертки на основе использования свойства разложения многозначного расширенного поля Галуа на подмножества по степени примитивного элемента соответствующего простого поля с использованием индикаторного представления одного из множителей. Доказано сокращение общего числа логических элементов и как следствие цены по Квайну и ранга схемы. Приведены примеры вычисления полевых свертки в многозначных полях Галуа различными методами. Дана сравнительная оценка разработанного метода.

Список источников

1. Рахман П.А. Кодирование информации с применением кодов Рида-Соломона. Уфа: УГНТУ, 2012. 167 с.
2. Когновицкий О.С. Теория, методы и алгоритмы решения задач в телекоммуникациях на основе двойственного базиса и рекуррентных последовательностей. Дис. ... докт. техн. наук. СПб.: СПбГУТ, 2011. 427 с. EDN:QFKYXL
3. Муттер В.М. Основы помехоустойчивой телепередачи информации. Л.: Энергоатомиздат. Ленингр. отд-ние, 1990. 288 с.
4. Когновицкий О.С., Сюрин В.Н., Ассанович Б.А. Метод вычисления логарифма в конечном поле $GF(2^m)$ // Девятая всесоюзная конференция по теории кодирования и передачи информации. Часть 1. Тезисы докладов. Одесса, 1988. С. 100–102.
5. Рахман П.А. Рекуррентный алгоритм вычисления усеченной свертки полиномов над полем Галуа и его аппаратная реализация // Международный журнал прикладных и фундаментальных исследований. 2015. № 12-2. С. 231–235. EDN:SZAEYV
6. Иванова И.В. Научные основы создания автоматизированных систем кодирования данных в конечных полях Галуа методами дискретной алгебры Клини. Дис. ... докт. техн. наук. СПб.: СЗТУ, 2005. 276 с. EDN:NNZFPD
7. Берлекэмп Э.Р. Алгебраическая теория кодирования. Пер. с англ. М.: Мир, 1971. 478 с.
8. Мак-Вильямс Ф.Дж., Спозэн Н.Дж.А. Теория кодов, исправляющих ошибки. Пер. с англ. М.: Связь, 1979. 744 с.
9. Касами Т., Токура Н., Ивадари Е., Ирагаки Я. Теория кодирования. М.: Мир, 1978. 576 с.
10. Рахман П.А. Арифметика двоичного поля Галуа на базе быстрого умножения и инвертирования элементов поля и ее аппаратная реализация // Международный журнал прикладных и фундаментальных исследований. 2015. № 12-3. С. 403–408. EDN:VBUMBj
11. Листопад Е.В., Петровский А.А. Особенности реализации операций умножения элементов поля Галуа на FPGA // 53-я научная конференция аспирантов, магистров и студентов БГУИР (Минск, Республика Беларусь, 2–6 мая 2017 г.). Минск: Белорусский государственный университет информатики и радиоэлектроники, 2017. С. 234–235.
12. Касперски К. Полиномиальная арифметика и поля Галуа, или информация, воскресшая из пепла II // Системный администратор. 2003;10(11):84–90. EDN:RDELQN
13. Салимов Г.Ю. Предложения по реализации умножения в поле Галуа над неприводимым многочленом на примере преобразования L в алгоритме ГОСТ Р 34.1 2 2015 // XXII научно-практическая конференция «РусКрипто 2020» (17–29 марта 2020 г.) 2020. URL: https://ruscrypto.ru/resource/archive/rc2020/files/14_salimov.pdf (дата обращения 17.04.2025)
14. Лидл Р., Нидеррайтер Г. Конечные поля. Пер. с англ. М.: Мир, 1988. 818 с.
15. Габидулин Э.М., Афанасьев В.Б. Кодирование в радиоэлектронике. М.: Радио и связь, 1986. 176 с.
16. Питерсон Ч., Уэлдон Э. Коды, исправляющие ошибки. Пер. с англ. М.: Мир, 1976. 594 с.
17. Кларк Дж., мл., Кейн Дж. Кодирование с исправлением ошибок в системах цифровой связи. Пер. с англ. М.: Радио и связь, 1987. 391 с.
18. Золотарев В.В. Теория и алгоритмы многопорогового декодирования. М.: Радио и связь, 2006. 266 с.
19. Трифонов П.В. Основы помехоустойчивого кодирования. СПб.: Университет ИТМО, 2022. 231 с.
20. Ишмухаметов Ш.Т., Латыпов Р.Х., Столов Е.Л., Рубцова Р.Г. Введение в теорию чисел и теорию кодирования: учебное пособие. Казань: Казанский университет, 2014. 65 с.
21. Ишмухаметов Ш.Т., Рубцова Р.Г. Вычисления в конечных полях: учебно-методическое пособие. Казань: Казанский университет, 2019. 23 с.
22. Назарова А.К., Локтионова О.Е., Спесивцев Г.А. Карты Карно // Международная научно-практическая конференция «Моделирование информационных систем и технологий» (Воронеж, Российская Федерация, 02 апреля 2024 г.). Воронеж: Воронежский государственный лесотехнический университет имени Г.Ф. Морозова, 2024. С. 116–121. EDN:JPIWIW
23. Исмагилова Е.И. Булевы функции и построение логических схем. М.: МИРЭА, 2015. 160 с.
24. Цирлер Н. Линейные возвратные последовательности // Кибернетический сборник. 1963. № 6. С. 31–48.

References

1. Rakhman P. *Information Coding Using Reed-Solomon Codes*. Ufa: Ufa State Petroleum Technical University Publ.; 2012, 167 p. (in Russ.)
2. Kognovitsky O. *Theory, Methods and Algorithms for Solving Problems in Telecommunications Based on a Dual Basis and Recurrent Sequences*. D.Sc Thesis. St. Petersburg: The Bonch-Bruевич Saint-Petersburg State University of Telecommunications Publ.; 2011. 427 p. (in Russ.) EDN:QFKYXL
3. Mutter V. *Fundamentals of Noise-Immune Television Transmission of Information*. Leningrad: Energoatomizdat; 1990. 288 p. (in Russ.)
4. Kognovitsky O.S., Syurin V.N., Assanovich B.A. Method for calculating logarithms in the traditional field $GF(2^m)$. *Proceedings of the 9th All-Union Conference on Theories of Coding and Transmission of Information. Part 1*. Odessa; 1988. p.100–102. (in Russ.)
5. Rakhman P.A. Algorithm for Computing the Truncated Convolution of the Polynomials Over Galois Field and Its Hardware Implementation. *Mezhdunarodnyi zhurnal prikladnykh i fundamentalnykh issledovaniy*. 2015;12-2:231–235. (in Russ.) EDN:SZAEYV
6. Ivanova I.V. *Scientific Foundations for Creating Automated Data Coding Systems in Finite Galois Fields Using Discrete Kleene Algebra Methods*. D.Sc Thesis. St. Petersburg: North-West Technical University Publ; 2005. 276 p. (in Russ.) EDN:NNZFPD

7. Berlekamp E. R. *Algebraic Coding Theory*. 1968.
8. McWilliam F.J., Sloane N.J.A. *The theory of error-correcting codes*. North-Holland Publishing; 1977.
9. Kasami T., Tokura N., Iwadari E., Iragaki Y. *Coding Theory*. Moscow: Mir Publ.; 1978. 576 p. (in Russ.)
10. Rahman P.A. Arithmetic of a binary Galois field based on fast multiplication and inversion of field elements and its hardware implementation. *Mezhdunarodnyi zhurnal prikladnykh i fundamentalnykh issledovaniy*. 2015;12-3:403–408. EDN:VBUMBJ
11. Listopad E.V., Petrovsky A.A. Features of the implementation of operations of multiplication of Galois field elements on FPGA. *Proceedings of the 53rd Scientific Conference of Graduate Students, Masters and Students of BSUIR, 2–6 May 2017, Minsk, Republic of Belarus*. Minsk: Belarusian State University of Informatics and Radioelectronics Publ.; 2017. p.234–235. (in Russ.)
12. Kasperski K. Polynomial arithmetic and Galois fields, or information resurrected from the ashes II. *Sistemnyi administrator*. 2003;10(11):84–90. (in Russ.) EDN:RDELQN
13. Salimov G.Yu. Proposals for the implementation of multiplication in a Galois field over an irreducible polynomial using the example of transformation L in the GOST R 34.1 2 2015 algorithm. *Proceedings of the XXIIInd Scientific and Practical Conference "RusCrypto 2020", 17–29 March 2020*. 2020. (in Russ.) URL: https://ruscrypto.ru/resource/archive/rc2020/files/14_salimov.pdf [Accessed 17.04.2025]
14. Lidl R., Niederreiter G. *Finite fields*. London: Addison-Wesley Publishing; 1983.
15. Gabidullin E.M., Afanasyev V.B. *Coding in Radio Electronics*. Moscow: Radio i sviaz Publ.; 1986. 176 p. (in Russ.)
16. Peterson W.W., Weldon-jr E.J. *Error-Correcting Codes*. MIT Press; 1972. 575 p.
17. Clark J., Jr., Kane J. *Error-Correcting Coding for Digital Communication*. New York: Plenum Press; 1981.
18. Zolotarev V.V. *Theory and Algorithms of Multithreshold Decoding*. Moscow: Radio i sviaz Publ.; 2006. 266 p. (in Russ.)
19. Trifonov P.V. *Fundamentals of Noise-Corrective Coding*. St. Petersburg: ITMO University Publ.; 2022. 231 p. (in Russ.)
20. Ishmukhametov Sh.T., Latypov R.Kh., Stolov E.L., Rubtsova R.G. *Introduction to Number Theory and Coding Theory*. Kazan: Kazan University Publ.; 2014. 65 p. (in Russ.)
21. Ishmukhametov Sh.T., Rubtsova R.G. *Calculations in Finite Fields*. Kazan: Kazan University Publ.; 2019. 23 p. (in Russ.)
22. Nazarova A., Loktionova O., Spesivtsev G. Carnot Cards. *Proceedings of the International Scientific and Practical Conference on Modeling of Information Systems and Technologies, 02 April 2024, Voronezh, Russian Federation*. Voronezh: Voronezh State Forestry University Publ.; 2024. p.116–121. (in Russ.) EDN:JPIWIW
23. Ismagilova E.I. *Boolean Functions and Construction of Logical Circuits*. Moscow: MIREA – Russian Technological University Publ.; 2015. 160 p.
24. Zierler N. Linear recurring sequences. *Journal of the Society for Industrial and Applied Mathematics*. 1959;7(1):31–48. (in Russ.)

Статья поступила в редакцию 09.12.2024; одобрена после рецензирования 19.03.2025; принята к публикации 20.03.2025.

The article was submitted 09.12.2024; approved after reviewing 19.03.2025; accepted for publication 20.03.2025.

Информация об авторе:

УЛЬЯНОВ
Илья Владимирович

кандидат технических наук, сотрудник Академии Федеральной службы охраны
Российской Федерации
 <https://orcid.org/0009-0009-7607-7868>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.

План издания научной литературы 2025 г., п. 10

Усл.-печ. л.
15,0

Формат
60×84_{1/8}

Заказ
№ 1627

Учредитель и издатель:

Федеральное государственное бюджетное образовательное учреждение
высшего образования "Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М.А. Бонч-Бруевича"

E-mail: tuzs@sut.ru Web: tuzs.sut.ru VK: vk.com/spbtuzs

