

Том 10. № 4
2024

ISSN 1813-324X (Print)
ISSN 2712-8830 (Online)

ТРУДЫ УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Темы номера:

- ✓ Позиционирование устройств в сетях шестого поколения 6G
- ✓ Пеленгация для БПЛА с применением виртуальных магнитных диполей
- ✓ Маршрутизация в сети динамических туманных вычислений

Vol. 10. Iss. 4
2024

PROCEEDINGS
OF TELECOMMUNICATION UNIVERSITIES

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 10. № 4

Proceedings of Telecommunication Universities

Vol. 10. Iss. 4

Санкт-Петербург

2024

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-Bruевич 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, +79643759970

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.**

Publisher: Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruевич 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, +79643759970

E-mail: tuzs@sut.ru

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

Executive Editor **Tatarnikova I.M.**
Commissioning Editor **Yashugin D.N.**
Design: **Korovin V.M.**

ВЫХОДНЫЕ ДАННЫЕ / IMPRINT

Дата выхода в свет: 06.09.2024
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: 06.09.2024
Circulation: 1000 copies. Free price.

Printed in the printing office

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



СОДЕРЖАНИЕ

CONTENTS

ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ

Березкин А.А., Вивчарь Р.М., Киричек Р.В.
Подход к обоснованию требований
к качеству видеопотока при FPV-управлении
беспилотными системами

7

Berezkin A.A., Vivchar R.M., Kirichek R.V.
Approach to video stream quality requirements
justification for FPV control of unmanned systems

Былина М.С., Глаголев С.Ф., Доценко С.Э.
Анализ методов управления дисперсией
для поддержания квазисолитонного режима
распространения импульсов в высокоскоростной
волоконно-оптической системе связи

16

Bylina M.S., Glagolev S.F., Dotsenko S.E.
Method for controlling dispersion in order
to maintain a quasi-solution pulse propagation
mode in a high-speed fiber-optic
communications system

Волков А.Н.
Задача маршрутизации в сети динамических
туманных вычислений

27

Volkov A.N.
Routing task in dynamic fog computing network

Елагин В.С., Гребенщикова А.А.
Прогнозирование трафика трехмерной сети
Интернета вещей высокой плотности
как многомерного случайного процесса

38

Elagin V.S., Grebenshchikova A.A.
Traffic prediction as a multidimensional random
process in a three-dimensional high-density
Internet of things network

**Ищенко Е.А., Пастернак Ю.Г.,
Фёдоров С.М., Баранников И.А.**
Реализация высокоомобильного комплекса
пеленгации для БПЛА с применением
виртуальных магнитных диполей

48

**Ishchenko E.A., Pasternak Y.G.,
Fedorov S.M., Barannikov I.A.**
Implementation of a highly mobile UAVs
direction finding complex
using virtual magnetic dipoles

**Кузнецов К.А., Парамонов А.И.,
Мутханна А.С.А., Кучерявый А.Е.**
Модель и методы маршрутизации трафика
в сети связи с использованием БПЛА

62

**Kuznetsov K.A., Paramonov A.I.,
Muthanna A.S.A., Koucheryavy A.E.**
Model and methods of traffic routing
in a communication networks using UAVs

Фокин Г.А.
Комплекс моделей позиционирования
устройств в сетях шестого поколения.
Часть 1. Обзор методов и постановка задачи

73

Fokin G.A.
A set of models for device positioning
in sixth generation networks.
Part 1. Methods survey and problem statement

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

**Астахова Т.Н., Колбанёв М.О., Лямин А.С.,
Маслов Н.С., Маслова Д.А.**
Энергоэффективный алгоритм выбора
маршрута передачи данных в беспроводных
сенсорных системах высокой плотности

100

**Astakhova T.N., Kolbanev M.O., Lyamin A.S.,
Maslov N.S., Maslova D.A.**
Energy-efficient algorithm for data path selection
in high-density wireless sensor networks

**Грызунов В.В., Крюков А.С.,
Шестаков А.В., Зикратов И.А.**
Обеспечение информационной безопасности
интегрируемых информационных систем
на основе доверия

110

**Gryzunov V.V., Krjukov A.S.,
Shestakov A.V., Zikratov I.A.**
Ensuring information security of information
systems to be integrated based on trust

**Коржик В.И., Яковлев В.А., Изотов Б.В.,
Старостин В.С., Буйневич М.В.**
Прогресс в теории прикладной криптографии:
обзор и некоторые новые результаты.
Часть 1. Ключевая криптография

126

**Korzhik V.I., Yakovlev V.A., Izotov B.V.,
Starostin V.S., Buinevich M.V.**
Advance in applied cryptography theory:
Survey and new results.
Part 1. Key cryptography

Самарин Н.Н., Тулинова А.В.
Интеллектуальный метод мутации входных
корпусов с обратной связью

142

Samarin N.N., Tulinova A.V.
Intelligent method for mutation of input cases
with feedback

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 004.05

<https://doi.org/10.31854/1813-324X-2024-10-4-7-15>

Подход к обоснованию требований к качеству видеопотока при FPV-управлении беспилотными системами

Александр Александрович Березкин✉, berezkin.aa@sut.ru

Роман Михайлович Вивчарь, vivchar.rm@sut.ru

Руслан 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 в различных отраслях экономики РФ играют гибридные орбитально-наземные сети связи, важной составной частью которых являются беспилотные системы на FPV-управлении. Важным элементом таких систем являются каналы информационного обмена, в составе которых широкое распространение получило использование нейросетевых кодеков. Главным критерием успешности решения ими своих целевых задач является удовлетворение требований к качеству восстановленного видеопотока, что обуславливает особенную **актуальность** обоснования этих требований.

Целью настоящей статьи является представление подхода к количественному обоснованию требований к качеству передаваемого видеопотока.

Сущность представленного подхода заключается в том, что требуемые значения показателей качества передаваемого изображения, используемые для обоснования приемлемых нейросетевых кодеков, определяются путем анализа различных видеопотоков, на основе которых формировались воздействия по управлению беспилотными системами, позволившие достичь целей их функционирования. Рассмотрены основные этапы подхода и их логическая взаимосвязь.

Предложенный подход базируется на использовании **методов** статистического и риск-анализа, теории планирования эксперимента и теории вероятностей.

Научная новизна предложенного подхода заключается в том, что требования к качеству передаваемого от беспилотной системы к оператору видеопотока, используемые для обоснования приемлемых нейросетевых кодеков, рассчитываются путем анализа совокупности видеопотоков, позволивших достичь целей ее функционирования, что дает возможность избавиться от субъективизма присущего использующимся в настоящее время для решения этой задачи экспертным методом.

Теоретическая значимость: доказана возможность использования имитационного моделирования каналов информационного обмена для решения задачи коррекции требований к показателям качества восстановленного видеопотока при FPV-управлении. Предложенный подход также вносит вклад в совершенствование научно-методического аппарата в области проектирования альтернативных кодеков сжатия видеопотока с потерями.

Практическая значимость предложенного подхода заключается в том, что полученные с помощью него требования к показателям качества передаваемого видеопотока **могут быть в дальнейшем использованы** для обоснования оптимальных проектных решений по созданию каналов информационного обмена между беспилотными системами и оператором, что существенно повысит эффективность использования этих систем.

Ключевые слова: гибридные сети, беспилотные средства, FPV-управление, качество восстановленного видеопотока, риск, метод ядерной оценки плотности, требования, нейросетевые кодеки

Источник финансирования: статья подготовлена в рамках прикладных научных исследований СПбГУТ, регистрационный номер 1023031600087-9-2.2.4;2.2.5;2.2.6;1.2.1;2.2.3 в ЕГИСУ НИОКТР.

Ссылка для цитирования: Березкин А.А., Вивчарь Р.М., Киричек Р.В. Подход к обоснованию требований к качеству видеопотока при FPV-управлении беспилотными системами // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 7–15. DOI:10.31854/1813-324X-2024-10-4-7-15. EDN:JQQCXK

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-7-15>

Approach to Video Stream Quality Requirements Justification for FPV Control of Unmanned Systems

 Aleksandr A. Berezkin✉, berezkin.aa@sut.ru Roman M. Vivchar, vivchar.rm@sut.ru Ruslan V. Kirichuk, kirichuk@sut.ru

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

Annotation

Currently, hybrid orbital-ground communication networks play a key role in various sectors of the Russian economy, an important part of which are unmanned systems based on FPV control. One of the key elements of such systems is information exchange channels. The main criterion for the success of their objectives is to meet the requirements for the quality of the transmitted video stream, which makes it particularly **relevant** to substantiate these requirements.

The purpose of this article is to present an approach to the quantitative justification of the requirements for the quality of the transmitted video stream.

The essence of the presented approach lies in the fact that the required values of the transmitted image quality indicators, used to justify acceptable neural network codecs, are determined by analyzing various video streams, on the basis of which the impacts on the control of unmanned systems that allowed to achieve the goals of their functioning were formed. The main stages of the approach and their logical interrelation are considered.

The proposed approach is based on the use of **methods** of statistical and risk analysis, the theory of experiment planning and probability theory.

Scientific novelty of the proposed approach lies in the fact that the requirements to the quality of the video stream transmitted from the unmanned system to the operator, used to justify acceptable neural network codecs, are calculated by analyzing the totality of video streams that allowed to achieve the goals of its functioning, which allows to get rid of the subjectivism inherent in the expert methods currently used to solve this problem.

Theoretical significance of the proposed approach lies in the fact that the necessity and possibility of not only substantiating the requirements for the quality indicators of the transmitted video stream for FPV control, but also their correction has been proved.

Practical significance. The requirements for the quality indicators of the transmitted video stream obtained using the proposed approach can be further used to justify optimal design solutions for creating information exchange channels between unmanned systems and the operator, which will significantly increase the efficiency of using these systems.

Keywords: hybrid networks, unmanned vehicles, FPV control, the quality of the restored video stream, risk, nuclear probability density estimation method, requirements, neural network codecs

Funding: The scientific article was prepared within the framework of applied scientific research SPbSUT, registration number 1023031600087-9-2.2.4;2.2.5;2.2.6;1.2.1;2.2.3 in the information system (<https://www.rosrid.ru/information>).

For citation: Berezkin A.A., Vivchar R.M., Kirichuk R.V. Approach to Video Stream Quality Requirements Justification for FPV Control of Unmanned Systems. *Proceedings of Telecommunication Universities*. 2024;10(4): 7–15. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-7-15. EDN:JQQCXK

Введение

В настоящее время все более существенную роль в различных отраслях экономики РФ играют

цифровые системы и сервисы. Это вызывает рост требований к услугам связи, особенно в части покрытия, скорости, задержки и надежности переда-

чи данных. Для удовлетворения названных требований активно внедряются новые поколения технологий связи, одним из которых являются создаваемые гибридные орбитально-наземные сети связи (ГОНСС) [1, 2]. В силу своих технологических особенностей и взаимодействия с наземной инфраструктурой связи такие спутниковые и орбитально-наземные системы связи в перспективе окажут существенное влияние на рынки спутниковой, фиксированной и мобильной связи не только в России, но и во всем мире, в соответствии со Стратегией развития отрасли связи Российской Федерации до 2035 г. (распоряжение Правительства РФ от 24 ноября 2023 г. № 3339-р).

Важным элементом ГОНСС выступают различные беспилотные средства (БС), основным назначением которых является сбор информации и осуществление различных действий за пределами прямой видимости оператора [3]. Использование глобальных сетей обеспечивает управление БС на больших расстояниях, что особенно актуально в различных сферах народного хозяйства, например, в сельскохозяйственной отрасли, а также для решения задач МЧС.

Так как задержка в каналах передачи связи при использовании ГОНСС соизмерима с задержкой в мобильных сетях связи, что подтверждается результатами испытаний Бюро 1440 своей низкоорбитальной спутниковой группировки, то широкое распространение нашли БС, оснащенные системой FPV-управления (*аббр. от англ. First Person View* – вид от первого лица) [4].

Одним из ключевых элементов таких систем являются каналы информационного обмена (КИО), назначением которых является передача видеопотока от БС к оператору, а также команд управления. Следовательно, объектом представленных исследований является процесс информационного обмена при FPV-управлении БС. Одним из элементов КИО являются нейросетевые кодеки (НК), предназначенные для сжатия и восстановления передаваемого видеопотока. Эффективность решения КИО своих целевых задач определяется успешностью их проектирования, одной из ключевых задач которого является выбор приемлемого НК. При этом главным критерием является удовлетворение требований к качеству восстановленных после сжатия и передачи изображений, составляющих видеопоток. Следовательно, в процессе проектирования КИО при выборе приемлемого НК одной из ключевых задач является обоснование требований к изменению качества восстановленного видеопотока.

Целью настоящей статьи является представление методики, позволяющей на количественном уровне оценить требования к изменению качества восстановленного видеопотока, что позволит осу-

ществить выбор приемлемого варианта НК и, в конечном итоге, повысит эффективность процесса проектирования БС на FPV-управлении.

Содержание и этапы подхода к обоснованию требований к качеству восстановленного видеопотока при FPV-управлении

В результате анализа критериев эффективности КИО при FPV-управлении было установлено, что наиболее часто используемыми показателями качества восстановленного видеопотока являются [5, 6]:

- показатель структурного сходства (SSIM, *аббр. от англ. Structure Similarity*), характеризующий тенденции к общности между исходным и сжатым изображением; достоинством критерия является учет не только яркости и цвета изображения, но также его текстуры и контрастности; его целесообразно использовать в качестве показателя, характеризующего пофрагментное сходство исходного и полученного кадров для оценки изменения качества кадров видеопотока при передаче его по каналам информационного обмена; это необходимо для определения требований к проектируемым видеокodeкам, используемым в задаче FPV-управления;

- Peak Signal-to-Noise Ratio (PSNR), характеризующий отношение сигнала к шуму между двумя изображениями [7];

- вероятность сбоя.

При этом, говоря о показателях качества восстановленного видеопотока в предлагаемой статье, понимаются критерии, которые характеризуют изменение качества исходного видеопотока при его прохождении по каналам информационного обмена (кодирование, декодирование, непосредственная передача по каналу).

Представленные выше показатели позволяют провести объективное тестирование, главным достоинством которого является возможность автоматизировать этот процесс, что дает возможность оценить качество изменения исходного видеопотока при большом количестве вариантов используемых НК. В качестве требуемого значения вероятности сбоя целесообразно использовать величину, обратную доверительной вероятности отсутствия сбоев при передаче видеопотока, которая может быть рассчитана при заданном уровне, принимающем решения, уровне значимости. Следовательно, говоря об обосновании требований к показателям качества восстановленного видеопотока, в дальнейшем будем понимать под этим обоснование требуемых значений SSIM и PSNR.

Сущность рассматриваемого подхода заключается в анализе различных видеопотоков, на основе которых осуществлялось управление БС, позволившее достичь целей его функционирования, а

также – в обосновании по результатам проведенного выше анализа требуемых значений показателей качества восстановленного видеопотока от БС к оператору. Основные этапы подхода и их логическая взаимосвязь представлена на рисунке 1.

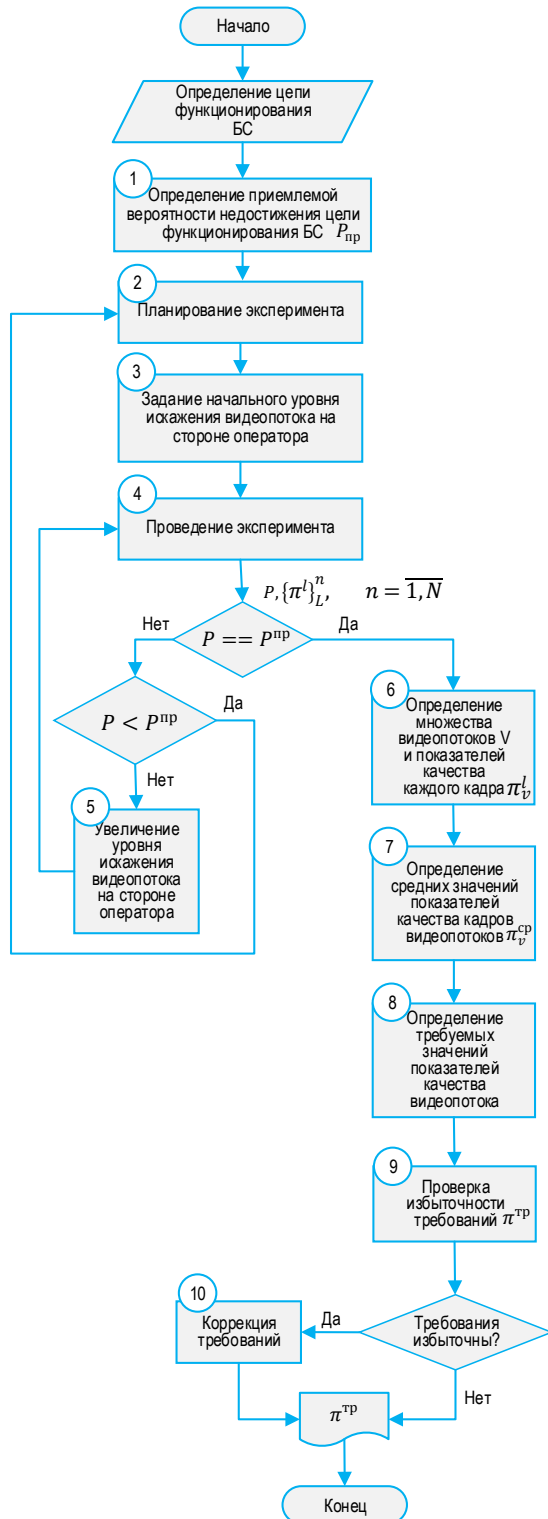


Рис. 1. Структурно-логическая схема подхода к обоснованию требований к качеству передаваемого видеопотока

Fig. 1. Structure and Logic Diagram of the Approach to Substantiation of Requirements to Transmitted Video Stream Quality

На этапе обоснования исходных данных для расчета требуемых значений показателей качества восстановленного видеопотока при FPV-управлении определяется цель функционирования БС, которая выступает в данном случае целью суперсистемы, которая представляет собой систему его функционирования. Например, прохождение БС динамического маршрута между разными точками пространства за установленное время, отыскание с помощью БС определенных объектов в течение установленного времени и другие. От точности определения цели функционирования БС зависит в конечном итоге адекватность обоснования требований к качеству восстановленного видеопотока, а, следовательно, успешность решения задачи выбора нужного варианта НК.

Задачей *первого этапа* является определение приемлемой вероятности недостижения цели функционирования БС, подобранной ранее, и сопутствующего при этом ущерба. Данные показатели выступают в качестве показателей риска недостижения цели функционирования БС [8]. Для определения этих показателей могут быть использованы различные методы, например, представленные в [10]. В данной статье предлагается использовать широко известный метод, основанный на вероятностной оценке показателей риска (PRA, аббр. от англ. Probabilistic Risk Assessment) [9, 10], в соответствии с которым показатели риска недостижения цели функционирования БС определяются путем выполнения последовательных действий:

- идентификации всех сценариев недостижения цели (I);
- определении вероятности появления каждого i -го сценария P_i и возможных последствий S_i .

Вероятности реализации сценариев определяются путем перемножения вероятности появления инициирующего нежелательного фактора на вероятности появления последующих нежелательных факторов при условии того, что предыдущие события уже наступили:

$$P_i = p(A_i^1)p(A_i^2|A_i^1)\dots p(A_i^{J_i}|A_i^1\dots A_i^{J_i-1}),$$

где $p(A_i^1)$ – вероятность появления инициирующего нежелательного фактора; J_i – количество нежелательных факторов, входящих в i -й сценарий; $p(A_i^2|A_i^1), \dots, p(A_i^{J_i}|A_i^1\dots A_i^{J_i-1})$ – условные вероятности появления последующих нежелательных факторов.

Определение вероятностей появления нежелательных факторов (джиттера, сетевой задержки, потери пакетов), в том числе и условных, предлагается проводить на основе анализа статистических данных или с помощью экспертных методов.

Вероятность недостижения цели функционирования БС вычисляется путем суммирования показателей риска всей совокупности сценариев:

$$P_{\text{пр}} = \sum_{i=1}^I P_i.$$

Необходимость оценивания приемлемых показателей риска недостижения цели функционирования БС связана с тем, что этот процесс подвержен влиянию различных факторов неопределенности, и, следовательно, это целесообразно учитывать при обосновании требований к показателям качества видеопотока.

Вторым этапом обоснования требований к показателям качества восстановленного видеопотока является планирование эксперимента, воспроизводящего процесс достижения цели функционирования БС. Такой эксперимент должен включать в себя как натурную часть, так и вычислительную. Натурная часть эксперимента заключается в воссоздании реальных условий для функционирования БС. Предлагается для этого разработать специальный динамически меняющийся маршрут прохождения БС от одной точки к другой в условиях воздействия различных неблагоприятных факторов, например, таких как появление на маршруте различных препятствий, изменение освещения и т. д. Успешным достижением цели функционирования БС в этом случае будет безаварийное прохождение маршрута в течение заданного времени. Вычислительная часть эксперимента заключается в имитации того же самого маршрута с помощью средств вычислительной техники. Для этого предлагается использовать имитационную модель, представленную в [7]. Данная модель будет в дальнейшем необходима для проверки корректности задания требований к показателям качества восстановленного видеопотока, о чем будет сказано ниже.

Важным элементом планирования эксперимента является обоснование количества опытов (N), позволяющих в полной мере учесть стохастичность воспроизводимого процесса функционирования БС. Для определения числа опытов могут быть использованы детерминированный и вероятностный методы.

Недостатком применения детерминированного метода является тот факт, что для определения числа опытов используются приблизительные методы расчета и оценить объективно точность результатов исследования не представляется возможным. В этой связи актуальным становится использование для определения числа опытов вероятностного метода, в общем случае представляющего собой сложный процесс, включающий ряд шагов.

Шаг 1. Оценка факторов, влияющих на объем выборки.

Шаг 2. Подбор метода расчета размера выборки.

Шаг 3. Расчет размера выборки.

Шаг 4. Оценка стандартного отклонения от среднего значения в выборочной совокупности.

Шаг 5. Расчет предельной ошибки выборки.

Шаг 6. Оценка среднего значения признака в генеральной совокупности.

Использование вероятностного метода позволит определить такое число опытов, которое в полной мере обеспечит репрезентативность выборки передаваемых видеопотоков, что, в свою очередь, повысит адекватность определения требуемых показателей их качества.

На третьем этапе определяется начальный уровень искажения видеопотока на стороне оператора, что позволит смоделировать возможное воздействие нежелательных факторов при прохождении видеопотока по КИО от БС к оператору. Для этого предлагается использовать специализированное программное обеспечение – эмулятор сети NetDisturb [11], которое позволяет вносить в поток пакетов (IPv4, IPv6) искажения, связанные с задержками, джиттером, дублированием и потерями пакетов, а также ограничением полосы пропускания канала связи. NetDisturb позволяет искажать потоки по IP-сетям, помогая изучать поведение приложений, устройств или служб в «нарушенной» сетевой среде. Данные искажения позволяют ухудшить качество итогового изображения, выводимого на средства отображения оператора, на уровне приложения FPV-управления в соответствии с логикой работы кодеков стандартов h264/h265.

Обоснованность использования NetDisturb в ходе моделирования воздействий нежелательных факторов обусловлена тем, что ГОНСС для управления БС включает в себя наземный сегмент сети связи, реализуемый на базе существующих сотовых сетей связи, а также космический сегмент, включающий в себя механизмы передачи трафика космических аппаратов, находящихся в том числе на различных орбитах, с целью глобального покрытия. В состав таких механизмов входит и маршрутизация трафика, что обеспечивается на сетевом уровне и представляет собой пакеты данных. Поэтому ГОНСС с точки зрения сети связи обеспечивает маршрутизацию пакетов данных и может рассматриваться как классическая сеть с коммутацией пакетов: моделирование задержек, джиттера и ограничения полосы пропускания с помощью пакета NetDisturb позволяет управлять потоком пакетов на уровне приложения FPV-управления БС.

Количество опытов N и начальный уровень искажения видеопотока на стороне оператора являются исходными данными для проведения натурной части эксперимента, являющегося *четвертым этапом* обоснования требований. Результатом проведения натурного эксперимента должны быть:

– фактическая вероятность недостижения цели функционирования БС, которая вычисляется следующим образом:

$$P = \frac{N^{\text{ну}}}{N},$$

где $N^{\text{ну}}$ – количество опытов, в которых цель функционирования БС не была достигнута;

– множество значений показателей качества всех кадров видеопотоков в каждом опыте:

$$\{\pi^l\}_L^n, \quad n = \overline{1, N},$$

где π^l – значение показателя качества (SSIM или PSNR) l -го кадра видеопотока; L – общее количество кадров видеопотока.

Для определения показателей качества необходимо использовать выражения, представленные в [7]; переход к дальнейшим этапам будет зависеть от ряда условий:

$$P = P_{\text{пр}}, \quad (1)$$

$$P < P_{\text{пр}}. \quad (2)$$

В случае, если условия (1) и (2) не выполняются, то реализуется *пятый этап*, заключающийся в увеличении степени искажения видеопотока на стороне оператора, после которого необходимо вновь перейти к четвертому этапу. Если же условие (2) выполняется, то это говорит о том, что неправильно определены характеристики эксперимента, в этом случае необходимо их уточнение на втором этапе.

В случае, если условие (1) выполняется, то осуществляется переход на *шестой этап* – определение множества видеопотоков V , передаваемых оператору, которые позволили достичь цели функционирования БС.

Обработка полученного в результате эксперимента множества $\{\pi^l\}_L^n$, $n = \overline{1, N}$ позволит на *седьмом этапе* получить математическое ожидание показателя качества π_v^{cp} для видеопотока в целом.

Это становится возможным за счет определения с помощью метода ядерной оценки [12, 13] плотностей вероятностей SSIM и PSNR $f_v(\pi_v)$ и вычисления на их основе математического ожидания этих показателей:

$$\pi_v^{\text{cp}} = \int_{-\infty}^{\infty} \pi_v f_v(\pi_v) d\pi_v. \quad (3)$$

Полученные математические ожидания (3) позволяют на *восьмом этапе* определить требуемые значение каждого из показателей качества восстановленного видеопотока:

$$\pi^{\text{тр}} = \frac{\sum_{v=1}^V \pi_v^{\text{cp}}}{V}.$$

Далее важным этапом обоснования требований к показателям качества восстановленного видеопотока при FPV-управлении является их проверка на избыточность (*девятый этап*). Сущность такой проверки заключается в нахождении зависимости вероятности достижения всех требований к показателям качества восстановленного видеопотока, полученных ранее, от вариантов, используемых НК (4), и оценивании количества таких вариантов, при которых значение вероятности достижения всех требований близко к единице. В качестве такой зависимости может быть использована модель, структурно-логическая схема которой представлена на рисунке 2.

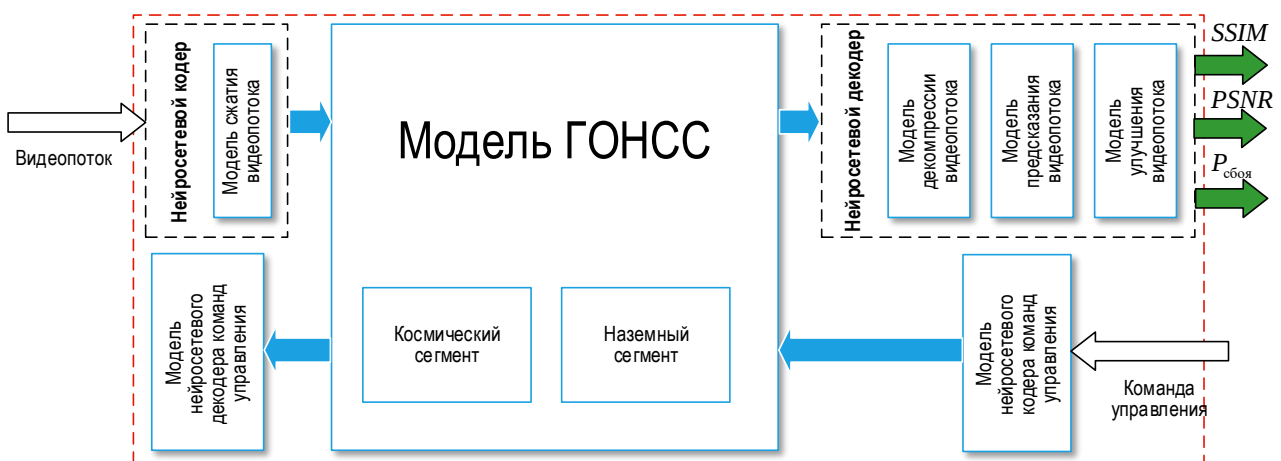


Рис. 2. Структурно-логическая схема модели информационного обмена между беспилотными средствами и оператором

Fig. 2. Structure and Logic Diagram of the Information Exchange Model between Unmanned Vehicles and Operator

В качестве моделей сжатия, декомпрессии, предсказания и улучшения видеопотока используются модели, представленные в [3]:

$$P[(SSIM(\mathbf{w}_\theta) \geq SSIM^{TP}), (PSNR(\mathbf{w}_\theta) \geq PSNR^{TP})] = f(\mathbf{w}_\theta), \quad (4)$$

где P – вероятность достижения всех требований к показателям качества восстановленного видеопотока; $\mathbf{w}_\theta \in W$ – θ -ый вариант НК; W – множество вариантов НК.

Зависимость (4) может быть получена путем проведения имитационного моделирования информационного обмена между БС и оператором при использовании различных вариантов НК ($\mathbf{w}$).

Целесообразность использования в качестве критерия проверки избыточности задания требований к показателям качества восстановленного видеопотока вероятности достижения всех требований одновременно определена в [14].

Для получения зависимости (4) должна быть использована имитационная модель, упомянутая при описании второго этапа. Если в результате анализа зависимости (4) окажется более одного решения, при котором вероятность достижения всех требований близка к единице, как, например, на рисунке 3 (решения 2, 3, 4, 7, 10, 12), то это позволяет сделать вывод о слишком широком диапазоне требуемых значений показателей качества восстановленного видеопотока.

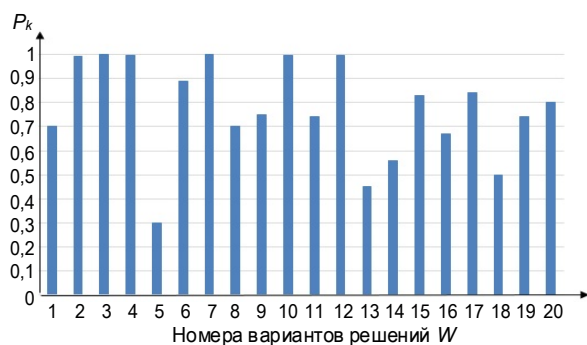


Рис. 3. Зависимость вероятности достижения всех требований к показателям качества видеопотока от вариантов решений по созданию КИО

Fig. 3. Dependence of Achieving All the Requirements Probability to Video Stream Quality Indicators on the Solutions Options for the Creation of IECs

Данное обстоятельство объясняется тем, что при полученных выше решениях весь диапазон значений плотности вероятности системы случайных величин SSIM и PSNR соответствуют таким их значениям, которые лежат внутри области заданных требований. Если в этом случае построить плотности вероятности SSIM и PSNR при параметрах КИО, соответствующих вариантам решений 2, 3, 4, 7, 10, 12, и рассматривать при этом SSIM и PSNR не как систему случайных величин, а как от-

дельные случайные величины, то можно увидеть следующую картину (рисунок 4).

При всех вариантах НК почти все возможные значения SSIM и PSNR принадлежат области заданных требований (см. рисунок 4). При этом в случае варианта 12 ($\mathbf{w}_{12}$) минимальные значения SSIM меньше, чем при других вариантах и меньше требуемого значения, равного 0,85. Таким образом видно, что если оставить требования к значениям показателей качества видеопотока в исходном виде, то можно получить множество решений, которые будут обладать высокой вероятностью достижения этих требований, но при этом характеризоваться меньшим значением показателя SSIM.

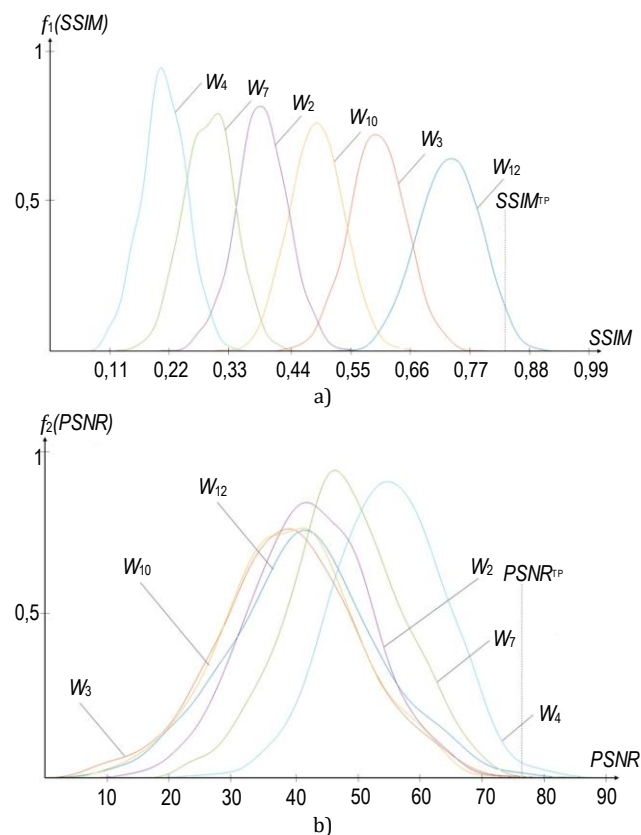


Рис. 4. Плотности вероятности SSIM (а) и PSNR (б) при различных вариантах решений по созданию КИО

Fig. 4. Probability Densities of SSIM (a) and PSNR (b) under Different IECs Solutions Options

Коррекция требований позволит получать только одно решение (рисунок 5), при котором вероятность достижения требуемых показателей качества видеопотока будет такой же высокой, как и при начально заданных требованиях, но такое решение позволит повысить либо показатель SSIM, либо PSNR, либо и то, и другое. Из рисунка видно, что при изменении требований удастся все равно получить оптимальное решение при $\mathbf{w}_7$, при котором $P_k(\mathbf{w}) \approx 1$.

Таким образом можно сделать вывод, что предложенный научно-методический аппарат позво-

ляет наряду с обоснованием требуемых значений показателей качества восстановленного видеопотока при FPV-управлении также корректировать эти значения в случае необходимости.

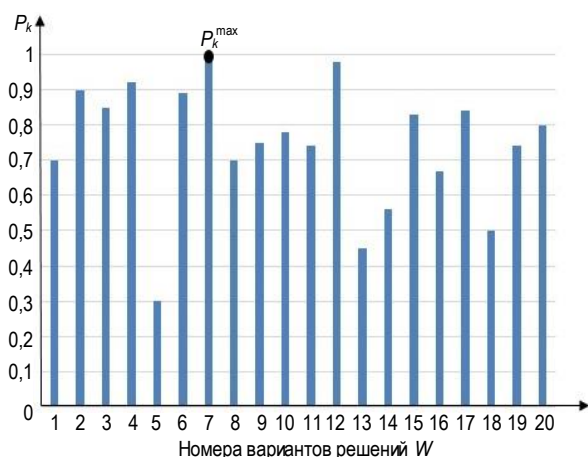


Рис. 5. Зависимость вероятности достижения всех требований к показателям качества от вариантов решений по созданию КИО при уменьшении требуемого значения показателя SSIM

Fig. 5. Dependence of Achieving All Requirements Probability to Video Stream Quality Indicators on the Solutions Variants for the Creation of KIO when the Required SSIM Indicator Value Decreases

Заключение

В результате проведенных исследований был разработан оригинальный подход к обоснованию требований к показателям качества восстановленного видеопотока при FPV-управлении, которые позволяют в процессе проектирования КИО осуществить выбор приемлемого варианта НК.

Научная новизна предложенного подхода заключается в том, что требуемые значения показателей, характеризующих качество восстановленного видеопотока, с помощью которого формируются воздействия по управлению БС, рассчитываются путем анализа различных видеопотоков, позволивших достичь целей его функционирования. Это позволяет избавиться от субъективизма, присущего используемым в настоящее время для решения этой задачи экспертным методам.

Кроме того, предложенный научно-методический аппарат позволяет не только обосновать требования к показателям качества восстановленного видеопотока при FPV-управлении с точки зрения достижения цели функционирования БС, но и корректировать эти требования в случае, если они заданы некорректно.

Список источников

1. Гриценко А.А. Гибридные радиосети, или Третья технологическая волна в развитии спутниковых систем // Connect. 2023. № 11-12. С. 48–53.
2. Тихвинский В., Стрелец М. Перспективы создания спутникового сегмента 5G // Первая миля. 2018. № 1(70). С. 16–25. DOI:10.22184/2070-8963.2018.70.1.16.25. EDN:YRTPCM
3. Березкин А.А., Вивчарь Р.М., Слепнев А.В., Киричек Р.В., Захаров А.А. Метод сжатия видеопотока при управлении беспилотными системами в гибридных орбитально-наземных сетях связи // Электросвязь. 2023. № 10. С. 48–56. DOI:10.34832/ELSV.2023.47.10.007. EDN:HRBGLL
4. Матарас А.А., Гуляев И.Ю. Анализ применения FPV дронов в ходе боевых действий 2014-2023 гг. // Актуальные вопросы повышения эффективной огневой подготовки в силовых структурах: теория и практика (III Макаровские чтения), Пермь, Россия, 23 мая 2023. Всероссийский сборник научно-практических материалов. Пермь: Пермский военный институт войск национальной гвардии Российской Федерации, 2023. Т. 3. С. 135–141. EDN:XCJAX
5. Kumar P., Parmar A. Versatile Approaches for Medical Image Compression: A Review // Procedia Computer Science. 2020. Vol. 167. PP. 1380–1389. DOI:10.1016/j.procs.2020.03.349
6. Старовойтов В.В. Уточнение индекса SSIM структурного сходства изображений // Информатика. 2018. Т. 15. № 3. С. 41–55. EDN:XZOOHR
7. Березкин А.А., Вивчарь Р.М., Киричек Р.В. Модель системы управления мобильными роботизированными комплексами различного назначения // Электросвязь. 2023. № 8. С. 12–18. DOI:10.34832/ELSV.2023.45.8.002. EDN:XXOJNM
8. Звягин В.И., Птушкин А.И., Трудов А.В. Риск как одно из свойств качества решений, принимаемых в условиях неопределенности // Надежность. 2018. Т. 18. № 4(67). С. 45–50. EDN:VNPHJG
9. ГОСТ Р ИСО 11231-2013 Менеджмент риска. Вероятностная оценка риска на примере космических систем. М.: Стандартинформ, 2014.
10. NASA/SP-2011-3421 (12/2011) Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners NASA.
11. NetDisturb // ZTI Communications. URL: <https://www.zti-communications.com/netdisturb> (дата обращения 08.07.2024)
12. Поршенев С.В., Копосов А.С. Использование аппроксимации Розенблатта-Парзена для восстановления функции распределения непрерывной случайной величины с ограниченным одномодальным законом распределения // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013. № 92. С. 1–27. EDN:RNEGGN
13. Маслаков М.Л., Терновская А.К. Построение плотности распределения вероятностей КАМ сигналов // Цифровая обработка сигналов. 2021. № 3. С. 36–40. EDN:FIACYQ
14. Березкин А.А., Вивчарь Р.М., Киричек Р.В. Многокритериальная оценка эффективности управления беспилотными системами в гибридных сетях связи // Труды учебных заведений связи. Т. 10. № 1. 2024. С. 18–25. DOI:10.31854/1813-324X-2024-10-1-18-25. EDN:VLZDQC

References

1. Gritsenko A.A. Hybrid radio networks, or the Third technological wave in the development of satellite systems. *Connect.* 2023;11-12:48–53. (in Russ.)
2. Tikhvinsky V., Strelets M. Prospects for creation of satellite segment for 5G. *Last Mile.* 2018;1(70):16–25. DOI:10.22184/2070-8963.2018.70.1.16.25. (in Russ.) EDN:YRTPCM
3. Berezkin A.A., Vivchar R.M., Slepnev A.V., Kirichek R.V., Zaharov A.A. Method of video stream compression when controlling unmanned systems in hybrid orbital-terrestrial communication networks. *Electrosvyaz.* 2023;10:48–56. (in Russ.) DOI:10.34832/ELSV.2023.47.10.007. EDN:HRBGLL
4. Mataras A.A., Gulyaev I.U. Evolution of FPV drone use during combat operations 2014–2023. *Proceedings of the All-Russian Scientific and Practical Conference Actual Issues of Increasing Effective Firearms Training in Power Structures: Theory and Practice (III Makarov Readings), 23 May 2023, Perm, Russia, vol.3.* Perm: Perm Military Institute of the National Guard Troops of the Russian Federation Publ.; 2023. p.135–141. (in Russ.) EDN:XCNJAX
5. Kumar P., Parmar A. Versatile Approaches for Medical Image Compression: A Review. *Procedia Computer Science.* 2020; 167:1380–1389. DOI:10.1016/j.procs.2020.03.349
6. Starovoitov V.V. Enhancement of the structural similarity index SSIM. *Informatics.* 2018;15(3):41–55. (in Russ.) EDN:XZOOHR
7. Berezkin A.A., Vivchar R.M., Kirichek R.V. Model of the mobile robotic complex management system. *Electrosvyaz.* 2023;8: 12–18. (in Russ.) DOI:10.34832/ELSV.2023.45.8.002. EDN:XXOJNM
8. Zviagin V.I., Ptushkin A.I., Trudov A.V. Risk as one of the properties of decisions taken under uncertainty. *Dependability.* 2018;18(4):45–50. (in Russ.) EDN:VNPHJG
9. ISO 11231:201 Space systems. Probabilistic risk assessment. Moscow: Standardinform Publ.; 2014. (in Russ.)
10. NASA/SP-2011-3421 Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners NASA. December 2011.
11. ZTI Communications. NetDisturb. URL: <https://www.zti-communications.com/netdisturb> [дата обращения 08.07.2024]
12. Porshnev S.V., Koposov A.S. Using Rozenblatt-Parzen approximation for recovering a cumulative distribution function of continuous random variable with a bounded single-mode distribution rule. *Scientific journal of KubSAU.* 2013;92:1–27. (in Russ.) EDN:RNEGGN
13. Maslakov M.L., Ternovaya A.K. Estimation of the probability density function of QAM signals. *Digital Signal Processing.* 2021;3:36–40. (in Russ.) EDN:FIACYQ
14. Berezkin A., Vivchar R., Kirichek R. Multi-Criteria Evaluation of UAV Control Efficiency in Hybrid Communication Networks. *Proceedings of Telecommunication Universities.* 2024;10(1):18–25. (in Russ.) DOI:10.31854/1813-324X-2024-10-1-18-25. EDN:VLZDQC

Статья поступила в редакцию 29.07.2024; одобрена после рецензирования 26.08.2024; принята к публикации 27.08.2024.

The article was submitted 29.07.2024; approved after reviewing 26.08.2024; accepted for publication 27.08.2024.

Информация об авторах:

БЕРЕЗКИН
Александр Александрович

кандидат технических наук, доцент кафедры программной инженерии и вычислительной техники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-1748-8642>

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 621.39, 530.182

<https://doi.org/10.31854/1813-324X-2024-10-4-16-26>

Анализ методов управления дисперсией для поддержания квазисолитонного режима распространения импульсов в высокоскоростной волоконно-оптической системе связи

Мария Сергеевна Былина¹, Bylina.Maria@sut.ru
Сергей Федорович Глаголев¹, Glagolev.Sergey@sut.ru
Сергей Эдуардович Доценко²✉, dotsenko.sergei@ya.ru

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

²ООО «Газпром трансгаз Санкт-Петербург», Санкт-Петербург, 196128, Российская Федерация

Аннотация

Актуальность. С каждым годом возрастает необходимость в увеличении пропускной способности и дальности передачи данных в волоконно-оптических системах связи. Исследование методов управления дисперсией является актуальным и способствует поддержанию квазисолитонного режима, что важно для обеспечения высокоскоростной и надежной передачи данных в современных телекоммуникационных сетях. Проведенное моделирование и расчеты параметров волоконно-оптической системы связи делает исследование практическим и прикладным, что позволяет инженерам и исследователям точно прогнозировать поведение системы и оптимизировать ее характеристики до внедрения в реальных сетях.

Постановка задачи: исследование процессов поддержания квазисолитонного режима в волокнах с уменьшающейся хроматической дисперсией и при чередовании волокон с разными знаками дисперсии, как с начальным чирпингом, так и без него.

Цель работы: разработка и анализ методов управления дисперсией для поддержания квазисолитонного режима распространения импульсов в одномодовом оптическом волокне.

Используемые методы: исследование процессов поддержания квазисолитонного режима проводилось с помощью методов математического и численного моделирования. Для апробации методов был проведен теоретический анализ и расчеты, а также разработаны схемы квазисолитонной волоконно-оптической системы связи и проведено имитационное моделирование. Анализ **результатов** показал эффективность предложенного метода, продемонстрировав работоспособность и устойчивость предложенных решений в условиях, близких к реальным.

Новизна: разработаны модели волоконно-оптических систем связи для поддержания квазисолитонного режима и методики их исследования, проведен анализ наиболее эффективных методов поддержания длительности и пиковой мощности квазисолитонных импульсов на больших расстояниях.

Практическая значимость: разработанные модели и методики их исследования могут быть использованы в учебном процессе факультета инфокоммуникационных сетей и систем СПбГУТ и при создании реальных волоконно-оптических систем связи.

Ключевые слова: одномодовое волокно с уменьшающейся хроматической дисперсией, дисперсия групповых скоростей, фундаментальный солитон, квазисолитонный режим, фазовая самомодуляция, чирпинг, дискретный оптический усилитель

Ссылка для цитирования: Былина М.С., Глаголев С.Ф., Доценко С.Э. Метод управления дисперсией для поддержания квазисолитонного режима распространения импульсов в высокоскоростной волоконно-оптической системе связи // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 16–26. DOI:10.31854/1813-324X-2024-10-4-16-26. EDN:KIGLSY

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-16-26>

Method for Controlling Dispersion in Order to Maintain a Quasi-Soliton Pulse Propagation Mode in High-Speed Fiber-Optic Communications System

 **Maria Sergeevna Bylina**¹, Bylina.Maria@sut.ru
 **Sergey Fedorovich Glagolev**¹, Glagolev.Sergey@sut.ru
 **Sergey Eduardovich Dotsenko**²✉, dotsenko.sergei@ya.ru

¹The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

²Gazprom transgaz Saint-Petersburg LLC,
St. Petersburg, 196128, Russian Federation

Annotation

Relevance: Every year, there is an increasing need to enhance the bandwidth and range of fiber-optic communication systems. The study of methods to control dispersion is relevant and helps maintain a quasi-soliton regime, which is essential for ensuring high-speed, reliable data transmission. Conducted modeling and calculations make this research practical and applicable. This will enable engineers to accurately predict the system's behavior and optimize its performance before implementing it in real-world networks.

Problem statement: Investigation of the processes of maintaining a quasi-soliton regime in fibers with decreasing chromatic dispersion and alternating fibers with different signs of chromatic dispersion, both with and without initial chirping.

Goal of the work: The development and analysis of techniques to control dispersion in order to maintain a quasi-soliton mode of light propagation in single-mode optical fibers.

Methods: The study of the processes of maintaining a quasi-soliton regime was carried out by mathematical and numerical modeling. To substantiate the methods, theoretical analysis and calculations were carried out and schemes for a quasi-soliton fiber-optic communication system were developed and modeling was carried out.

Result: The analysis of the results demonstrated the effectiveness of the proposed method and showed the efficiency and stability of the solutions in conditions that were close to real-world scenarios.

Novelty: Models for maintaining a quasi-soliton regime and methods for studying them have been developed. The most effective ways to preserve quasi-soliton pulses over long distances have been analyzed.

Practical significance: The developed models and research methods can be applied in the educational process and the development of real fiber-optic communication systems.

Keywords: single-mode fiber with decreasing chromatic dispersion, dispersion of group velocities, fundamental soliton, quasi-soliton mode, phase modulation, chirping, discrete optical amplifier

For citation: Bylina M.S., Glagolev S.F., Dotsenko S.E. Method for Controlling Dispersion in Order to Maintain a Quasi-Soliton Pulse Propagation Mode in High-Speed Fiber-Optic Communications System. *Proceedings of Telecommunication Universities*. 2024;10(4):16–26. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-16-26. EDN:KIGLSY

1. Поддержание квазисолитонного режима при использовании волокон с уменьшающейся хроматической дисперсией

Основной недостаток солитонных волоконно-оптических систем связи (ВОСС) с управлением затухания – ограничение длины усилительного участка (УУ) [1]. В 1987 г. была предложена схема

солитонной ВОСС, использующей дискретные оптические усилители (ОУ), которая позволяет увеличить расстояние L_A между усилителями. Эта схема предполагает использование специального одномодового оптического волокна (ОМ ОВ) – волокна с уменьшающейся дисперсией (ВУД), которое было разработано, изготовлено и исследовано

[2, 3]. В этих ОВ модуль отрицательной дисперсии групповых скоростей (ДГС) β_2 по утверждению авторов уменьшается с увеличением расстояния Z .

Рассмотрим процессы распространения по ВУД оптических импульсов, имеющих форму гиперболического секанса, которая характерна для фундаментальных солитонов.

Известно, что для поддержания формы фундаментального солитона в ОВ необходимо, чтобы его пиковая мощность P_m в каждом сечении ОМ ОВ удовлетворяла условию [2]:

$$P_m(Z) = \frac{|\beta_2|(Z)}{(\gamma(Z) \cdot T_0^2)} = \frac{A(Z)}{T_0^2} = P_m(0) \cdot \exp(-\alpha \cdot Z), \quad (1)$$

где γ – коэффициент нелинейности ОМ ОВ; T_0 – каноническая полуширина импульса, имеющего форму гиперболического секанса $\text{sech}(T/T_0)$ (далее секансного импульса); α – коэффициент затухания ВУД; T – внутриимпульсное время, отсчитываемое от вершины импульса при любом значении Z ; $A(Z) = |\beta_2|(Z)/\gamma(Z)$ – параметр ВУД, зависящий от Z .

Из (1) следует, что для сохранения солитонного режима в ВУД необходимо обеспечить выполнение следующих условий:

$$A(Z) = |\beta_2(Z)|/\gamma(Z) = [|\beta_2(0)|/\gamma(0)] \cdot \exp(-\alpha \cdot Z), \quad (2)$$

$$P_m(0) = \frac{|\beta_2|(0)}{(\gamma(0) \cdot T_0^2)} = \frac{A(0)}{T_0^2}. \quad (3)$$

Процесс изготовления ВУД включал уменьшение диаметра сердцевин ОВ в процессе его вытягивания [3]. Это изменение диаметра сердцевин влияет на волноводную составляющую ДГС и коэффициент нелинейности γ , что приводит к уменьшению значения $A(Z) = |\beta_2(Z)|/\gamma(Z)$. В ходе проведенного исследования $A(Z)$ в ВУД изменялась в 10 раз на расстоянии 40 км. Во столько же раз изменялось требуемое значение пиковой мощности P_m . Фундаментальные солитоны сохраняли свою длительность и форму, даже при наличии потерь, составляющих 10 дБ [3].

Следуя выражениям (2–3) и используя результаты [3], рассмотрим гипотетический пример реализации одного УУ солитонной ВОСС со скоростью $V = 10$ Гбит/с на основе ВУД длиной $L_A = 100$ км и коэффициентом затухания $\alpha = 0,2$ дБ/км. Для последующих расчетов полагаем, что длительность импульсов на уровне половины амплитуды $t_u = 20$ пс, а требуемая для формирования фундаментального солитона на входе в ВУД пиковая мощность $P_m(0) = 17$ мВт (12,3 дБм).

Каноническая полуширина секансного импульсов при этом будет равна [4, 5]:

$$T_0 = t_u/1,763 = 11,34 \text{ пс}. \quad (4)$$

Тогда в соответствии с (3) $A(0) = 2,19$ Вт·пс², $A(100) = 0,0219$ Вт·пс², $P_m(100) = 0,17$ мВт. На рисунке 1 показана зависимость $P_m(Z)$ на УУ с ВУД (синяя кривая). Этот же график одновременно описывает изменения реальной пиковой мощности и требуемой для сохранения фундаментального солитона в ВУД на УУ. Важно отметить, что солитонный режим в такой ВОСС с ВУД будет сохраняться на больших расстояниях при установке дискретных ОУ с коэффициентом усиления $\alpha_{LA} = 20$ дБ через каждые $L_A = 100$ км.

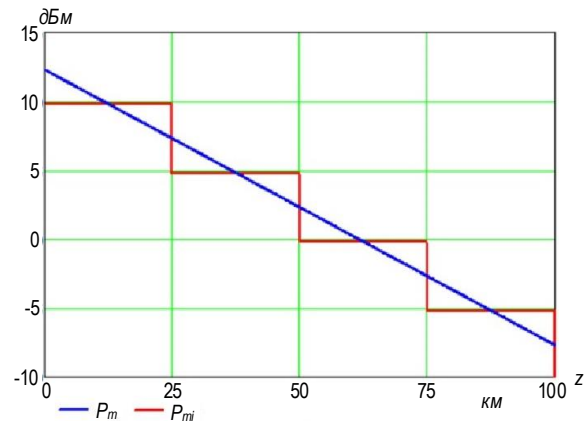


Рис. 1. Изменение пиковой мощности фундаментального солитона в ВУД (синяя линия) и средней пиковой мощности в четырех ОВ с дискретно уменьшающейся дисперсией (красная линия) в зависимости от расстояния на УУ

Fig. 1. The Change in the Peak Power of the Fundamental Soliton in a Fiber with Decreasing Dispersion (Blue Line) and the Average Peak Power in Four Optical Fibers with Discretely Decreasing Dispersion (Red Line) Depending on the Distance in the Amplification Section

В настоящее время ВУД серийно не производятся и ВОСС с их использованием не существует. Рассмотрим возможность создания аналога ВОСС с ВУД, используя УУ с n последовательно соединенными ОМ ОВ с номерами от $i = 0$ до $n-1$ дискретно уменьшающимся параметром A_i . Рассмотрим в качестве примера УУ протяженностью $L_A = 100$ км, разбитый на 4 одинаковых по протяженности ОВ длиной $l_{ob} = 25$ км с номерами от $i = 0$ до 3 с параметрами ОВ A_i , которые соответствуют средним значениям для уже рассмотренного УУ на основе ВУД на каждом 25-километровом участке. Ступенчатые изменения средних требуемых пиковых мощностей P_{mi} вдоль УУ представлены на рисунке 1 (красная кривая). Значение средней требуемой пиковой мощности будет уменьшаться на $\alpha l_{ob} = 5$ дБ при переходе к каждому следующему ОВ.

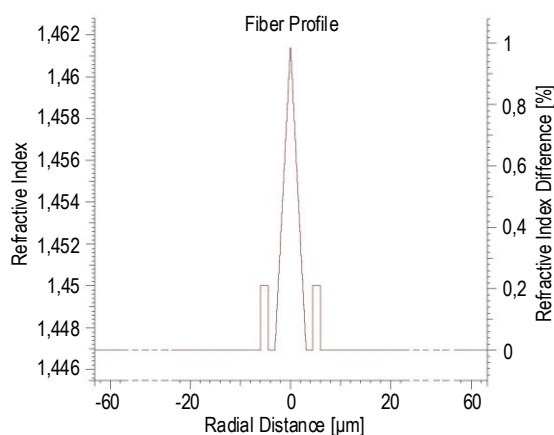
В таблице 1 приведены результаты расчетов требуемых значений средних для i -го ОВ пиковых мощностей P_{mi} и параметров ОВ A_i с учетом $A_0 = A(0) = 2,19$ Вт·пс² (3) по выражению:

$$P_{mi} = \frac{|\beta_{2i}|}{(\gamma_i \cdot T_0^2)} = \frac{A_i}{T_0^2}. \quad (5)$$

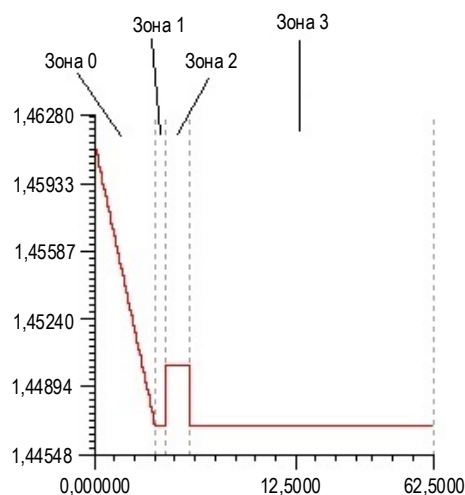
ТАБЛИЦА 1. Расчетные параметры УУ ВОСС с дискретно изменяющимися параметрами ОВ

TABLE 1. Calculated Parameters of the Amplifying Section of a Fiber-Optic Communication System with Discretely Varying Optical Fiber Parameters

Номер ОВ, i	0	1	2	3
Расстояние до начала i -го ОВ, км	0	25	50	75
Средн. пик. мощн. в i -ом ОВ P_{mi} , мВт	9,55	3,02	0,95	0,30
Уров. средн. пик. мощности p_{mi} , дБм	9,80	4,80	-0,20	-5,20
Параметр A_i , Вт·пс ²	1,23	0,39	0,12	0,04



а)



б)

Рис. 2. Профиль ППП для ОМ ОВ типа DSF (а) и изменение параметров ППП в OptiFiber (б)

Fig. 2. Refractive Index Profile for Single-Mode Optical Fiber Type DSF (a) and Changing the Parameters of the Refractive Index Profile in OptiFiber (b)

Программа OptiFiber использует метод численного решения и другие специализированные модели, которые позволяют оценивать широкий спектр параметров оптических волокон, включая хроматическую дисперсию (ХД) D_x , нелинейный показатель преломления n_2 и площадь модового поля A_{ef} . Параметры D_x , n_2 и A_{ef} , подобранные с помощью программы OptiFiber и обеспечивающие требуемые значения A , приведены в таблице 2. Величина ДГС β_{2i} и коэффициент нелинейности γ_i рассчитывались для синтезируемых ОВ по нижеприведенным выражениям [6] и помещались в таблицу 2:

$$\beta_{2i} = -D_{xi} \cdot \frac{\lambda_0^2}{2\pi \cdot c}, \quad (6)$$

$$\gamma_i = \frac{2\pi}{\lambda_0} \cdot \frac{n_2}{A_{ef}}, \quad (7)$$

где c – скорость света; λ_0 – длина волны излучения.

Для подтверждения возможности создания квазисолитонной ВОСС проведено моделирование процессов распространения квазисолитонных им-

Для подтверждения возможности создания ОМ ОВ с заданными в таблице 1 параметрами для реализации квазисолитонной ВОСС с УУ длиной 100 км использовалась программа OptiFiber. В качестве прототипа для синтезируемых ОВ в программе было выбрано ОВ со смещенной дисперсией (DSF, аббр. от англ. Dispersion-Shifted Single Mode Fiber), которое имеет профиль показателя преломления (ППП), показанный на рисунке 2а. Для синтеза ОВ с заданными свойствами изменялись геометрические параметры ППП, в частности радиусы зон 0–2 (рисунок 2б).

пульсов в многопролетной схеме ВОСС (рисунок 3) в программе OptiSystem.

Программное обеспечение OptiSystem предназначено для проектирования и оптимизации систем оптической связи. Оно состоит из более чем 600 различных компонентов, что позволяет детально планировать, тестировать и моделировать разнообразные типы оптических линий связи, включая сверхдальние ВОСС.

Каждый пролет волоконно-оптического линейного тракта (ВОЛТ) составлен из четырех ОМ ОВ и ОУ типа EDFA. В качестве источника сигнала использован генератор периодической битовой последовательности (0101101110) со скоростью передачи данных $B = 10$ Гбит/с, который модулирует оптический генератор секансных импульсов. Длительность такта $\Delta T = 100$ пс, длительность импульсов на уровне половины пиковой мощности $t_u = 20$ пс. Для изменения количества УУ N_u в OptiSystem используется петлевой элемент. Для минимизации шумов и улучшения качества передающего сигнала перед фотоприемным устройством установлен оптический полосовой фильтр.

ТАБЛИЦА 2. Результаты моделирования в программе OptiFiber и расчетов параметров синтезируемых ВО
TABLE 2. The Results of Modeling in the OptiFiber Program and Calculations of Parameters of Synthesized Optical Fibers

Тип ВО	Радиус зон ППП ОМ ВО, мкм				Результаты моделирования на длине волны 1550 нм			Расчетные значения		
	Зона 0	Зона 1	Зона 2	Зона 3	$D\chi$, пс/нм·км	A_{ef} , мкм ²	n_2 , м ² /ВТ	β_2 , пс ² /км	γ , (Вт·км) ⁻¹	A , Вт·пс ²
DSF	3,10	1,32	1,50	57,30	0,00	53,50	$18,35 \cdot 10^{-21}$	0,000	1,390	0,000
OB ₀	3,42	1,00	1,50	56,58	1,63	48,55	$19,49 \cdot 10^{-21}$	-2,071	1,627	1,273
OB ₁	3,21	1,21	1,50	56,58	0,43	51,75	$18,72 \cdot 10^{-21}$	-0,548	1,466	0,374
OB ₂	3,14	1,28	1,50	56,58	0,13	53,18	$18,43 \cdot 10^{-21}$	-0,171	1,405	0,122
OB ₃	3,12	1,31	1,50	56,58	0,04	53,76	$18,32 \cdot 10^{-21}$	-0,056	1,381	0,041

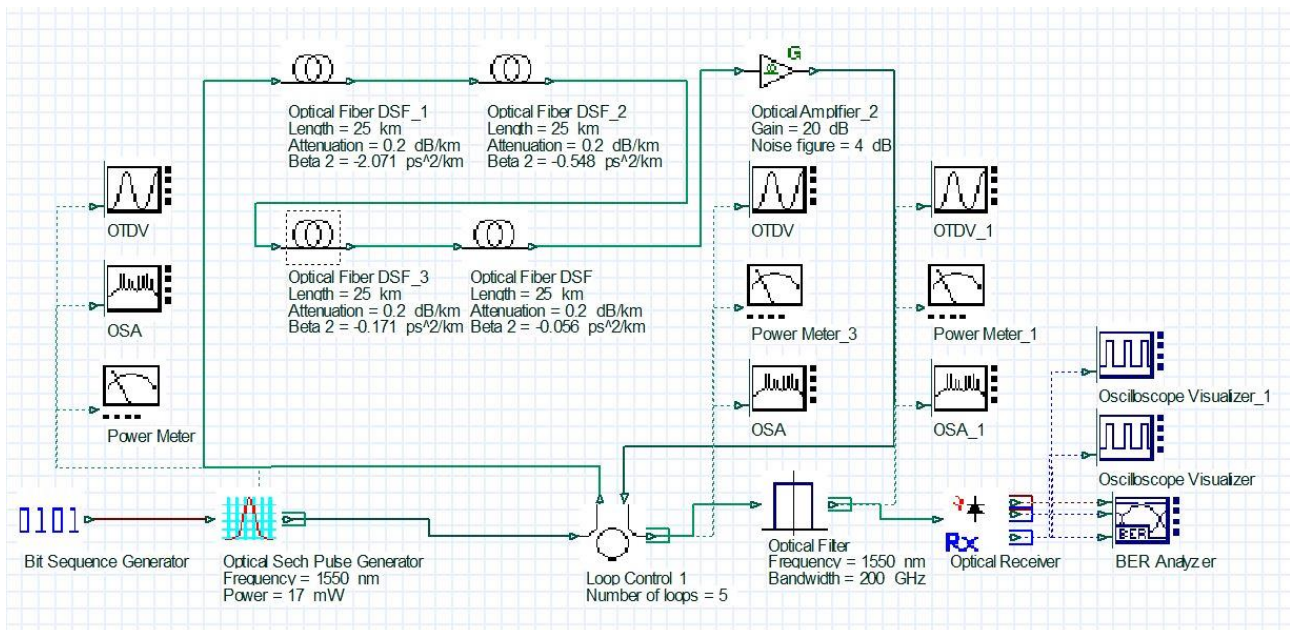


Рис. 3. Схема квазисолитонной многопролетной ВОСС с управлением дисперсией

Fig. 3. Schematic of Quasi-Soliton Multispan Fiber-Optic Communication System with Dispersion Control

Исходные данные и результаты исследований при длительности входного импульса $t_{ul} = 20$ пс и пиковой мощности $P_{m0} = 17$ мВт приведены в таблице 3 и на рисунке 4. Видно, что проведенное в программе OptiSystem моделирование многопролетной ВОСС с четырьмя ВО в пролете, подтверждает высокую эффективность такого способа поддержания квазисолитонного режима. Максимальное отклонение пиковой мощности P_{m1} на вы-

ходе квазисолитонной ВОСС длиной до 1000 км составляет не более 2,9 %, а отклонение длительности t_{ul} выходного импульса – не более 1,5 %.

Недостатком такого способа управления дисперсией является необходимость использования сравнительно большого количества специальных ВО и сравнительно большой пиковой мощности на входе в пролет.

ТАБЛИЦА 3. Результаты исследований квазисолитонной ВОСС
TABLE 3. Research Results of Quasi-Soliton Fiber Optic Communication System

№ п/п	Исходные данные		Измеренные величины		
	Кол-во пролетов, N_{up}	Общая длина, L , км	Вых. пик. мощн., P_{m1} , мВт	Длит. вых. импульса, t_{ul} , пс	Q
1	1	100	17,15	19,9	99
2	3	300	17	20,2	62
3	5	500	16,95	20,2	51
4	10	1000	16,60–16,80	20,3	30

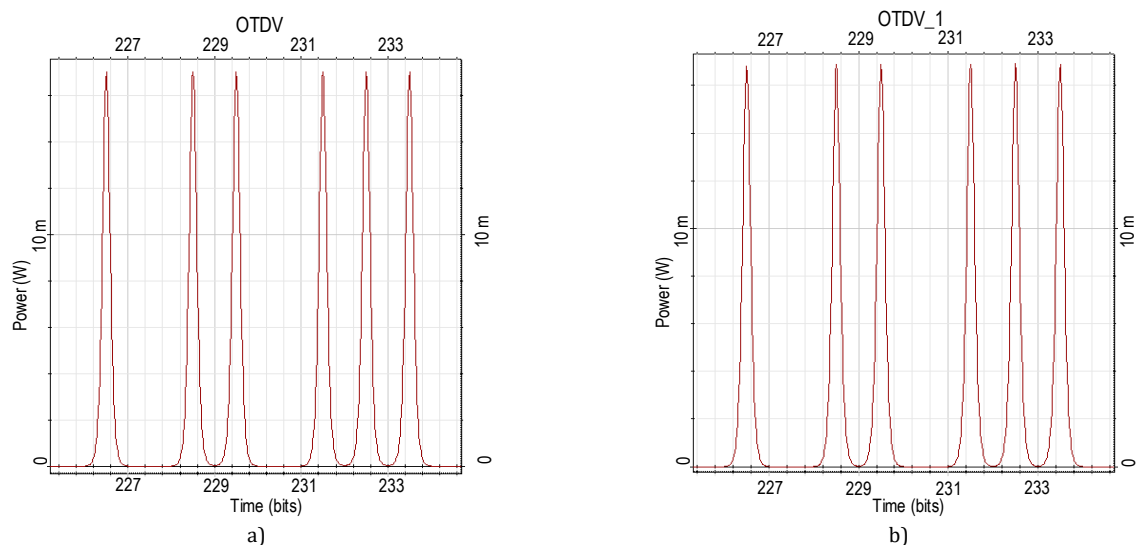


Рис. 4. Импульсы на входе (а) и выходе (б) моделируемой ВОСС длиной 1000 км

Fig. 4. Pulses at Input (a) and Output (b) of a Simulated 1000 km Fiber Optic Line

2. Поддержание квазисолитонного режима при использовании чередования волокон без потерь с разными знаками ХД

2.1. Использование источников излучения без начального chirping для ВОЛТ

В [7] было показано, что чередование двух ОМ ОВ с разными знаками ДГС позволяет реализовать квазисолитонный режим даже при использовании волокон со сравнительно большими ее отрицательными (нормальными) β_{2n} и положительными (аномальными) β_{2a} значениями, но при низкой величине средней ДГС:

$$\Delta\beta_2 = \frac{\beta_{2n} \cdot l_n + \beta_{2a} \cdot l_a}{l_n + l_a}, \quad (8)$$

где l_n и l_a – длины ОМ ОВ с нормальной и аномальной ДГС.

Из выражения (8) следует, что остаточная ДГС может иметь очень малое отрицательное значение

при достаточно больших ДГС каждого ОМ ОВ в отдельности. Это существенно снижает влияние эффектов четырехволнового смешения и ХД 3-го порядка в каждом отдельном ОМ ОВ. В результате в каждой секции из двух ОМ ОВ можно обеспечить небольшую отрицательную остаточную ДГС, которую затем компенсируют нелинейным эффектом фазовой самомодуляции. В пределах секции пиковая мощность, длительность, chirping и форма квазисолитонного импульса будут периодически осциллировать. Такие квазисолитоны получили название солитонов с управляемой дисперсией [7].

Оценим значение требуемой входной пиковой мощности секансных P_{0m} импульсов, необходимой для компенсации остаточной средней ДГС $\Delta\beta_2$ в секции, состоящей из двух ОМ ОВ без потерь [2]:

$$P_{0m} = \frac{|\Delta\beta_2|}{\gamma \cdot T_0^2}. \quad (9)$$

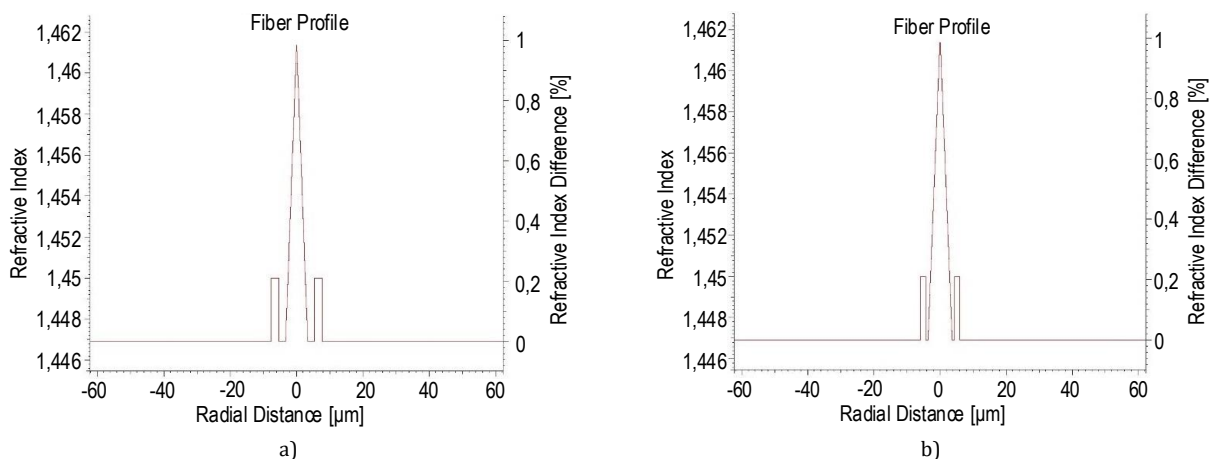


Рис. 5. ППП ОМ ОВ с ДГС 4,5 пс²/км (а) и -4,5 пс²/км (б)

Fig. 5. The Refractive Index Profile of an Optical Fiber with a Group Velocity Dispersion of 4,5 ps²/km (a) and -4,5 ps²/km (b)

Используя программу OptiFiber, проведем синтез двух ОМ ОВ, которые имеют ДГС разных знаков $\pm 4,5$ пс²/км. На рисунке 5 показаны ППП для таких ОМ ОВ. Параметры синтезированных ОМ ОВ приведены в таблице 4.

В программе OptiSystem проведено моделирование квазисолитонной ВОСС с чередующимися ОМ ОВ без потерь с параметрами, приведенными в таблице 4, с источником излучения без начального chirping (рисунок 6а).

Для реализации квазисолитонного режима в секции из двух ОМ ОВ с остаточной ДГС секции $\Delta\beta_2 = -0,4$ пс²/км выберем длину первого ОМ ОВ с нормальной ХД $l_n = 10$ км и рассчитаем по (8) длину второго ОМ ОВ l_a с аномальной дисперсией:

$$l_a = \frac{l_n \cdot (\beta_{2n} - \Delta\beta_2)}{\Delta\beta_2 - \beta_{2a}} = 11,84 \text{ км.}$$

ТАБЛИЦА 4. Результаты моделирования в программе OptiFiber и расчетов параметров, синтезируемых ОВ

TABLE 4. The Results of Modeling Using the OptiFiber Software and Calculations of the Parameters of Synthesized Optical Fibers

Тип ОВ	С норм. ХД	С аном. ХД
ДГС, β_2 , пс ² /км	4,512	-4,537
Коэффициент ХД, $D\chi$, пс/(нм км)	-3,54	3,56
Площадь модового поля, A_{eff} , мкм ²	52,50	46,08
Коэффициент затухания, α , дБ/км	0,2	0,2
Нелинейный показатель преломления, n_2 , м ² /Вт	18,54	20,3
Коэффициент нелинейности, γ , (Вт·км) ⁻¹	1,432	1,786
Параметр A , Вт·пс ²	3,152	2,541
Пиковая мощность P_{mi} , мВт	24	20

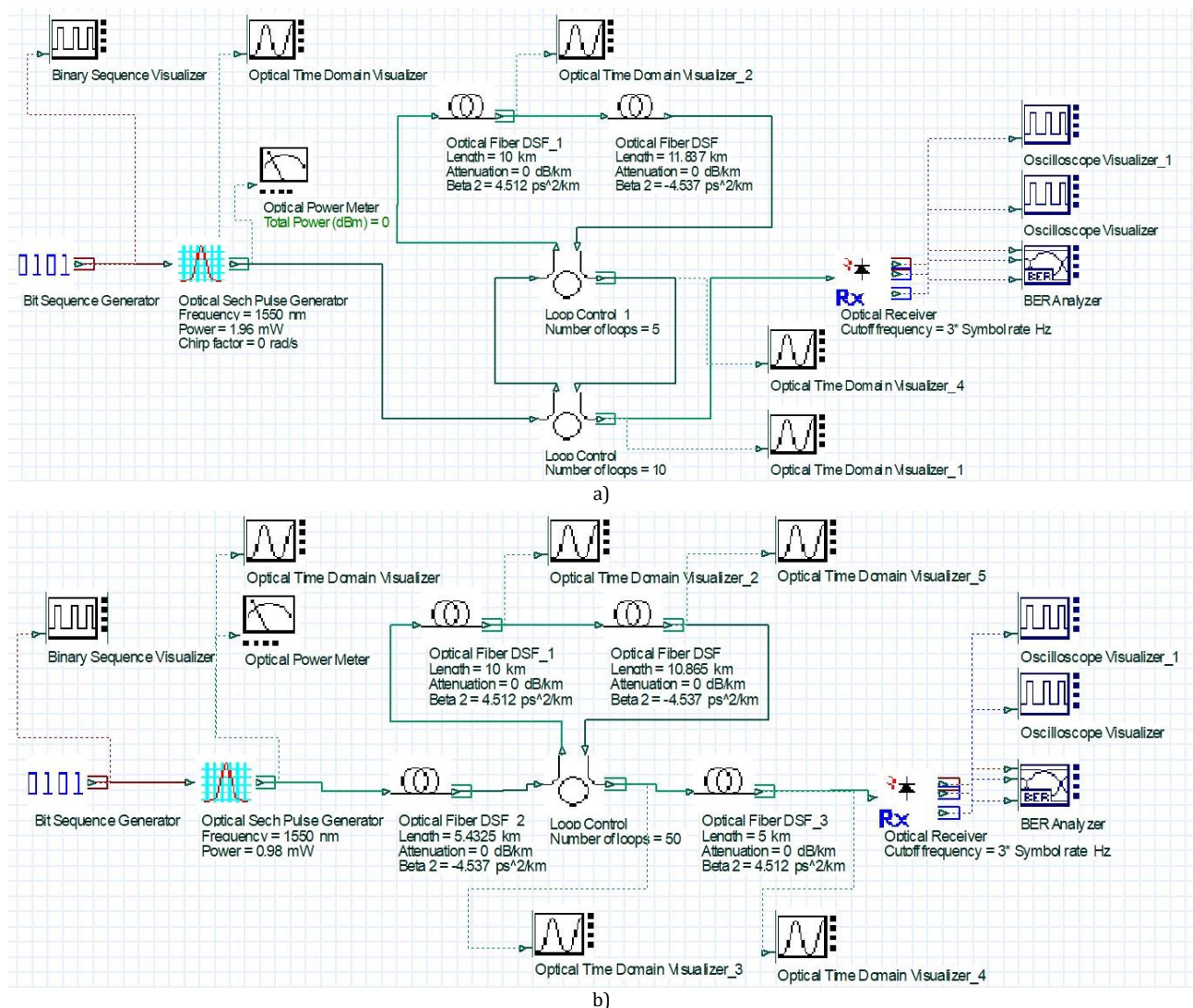


Рис. 6. Схема исследования квазисолитонной ВОСС с чередованием ОМ ОВ без потерь и без начального chirping (а) и с ОМ ОВ без потерь с начальным chirpingом, создаваемым волокном (б)

Fig. 6. The Study of a Quasi-Soliton Fiber-Optic Communication System Using Alternating Optical Fibers with No Loss and No Initial Chirping (a) and Using Lossless Optical Fiber, with Initial Chirping Caused by the Fiber (b)

Необходимые пиковые мощности для поддержания солитонного режима в синтезированных ОМ ОВ без потерь по уравнению (9) составили с нормальной ХД $P_{0m} = 2,17$ мВт, с аномальной ХД $P_{0m} = 1,74$ мВт. Для моделирования использовалась средняя мощность на секцию из двух ОМ ОВ $P_{0m} = 1,96$ мВт. Результаты исследований солитонного режима в ОМ ОВ без потерь и без начального chirpinga при разных значениях $\Delta\beta_2$ и l_n для длины линии порядка 1000 км приведены в таблице 5а (количестве пролетов $N = 10$).

Из таблицы 5а видно, что выходной импульс сохраняет свою форму при значениях остаточной дисперсии $\Delta\beta_2$ в пределах от $-0,2$ до $-0,4$ пс²/км и длины $l_n \leq 10$ км. При $l_n > 10$ км квазисолитонный режим поддерживается значительно хуже, учиты-

вая, что дисперсионная длина ОМ ОВ больше и равняется $L_D = 28$ км. С увеличением модуля $\Delta\beta_2 > -0,4$ пс²/км наблюдается джиттер, который приводит к колебаниям пиковой мощности и длительности импульсов от импульса к импульсу в пределах 5 %.

2.2. Использование на входе секции излучения с начальным chirpingом, создаваемым ОВ без потерь

Схема моделирования аналогична схеме на рисунке 6а, но с добавлением ОМ ОВ длиной $l_a/2$ для создания начального chirpinga в начало ВОЛТ и ОМ ОВ длиной $l_n/2$ для компенсации остаточного chirpinga в конце ВОЛТ (рисунок 6б) [8]. Результаты моделирования приведены в таблице 5б.

ТАБЛИЦА 5. Результаты исследований квазисолитонной ВОСС с чередованием ОВ без потерь и без начального chirpinga (а) и с использованием ОМ ОВ без потерь и с начальным chirpingом (б)

TABLE 5. The Research Results of a Quasi-Soliton Fiber-Optic Communication System Using Alternating Optical Fibers with No Loss (a) and Using Lossless Optical Fibers with Initial Chirping (b)

N , п/п	$\Delta\beta_2$	l_n , км	l_a , км	Кол-во секций на пролете	Длина пролета, L_N , км	Длина общая, L , км	Вх. пик. мощн., P_{0m} , мВт	Вых. пик. мощн., P_{lm} , мВт	Длит. вых. имп., t_{ul} , пс	Q
1	-0,1	10	10,39	5	101,9	1019	0,49	0,51	19,5	69
2	-0,2	5	5,43	10	104,3	1043	0,98	1,02	19,8	119
3	-0,2	10	10,87	5	104,3	1043	0,98	1,00	20,1	119
4	-0,2	15	16,30	4	125,2	1252	0,98	0,97	21,1	116
5	-0,3	5	5,68	10	106,8	1068	1,47	1,52	19,6	157
6	-0,3	10	11,36	5	106,8	1068	1,47	1,49	20,3	156
7	-0,3	15	17,04	4	128,1	1281	1,47	1,39	21,9	134
8	-0,4	5	5,94	10	109,4	1094	1,96	2,02	19,7	182
9	-0,4	10	11,84	5	109,2	1092	1,96	1,96	20,3	169
10	-0,4	15	17,81	4	131,2	1312	1,96	1,72–1,77	22,9–24	89
11	-0,6	10	12,99	5	114,9	1149	2,94	2,75–2,89	20,7–21,7	81
12	-0,8	10	14,22	5	121,1	1211	3,92	3,67–3,72	20,9–21,2	177

а)

1	-0,2	5	5,43	10	109,5	1095	0,98	0,98	20,0	122
2	-0,2	10	10,87	5	114,8	1148	0,98	0,97	20,4	121
3	-0,2	15	16,30	4	140,8	1408	0,98	0,93	21,2	120
4	-0,3	5	5,68	10	112,1	1121	1,47	1,49	19,9	164
5	-0,3	10	11,36	5	117,5	1175	1,47	1,45	20,3	159
6	-0,3	15	17,04	4	144,2	1442	1,47	1,40	21	162
7	-0,4	5	5,94	10	114,8	1148	1,96	2,00	19,8	201
8	-0,4	10	11,84	5	120,1	1201	1,96	2,03	19,5	207
9	-0,4	15	17,81	4	147,6	1476	1,96	1,89	21,1	168
10	-0,6	10	12,99	5	126,4	1264	2,94	4	14,8	309

б)

При моделировании квазисолитонного режима в ВОЛТ с начальным чирпингом выходной импульс сохраняет свою форму при значениях остаточной дисперсии $\Delta\beta_2$ в пределах от $-0,2$ до $-0,4$ пс²/км и длины $l_n \leq 10$ км. При $l_n > 10$ км квазисолитонный режим поддерживается хуже, но Q -фактор остается высоким.

3. Поддержание квазисолитонного режима в ВОСС с чередованием ОМ ОВ с потерями и разными знаками ХД

3.1. Использование источников излучения без начального чирпинга

Проведем исследование солитонного режима в ВОСС с чередованием ОМ ОВ с коэффициентом затухания $\alpha = 0,20$ дБ/км и без начального чирпинга. Схема моделирования и параметры ВОСС аналогичны схеме на рисунке 6а, но с включением в

каждый пролет ВОЛТ ОУ с коэффициентом усиления G , полностью компенсирующим затухание на ОУ, которое рассчитывается в зависимости от количества секций M из двух ОМ ОВ l_n и l_a :

$$G = \alpha \cdot \sum_{i=1}^M (\alpha \cdot l_n + \alpha \cdot l_a). \quad (10)$$

Для поддержания на ОУ средней пиковой мощности на его вход надо подавать существенно большую пиковую мощность. Ее можно определить по выражению [2]:

$$P_m = K \cdot P_{0m}, \quad (11)$$

где $K = g(G) \cdot \ln(g(G)) / (g(G) - 1)$ и $g(G) = 10^{\frac{G}{10}}$ раз.

Результаты исследования и расчетов по выражению (10 и 11) представлены в таблице 6, где количество ОУ – 10.

ТАБЛИЦА 6. Результаты исследований квазисолитонной ВОСС с использованием ОМ ОВ с потерями без начального чирпинга (а) и с использованием ОМ ОВ с потерями и начальным чирпингом (б)

TABLE 6. Research Results of a Quasi-Soliton Fiber Optic Communication System Using Optical Fibers with Attenuation without Initial Chirping (a) and Using Optical Fibers with Attenuation and Initial Chirping (b)

N, п/п	$\Delta\beta_2$	l_n , км	l_a , км	Кол-во секций на ОУ	Длина ОУ, L_n , км	Длина общая, L , км	Результаты расчета		Результаты моделирования		
							G , дБ	Вх. пик. мощн., P_{0m} , мВт	Вых. пик. мощн., P_{lm} , мВт	Длит. вых. имп., t_{ul} , пс	Q
1	-0,2	5	5,43	10	104,3	1043	20,86	4,75	4,85	19,9	13
2	-0,2	10	10,87	5	104,3	1043	20,87	4,75	4,9	20,3	13
3	-0,2	15	16,30	4	125,2	1252	25,04	5,67	5,48	21,9	8
4	-0,3	5	5,68	10	106,8	1068	21,36	7,28	7,45	20	14
5	-0,3	10	11,36	5	106,8	1068	21,36	7,28	7,15	21,1	14
6	-0,3	15	17,04	4	128,1	1281	25,63	8,70	7,8–8,1	21,6–23,2	8
7	-0,4	5	5,94	10	109,4	1094	21,87	9,94	9,9	20,1	14
8	-0,4	10	11,84	5	109,2	1140	21,84	9,92	9,6	21,0	14
9	-0,4	15	17,81	4	131,2	1312	26,25	11,87	9,3–9,8	25–27,5	8
10	-0,6	10	12,99	5	114,9	1149	22,99	15,64	12,8–14,2	22,2–24,9	15

а)

1	-0,2	5	5,43	10	109,5	1095	21,91	4,98	5,08	19,8	11
2	-0,2	10	10,87	5	114,8	1147	22,95	5,20	5,20	20,2	10
3	-0,2	15	16,30	4	140,8	1406	28,17	6,37	6,21	20,8	6
4	-0,3	5	5,68	10	112,1	1121	22,43	7,64	7,82	19,8	13
5	-0,3	10	11,36	5	117,5	1175	23,49	7,99	8,01	20,2	12
6	-0,3	15	17,04	4	144,2	1442	28,83	9,77	9,53	20,8	7
7	-0,4	5	5,94	10	114,8	1148	22,97	10,42	10,61	20	14
8	-0,4	10	11,84	5	120,1	1119	24,02	10,88	11,01	20	13
9	-0,4	15	17,81	4	147,6	1477	29,53	13,34	12,6	21,1	7
10	-0,6	10	12,99	5	126,4	1264	25,28	17,16	16,3–17,1	20–20,9	14

б)

При моделировании квазисолитонного режима в ВОЛТ с затуханием без начального чирпинга выходной импульс сохраняет свою форму при значениях остаточной дисперсии $\Delta\beta_2$ в пределах от $-0,2$ до $-0,4$ пс²/км и длине $l_n \leq 5$ км. При $l_n = 10$ км квазисолитонный режим поддерживается только при остаточной дисперсии $\Delta\beta_2 = -0,2$ пс²/км, в остальных случаях импульс значительно расширяется.

3.2. Использование на входе секции излучения с начальным чирпингом, создаваемым ОМ ОВ

Проведем исследование квазисолитонного режима в ОМ ОВ с потерями и начальным чирпингом, создаваемым ОМ ОВ. Схема моделирования и параметры ВОСС аналогичны схеме без затухания и начальным чирпингом (рисунок 6б), с установленным линейным ОУ.

Для поддержания на УУ средней пиковой мощности на вход линейного тракта надо подавать большую пиковую мощность с учетом затухания в ОМ ОВ для создания и компенсации чирпинга. Результаты расчетов коэффициента усиления по (10) и пиковой мощности по (11) представлены в таблице 6б.

Из таблицы 6б видно, что при моделировании квазисолитонного режима в ВОЛТ с затуханием и начальным чирпингом выходной импульс сохраняет свою форму при значениях остаточной дисперсии $\Delta\beta_2$ в пределах от $-0,2$ до $-0,4$ пс²/км и длине $l_n \leq 10$ км. При $l_n > 10$ км квазисолитонный режим поддерживается хуже.

Предварительное чирпирование не оказывает заметного влияния на качество поддержания квазисолитонного режима. Доказано, что возможно существенно снизить пиковую мощность на входе в пролеты по сравнению с применением ВУД и ОВ с дискретно уменьшающейся дисперсией на ВОЛТ.

Выводы

В статье исследованы методы управления дисперсией для поддержания квазисолитонного режима распространения импульсов в высокоскоростной ВОСС при использовании волокон с уменьшающейся ХД и при чередовании ее ОМ ОВ с разными знаками ХД. Все методы могут быть использованы в высокоскоростных ВОСС.

Наиболее эффективным методом поддержания длительности и пиковой мощности квазисолитонных импульсов на больших расстояниях является чередование ОМ ОВ с разными знаками ХД. Это решение позволяет обеспечить максимальную длину УУ до 120 км при минимальном значении входной пиковой мощности от 5 до 11 мВт.

Изменение длительности выходного импульса, при котором поддерживается квазисолитонный режим для импульса с начальным чирпингом, происходит в пределах от 19,8 до 20,2 пс при $\Delta\beta_2$ от $-0,2$ до $-0,4$ пс²/км и длине $l_n \leq 10$ км. Для импульса без начального чирпинга изменение длительности выходного импульса происходит в пределах от 19,9 до 20,3 пс, но в меньшем диапазоне $l_n \leq 5$ км при $\Delta\beta_2$ от $-0,2$ до $-0,4$ пс²/км и $l_n = 10$ км при $\Delta\beta_2 = -0,2$.

При поддержании квазисолитонного режима на ВОЛТ с длиной УУ 100 км и использовании волокон с уменьшающейся ХД изменение длительности выходного импульса происходит в пределах от 19,9 до 20,3 пс. Такое решение имеет более сложную схему, которая требует подбор ОМ ОВ с определенными параметрами.

Разработанные схемы квазисолитонных ВОСС и результаты их исследования дают представления о возможностях применения таких ВОСС в сетях связи, а также могут быть использованы в учебном процессе для подготовки специалистов по волоконно-оптической связи.

Список источников

1. Андреева Е.И., Былина М.С., Глаголев С.Ф., Чаймарданов П.А. Свойства временных оптических солитонов в оптических волокнах и возможность их использования в телекоммуникациях. Часть 2 // Труды учебных заведений связи. 2018. Т. 4. № 2. С. 26–35. DOI:10.31854/1813-324x-2018-2-26-35. EDN:XRVQCL
2. Андреева Е.И., Былина М.С., Глаголев С.Ф., Доценко С.Э., Чаймарданов П.А. Свойства временных оптических солитонов в оптических волокнах и возможность их использования в телекоммуникациях. Часть 3 // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 5–16. DOI:10.31854/1813-324x-2018-4-3-5-16. EDN:МАКQKL
3. Глаголев С.Ф., Лыков Г.Р. Исследование возможности практической реализации многопролетной квазисолитонной ВОСС с чередованием нескольких оптических волокон // Всероссийская научно-техническая и научно-методическая конференция магистрантов и их руководителей «Подготовка профессиональных кадров в магистратуре для цифровой экономики (ПКМ-2023)» (Санкт-Петербург, Российская Федерация, 05–07 декабря 2023). Сборник лучших докладов. СПб.: СПбГУТ, 2024. С. 283–287. EDN:TZOZEG
4. Агравал Г. Нелинейная волоконная оптика. М.: Мир, 1996. 323 с.
5. Agrawal G.P. Fiber-Optic Communication Systems. Wiley, 2010. 626 p.
6. Трещиков В.Н., Листвин В.Н. DWDM системы. М.: Техносфера, 2021. 420 с.
7. Кившарь Ю.С., Агравал Г.П. Оптические солитоны. От волоконных световодов до фотонных кристаллов. М.: ФИЗМАТЛИТ, 2005. 648 с.
8. Листвин А.В., Листвин В.Н., Швырков Д.В. Оптические волокна для линий связи. М.: ЛЕСАРпт, 2003. 288 с.

References

1. Andreeva E., Bylina M., Glagolev S., Chaimardanov P. Properties of Temporary Optical Solitons in Optical Fibers and the Possibility of Their Use in Telecommunications. Part 2. *Proceedings of Telecommunication Universities*. 2018;4(2):26–35. (in Russ.). DOI:10.31854/1813-324X-2018-4-3-5-16. EDN:XRVQCL
2. Andreeva E., Bylina M., Glagolev S., Dotsenko S., Chaimardanov P. Properties of Temporary Optical Solitons in Optical Fibers and the Possibility of Their Use in Telecommunications. Part 3. *Proceedings of Telecommunication Universities*. 2018;4(3):5–16. (in Russ.) DOI:10.31854/1813-324X-2018-4-3-5-16. EDN:MAKQKL
3. Glagolev S., Likov G. Investigation of the possibility of practical implementation of a multi-span quasi-soliton VOSS with alternating multiple optical fibers. *Proceedings of the All-Russian Scientific, Technical and Scientific-Methodical Conference of Undergraduates and Their Supervisors on Training of Professional Staff in the Master's Degree Program for the Digital Economy, 05-07 December 2023, St. Petersburg, Russian Federation*. Collection of the Best Reports. St. Petersburg: The Bonch-Bruevich Saint Petersburg State University of Telecommunications Publ.; 2024. p.283–287. (in Russ.) EDN:TZOZEG
4. Agrawal G. *Nonlinear fiber optics*. Moscow: Mir Publ.; 1996. 323 p. (in Russ.)
5. Agrawal G.P. *Fiber-Optic Communication Systems*. Wiley; 2010. 626 p.
6. Treschikov N., Listvin V.N. *DWDM Systems*. Moscow: Technosphaera Publ.; 2021. 420 p. (in Russ.)
7. Kishvar Yu.S., Agrawal G.P. *Optical Solitons. From Fiber Light Guides to Photonic Crystals*. Moscow: FIZMATLIT Publ.; 2005. 648 p (in Russ.)
8. Listvin A.V., Listvin V.N., Shvyrykov D.V. *Optical Fibers for Communication Lines*. Moscow: LESARart Publ.; 2003. 288 p. (in Russ.)

Статья поступила в редакцию 03.06.2024; одобрена после рецензирования 17.07.2024; принята к публикации 29.07.2024.

The article was submitted 03.06.2024; approved after reviewing 17.07.2024; accepted for publication 29.07.2024.

Информация об авторах:

БЫЛИНА
Мария Сергеевна

кандидат технических наук, заведующий кафедрой фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0001-9818-4060>

ГЛАГОЛЕВ
Сергей Федорович

кандидат технических наук, доцент кафедры фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-0664-9877>

ДОЦЕНКО
Сергей Эдуардович

ведущий инженер производственной лаборатории связи Службы связи ООО «Газпром трансгаз Санкт-Петербург»
 <https://orcid.org/0000-0003-0299-0469>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

Научная статья

УДК 621.391

<https://doi.org/10.31854/1813-324X-2024-10-4-27-37>



Задача маршрутизации в сети динамических туманных вычислений

Арте́м Николае́вич Волко́в, artem.nv@sut.ru

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

Аннотация

Актуальность. В условиях роста трафика, перехода к сетям IMT-2030 и услугам Телеприсутствия, особое место занимают задачи эффективного управления сетевыми и вычислительными ресурсами. Туманные вычисления, как следующая стадия декомпозиции архитектуры многоуровневых облачных вычислений, призваны в корне изменить модели и методы распределения вычислительных задач, повлияв, в том числе, на модели взаимодействия «пользователь-оператор». На данный момент существует целый пласт научных задач для раскрытия возможностей туманных вычислений. Их возможно разделить на ряд направлений, таких как: исследование моделей и методов реализации услуг сверхнадежных сетей связи с ультрамалыми задержками, определенные еще в сетях IMT-2020; исследование моделей и методов обеспечения качества обслуживания, включая качество восприятия; исследование методов живой миграции микросервисов, а также групп типовых микросервисов; исследование моделей и методов распределения ресурсов динамических туманных вычислений с обеспечением устойчивости форм туманных вычислений (кластера, туманности); одним из потенциально эффективных направлений является исследования в области объединения федеративного обучения (Federated Learning) с динамическими туманными вычислениями. В данной статье решается задача маршрутизации, которую можно отнести к направлению инфраструктурных исследований динамических туманных вычислений.

Постановка задачи: исследование эффективных методов определения маршрутов в сети динамических туманных вычислений, в том числе для задач миграции микросервисов услуг Телеприсутствия.

Цель работы: исследование и разработка эффективного метода поиска пути миграции микросервисов в сети туманных вычислений, который мог бы обеспечивать учет характеристик соединений (граней графа сети), вычислительные возможности, ограничения устройств туманных вычислений, и их особенности – динамичность устройств вычислений.

Используемые методы: для апробации предложенного метода была разработана модель в среде симулятора NS-3. Анализ полученных **результатов** показал эффективность предложенного метода в рамках поставленной задачи и различных сценариях применения.

Новизна: разработан метод миграции микросервисов в качестве нового протокола маршрутизации в среде динамических туманных вычислений, отличающийся от известных тем, что этот метод обеспечивает взаимодействие устройств туманных вычислений для миграции микросервисов, при этом позволяет достичь снижения потребляемой энергии устройствами туманных вычислений на 41 % и уменьшить долю потерянных пакетов в среднем до 34 %.

Практическая значимость: разработанный метод может быть использован при реализации туманных вычислений в условиях подвижности конечных устройств, в том числе с целью достижения требований перспективных услуг сетей IMT-2030.

Ключевые слова: IMT-2030, туманные вычисления, маршрутизация

Ссылка для цитирования: Волков А.Н. Задача маршрутизации в сети динамических туманных вычислений // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 27–37. DOI:10.31854/1813-324X-2024-10-4-27-37. EDN:QWBVQY

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-27-37>

Routing Task in Dynamic Fog Computing Network

 Artem N. Volkov, artem.nv@sut.ru

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

Annotation

Relevance. In the context of traffic growth, transition to IMT-2030 networks and Telepresence services, the tasks of efficient management of network and computing resources occupy a special place. Fog computing as the next stage of decomposition of the architecture of multi access edge cloud computing is designed to radically change the models and methods of distributing computing tasks, influencing, among other things, the user-operator interaction models. At the moment, there is a whole layer of scientific problems for revealing the possibilities of fog computing. They can be divided into a number of areas, such as: study of models and methods for implementing services of ultra-reliable and ultra-low latency communications, defined in IMT-2020 networks; study of models and methods for ensuring quality of service, including quality of experience; study of methods for live migration of microservices, as well as groups of typical microservices; study of models and methods for distributing resources of dynamic fog computing while ensuring the stability of fog computing forms (clusters, nebulae); one of the potentially effective areas is research in the field of combining federated learning with dynamic fog computing. This paper solves a routing problem that can be attributed to the direction of infrastructure research in dynamic fog computing.

Problem statement: research and develop the effective methods for routes determination in a dynamic fog computing network, including tasks of migrating microservices of telepresence services.

Goal of the work: research and development of an effective method for ways determination to migrate microservices in communication networks using fog computing technologies, which could take into account not only the characteristics of connections (edges of the network graph), but also the computing capabilities and limitations of fog computing devices, as well as their features - the dynamics of computing devices.

Methods: in order to test the proposed method, the program model was developed in the NS-3 modeling environment.

Result. Analysis of the results showed the effectiveness of the proposed method within the framework of the task and various application scenarios.

Novelty. A microservice migration method has been developed as a new routing protocol in a dynamic fog computing environment, which differs from the known ones in that this method ensures the interaction of fog computing devices for migrating microservices, while achieving a reduction in energy consumption by fog computing devices by 41% and reducing the share of lost packages on average up to 34%.

Practical significance: The developed method can be used to implement fog computing in conditions of mobility of end devices in order to achieve the requirements of promising services of IMT-2030 networks.

Keywords: IMT-2030, fog computing, routing

For citation: Volkov A.N. Routing Task in Dynamic Fog Computing Network. *Proceedings of Telecommunication Universities*. 2024;10(4):27–37. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-27-37. EDN:QWBVQY

Введение

В течение последних 10-ти лет активно развивалась архитектура многоуровневых облачных вычислений (МЕС, аббр. от англ. Multi Access Edge Computing), эффективность внедрения в сети которой была доказана как теоретически (во многих научных публикациях), так и на практике. Последующее масштабное внедрение многоуровневой облачной архитектуры осталось только вопросом времени,

где повышение эффективности распределения ресурсов может быть благодаря совершенствованию алгоритмов мониторинга и управления, а также модернизации самого программного обеспечения облачных платформ, обеспечивающих виртуализацию и контейнеризацию. В то же время многоуровневые облачные вычисления при всей доказанной эффективности являются статическим решением, заранее спроектированным и внедренным в сеть связи. Соответственно, каждый из уровней обладает жестко

закрепленным вычислительным и сетевым ресурсом тех серверов, которые были установлены оператором. Данная особенность имеет как положительные стороны (устойчивость, предсказуемость / детерминированность), так и недостатки (необходимость предварительного проектирования, внедрения и последующего обслуживания систем, ограниченность ресурсов и прочее) [1].

В то же время исследования в области интегрированных / гибридных сетей IMT-2030 и услуг телеприсутствия показали повышение потребности в вычислительных и сетевых ресурсах, поиске более оптимальных методов распределения задач, эффективных моделей интегрированных сетей, поиске решений в задачах управления ресурсами на основе инструментов искусственного интеллекта [2]. Например, в стратегии развития отрасли связи Российской Федерации (распоряжение Правительства РФ от 24 ноября 2023 года № 3339-р) приводится прогноз дальнейшего увеличения объема глобального Интернет трафика в среднем на 30 % в год, что может привести к росту в 30 раз в 2035 г., если сравнивать с 2022 г. Таким образом можно отметить возрастающую роль цифровых решений для обеспечения эффективности во всех отраслях экономики.

На данный момент, благодаря активному развитию микроэлектроники и повышению мощности вычислительных чипов, объему и скорости чтения-записи компонентов памяти (ПЗУ, ОЗУ), появлению мощных графических процессоров, в том числе с нейромодулями / нейропроцессорами, появилась возможность перехода к следующему уровню декомпозиции архитектуры облачных вычислений. Использование сетевых устройств последней мили, а также конечных пользовательских устройств в качестве вычислительных узлов сформировали новый вид распределенных облачных вычислений – туманные вычисления. Туманные вычисления на концептуальном уровне утверждены в консорциуме OpenFog в качестве спецификации в феврале 2017 г. Соглашение о создании консорциума было подписано в 2015 г. следующими участниками: ARM холдинг, Cisco, Dell, Intel, Microsoft и университет Принстона. В июле 2018 г. данный документ был принят IEEE в качестве стандарта OpenFog Reference Architecture. Стоит отметить, что данный документ заложил только технологический базис в области распределенных облачных вычислений, при этом на данный момент все еще идет стадия активных научных изысканий. В большинстве случаев в научных публикациях уделяется внимание исследованиям и разработкам в области динамических туманных вычислений, базирующихся на множестве пользовательских устройств и устройств Интернета вещей. В данном типе туманных вычис-

лений существует целый пласт научных задач: архитектурные, алгоритмические, интеграционные, протокольные и другие. В статье приводятся результаты исследований в области разработки эффективных методов маршрутизации в сети динамических туманных вычислений, в том числе для миграции микросервисов в условиях D2D (*аббр. от англ. Device-to-Device*) архитектуры сети.

Маршрутизация в сети туманных вычислений

В сетях связи известны различные методы маршрутизации, в частности динамической маршрутизации, которые основаны в первую очередь на данных о наличии маршрутов. При этом каждый маршрут имеет свою характеристику. Соответственно эти данные хранятся в маршрутизаторах и распределены в сети. Например, на практике в основе протоколов можно встретить такие алгоритмы, как:

- Distance Vector (DV), вектор расстояний, каждый маршрутизатор в сети хранит вектор расстояний до всех других сетей в сети (ярким примером является протокол RIP);
- Link-State (LS), состояние соединения, здесь каждый маршрутизатор собирает данные о состоянии всех соединений в сети и после вычисляет кратчайший путь до других сетей (примерами могут служить достаточно известные и часто применяемые протоколы IS-IS или OSPF);
- Path Vector (PV), вектор пути, данные алгоритмы подобны алгоритмам DV и при этом используют в дополнение информацию о маршрутах, а не только о существующих расстояниях (например, протокол маршрутизации BGP).

Выше в качестве примеров были приведены алгоритмы протоколов, относящиеся к опорным пакетным сетям связи. В то же время, с приходом концепции Интернета вещей [3] и предшествующей ей концепции M2M, были разработаны новые протоколы маршрутизации. Данные протоколы предназначены, в первую очередь, для беспроводных сетей связи, обладающих эффектом самоорганизации (предполагалась возможная динамика узлов сети, их гибридный режим работы с возможностью «сна» устройств и изменение архитектуры сети). Так, в качестве одного из критериев при разработке алгоритмов маршрутизации в сенсорных сетях является энергоэффективность. Вследствие данного требования и других особенностей сенсорных сетей протоколы маршрутизации стали иметь реактивный характер работы. То есть маршруты не хранятся в таблицах, а строятся заново при возникновении данной потребности. В качестве примера одного из таких алгоритмов можно привести алгоритм маршрутизации AODV (*аббр. от англ. Ad Hoc On-Demand Distance Vector* – протокол динамической маршрутизации для мобильных ad-hoc сетей),

где каждый узел сети самостоятельно строит маршрут в условиях архитектуры ячеистой сети. Для примера также стоит рассмотреть протокол RPL (*аббр. от англ. Routing Protocol for Low-Power and Lossy Networks* – протокол маршрутизации для сетей с низким энергопотреблением и потерями), стандартизированный в IETF как RFC 6550. Как в названии отмечено, данный протокол в первую очередь направлен на сети с низким энергопотреблением и потерями (LLN, *аббр. от англ. Low-Power and Lossy Networks*) – устройства могут иметь достаточно ограниченную мощность и склонны к потерям пакетов ввиду различных факторов. В основе протокола RPL лежат направленные ациклические графы DODAG (*аббр. от англ. Destination Oriented Directed Acyclic Graph*). Соответственно, данный протокол образует древовидную топологию, называемую DAG (*аббр. от англ. Directed Acyclic Graph*). При этом каждый узел в сети RPL имеет предпочтительного родителя, действующего как шлюз. Для оценки маршрутов протокол использует две целевые функции: на основе оценки потери пакетов и оценки количества переходов [4]. Отличительной возможностью RPL является Energy-Aware Routing (энергосберегающая маршрутизация), то есть маршруты в сенсорной сети выбираются с учетом энергопотребления узлов. Алгоритмы маршрутизации позволяют учитывать характеристику узлов сети. Например, такие параметры как: остаточный заряд батареи, уровень сигнала и расстояние до соседних узлов. Соответственно данный протокол – одно из тех гибридных решений в маршрутизации, которое позволяет учитывать не только «стоимость» маршрута / соединения между узлами, но и их возможность.

Представленный выше достаточно краткий обзор существующих подходов к маршрутизации позволяет оценить объемы проведенных исследований и разработок в области задачи определения оптимальных маршрутов для передачи трафика. В дополнение стоит отметить концепцию программно-конфигурируемых сетей (SDN, *аббр. от англ. Software-Defined Networks*), которая в вопросах маршрутизации имеет свои отличительные возможности, основанные на принципе централизации управления сегментом сети и наличия карты сети и ее ресурсов у единого элемента мониторинга и управления – контроллера SDN.

Стоит отметить, что вышеприведенные протоколы в первую очередь направлены на выстраивание маршрута для передачи данных при различных архитектурах и сценариях работы сети: древовидная, ячеистая. Туманные вычисления, в частности рассматриваемые динамические туманные вычисления, имеют достаточно противоречивую сущность. С одной стороны, это сеть из вычислительных устройств, представляющих собой вычислительный

кластер для развертывания программного обеспечения услуг, с другой стороны – это сеть динамически передвигающихся в пространстве устройств с ограниченными физическими ресурсами (мощность и энерговооруженность), а также ограниченными вычислительными и сетевыми ресурсами, что больше напоминает самоорганизующиеся сенсорные сети. В дополнение стоит отметить, что кроме задачи в области обеспечения связности сети устройств для передачи данных, особо остро стоит вопрос служебной маршрутизации в сети динамических туманных вычислений. Для обеспечения работы fog кластера, с учетом требований качества обслуживания QoS (*аббр. от англ. Quality of Service* – уровень сервиса или обслуживания), необходимо предусмотреть реализацию процессов живой миграции контейнеров микросервисов. Теме миграции микросервисов на данный момент посвящено немало научных работ, в основном, где ставится задача определения узла в вычислительном кластере для последующей миграции контейнера услуги. Существуют также патенты, например [5].

Вопросу маршрутизации в туманных вычислениях, в том числе, посвящены ряд работ, где в большинстве своем рассматриваются частные сценарии реализации Fog и / или интеграции с другими технологиями, такими как SDN. В данной статье поднимается вопрос обеспечения эффективной маршрутизации в динамических туманных вычислениях для решения задачи живой миграции микросервиса в сценарии D2D-взаимодействия узлов без предварительного копирования контейнера с централизованного хранилища. При этом предлагаемый алгоритм учитывает многопараметрическую оптимизацию целевой функции, где, кроме оценки состояния соединения, производится оценка возможности самих узлов для передачи контейнера по fog-сети. Ранее в работе [6] была поднята проблематика маршрутизации в сетях туманных вычислений, где рассматривается вопрос учитываемых параметров сети и узла, а также проводится сравнение с протоколом AODV. В текущей статье представлены результаты работы, где приводятся больше подробностей в части алгоритма маршрутизации, а также результаты более глубокого моделирования протокола, сравнение с аналогами и оценкой его эффективности.

Исследуемая модель и метод

На рисунке 1 модели сети туманных вычислений представлены следующие элементы:

- fog-устройства (FD); представлены следующим набором FD-1, FD-2, FD-3, FD-4, FD- n , FD- $(n+1)$;
- пользователи UD; представлены набором UD-1, UD-2, UD- m ;

- микросервис mS, располагающийся на устройствах FD-(n+1), FD-2 и мигрирующий между устройствами FD-n и FD-3;
- fog оркестратор, представлен набором серверов граничных вычислений Edge;
- NB-API (Northbound Application Programming Interface) – северный программный интерфейс, обеспечивающий связь между с и соответствующим fog оркестратором.

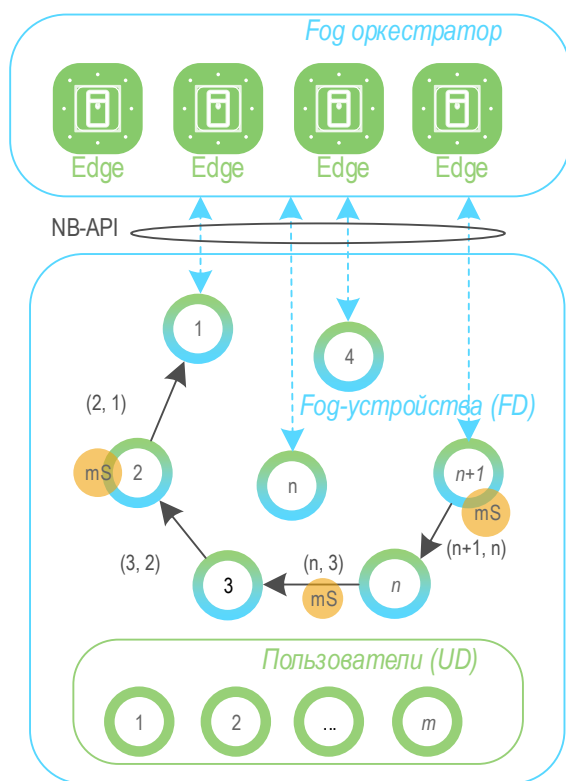


Рис. 1. Исследуемая модель

Fig. 1. The Model Under Research

На рисунке 1 приведены примеры связи между оркестратором и FD. Для полноценной работы сегмента туманных вычислений каждое устройство FD имеет соответствующую связь с оркестратором через NB-API для обмена в первую очередь служебной информацией.

В условиях динамических туманных вычислений, в первую очередь, рассматривается архитектура сети, где туманные устройства коммуницируют согласно принципу D2D, соответственно, важнейшей задачей здесь является выбор транзитных узлов для пересылки пакетов от отправителя к получателю.

Таким образом, главной характеристикой маршрута принимается стоимость миграции через него, и стоимость маршрута рассчитывается на основании группы важных параметров состояния сети. Ряд параметров относятся к определяющим QoS, на данный момент обозначенных как: использование ресурсов процессора узла, использование ресурсов

оперативной памяти узла, доля потерянных пакетов при передаче данных через канал связи, задержка при передаче данных через канал связи. Набор данных параметров отслеживается в реальном времени с помощью служебного северного API и пересылаются по сети по служебным каналам связи к оркестратору туманных вычислений. Таким образом, имея полную картину по нагрузке сети, а также анализ изменения спроса на тот или иной узел и / или микросервис, fog оркестратор контролирует процесс миграции. Вся получаемая статистика с устройства FD записывается в соответствующее хранилище данных и анализируется, и при возникновении необходимости в миграции микросервиса запускается алгоритм определения наиболее эффективного маршрута, с учетом как самих возможных путей, так и возможности самих транзитных узлов.

В работе [6] был определен алгоритм миграции микросервиса на базе анализа группы параметров. В частности, были исследованы вышеприведенные параметры и составлена фитнес-функция для расчета оценки пути. Отслеживаемые в данном случае характеристики состояния назовем метриками, учитываемыми предлагаемым протоколом. Некоторые из них также учитываются при расчете QoS:

- задержка на канале (D_{FD-n}) связи между узлами FD(n+1) и FD-n; измеряется с помощью специальных служб операционной системы и служебного ПО ведомой ноды оркестратора, запущенного на каждом узле, в миллисекундах (данные о задержке, при передаче данных через канал связи, являются достаточно важным параметром, так как миграция должна происходить достаточно быстро, чтобы остановка сервиса, перенос страниц его памяти и состояния, а также возобновление работы в новой среде, были незаметны для пользователей);

- доля потерянных пакетов (PL_{FD-n}) на канале связи между узлами FD(n+1) и FD-n; измеряется (в процентах) с помощью специальных служб, запущенных на каждом узле (важно знать долю потерянных пакетов при передаче данных через канал связи, так как доставка микросервиса на устройство назначения должна быть гарантированной, соответственно, недопустимо долгое ожидание завершения миграции в том числе потому, что сервер, на который происходит миграция сервиса, резервирует для него необходимое количество ресурсов, невозможность использования которых на протяжении продолжительного времени, при не останавливаемом процессе обработки данных может вызвать увеличение времени отклика на запросы и сбои в работе сети);

- нагрузка процессоров (CPU_{FD-n}) и оперативной памяти (RAM_{FD-n}) узла FD-n; измеряются (во флотах и килобайтах соответственно) посредством опроса каждого узла через оркестратор (важно

знать нагрузки процессора и оперативной памяти узла в активном состоянии, так как при предельном количестве используемых ресурсов у устройства может не хватить оставшихся для приема и пересылки пакетов при миграции микросервиса).

Выбор маршрута включает соответствующий процесс взаимодействия оркестратора сети туманных вычислений и n -го устройства FD, сам Алгоритм выбора оптимального маршрута миграции микросервиса на базе расчета фитнес-функции. Алгоритм осуществляет сбор статистики по использованию сетевых и вычислительных ресурсов, хранение данной статистики и последующее вычисление оптимального маршрута миграции микросервиса среди всех доступных маршрутов от узла $FD(n+1)$ до $FD-n$. Данный процесс отражен на рисунке 2.

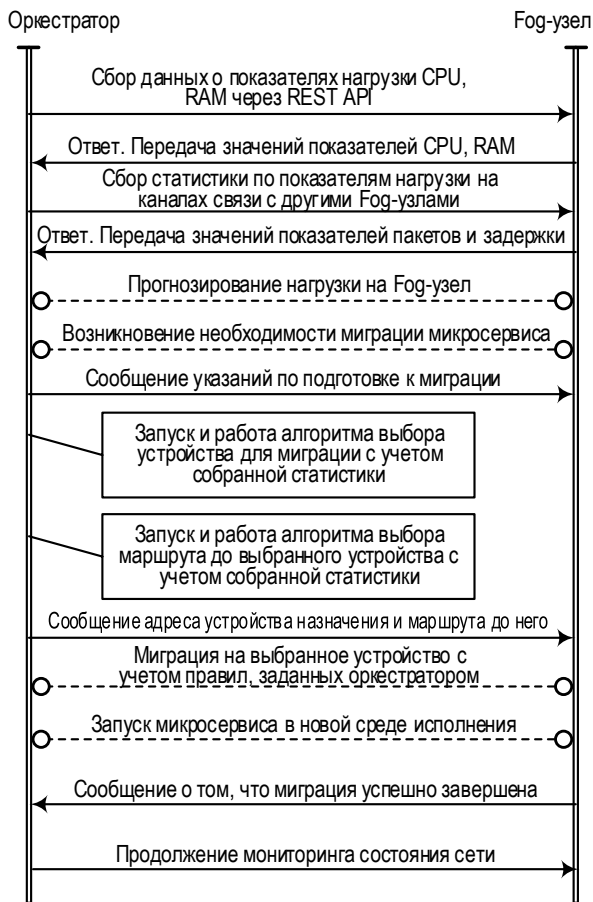


Рис. 2. Диаграмма взаимодействия оркестратора и FD- n

Fig. 2. Diagram of Interaction Between the Orchestrator and FD- n

Используется следующая фитнес-функция для нахождения стоимости одного ребра:

$$f_{FD(n+1),FD-n} = CPU_{FD-n}w_{CPU} + RAM_{FD-n}w_{RAM} + D_{FD-n}w_D + PL_{FD-n}w_{PL}, \quad (1)$$

где определены следующие веса:

$$w_{CPU} = w_D = w_{PL} = 0,275, w_{RAM} = 0,175. \quad (2)$$

Исследование влияния данных весов на конечный результат, а также обоснование их значений представлено в работе [3]. Веса должны определяться с помощью метода сравнения совокупности объектов друг с другом или метода опроса экспертов, исходя из условий поставленной задачи. Существующие веса были рассчитаны на основе метода парных сравнений. Сам метод парных сравнений предполагает сравнение всех исследуемых объектов во всех возможных парах среди этих объектов по заданному соответствующему критерию. В контексте рассматриваемой задачи были выдвинуты следующие условия:

1) нагрузка процессора узла имеет больший приоритет, чем нагрузка оперативной памяти узла (в текущей задаче мощность процессора узла туманных вычислений необходима для обработки процесса получения микросервиса и его пересылки к следующему участнику маршрута);

2) нагрузка процессора узла приоритетом эквивалентна задержке канала связи и оценке потерянных пакетов на этом канале связи (данные параметры в задаче выбора маршрута имеют одинаковую значимость и соответственно влияние на конечный результат);

3) оперативная память узла туманных вычислений имеет чуть меньший приоритет по отношению к задержке на канале связи и оценке потерянных пакетов на этом же канале связи;

4) задержка канала связи и доля потерянных пакетов на этом канале связи имеют идентичную значимость.

Функция расчета стоимости маршрута в общем виде [6] представлена следующей формулой:

$$F_{route}(1, n+1) = \sum f_{FD(n+1),FD-n}, \quad (3)$$

а искомый путь миграции микросервиса [6] в сети динамических туманных вычислений определяется как:

$$route_{migration} = \min_{CPU, RAM, D, PL} \{F_{route}(1, n+1)\}. \quad (4)$$

Полученный маршрут (4) принимается за оптимальный и сообщается вместе с остальной служебной информацией (способ миграции, узел-отправитель, узел-получатель и т. д.) соответствующим узлам туманных вычислений в сети по служебным каналам связи. Узлы принимают от оркестратора указания, после чего происходит миграция микросервиса с узла $(n+1)$ на узел 1 (см. рисунок 1). Таким образом, можно утверждать, что миграция микросервиса будет происходить по наиболее эффективному маршруту среди всех доступных в сети между узлами $FD(n+1)$ до $FD-n$.

Диаграмма взаимодействия между узлом туманных вычислений и соответствующим управляющим оркестратором отражена на рисунке 2.

Предлагаемый протокол маршрутизации микросервисов сети устройств туманных вычислений имеет также функционал обнаружения новых устройств для их последующего подключения к вычислительному кластеру. Процесс обнаружения новых устройств в сети является важной частью стабильного функционирования динамической сети туманных вычислений. Предполагается, что подключение нового устройства в сеть будет инициироваться самим новым устройством путем регистрации на специальном сервере своих учетных данных для дальнейшего добровольного предоставления свободных ресурсов устройства под задачи обработки данных. Отдельное внимание стоит уделить вопросу безопасности: необходимо обеспечить защиту от несанкционированного доступа в сеть туманных вычислений или подключение с целью нанесения ущерба, или хищения данных. Эту проблему должна решить технология цифровых сертификатов, позволяющая однозначно верифицировать и идентифицировать конкретное подключаемое устройство. Данный вопрос активно обсуждается в работе [11] Исследовательской комиссии МСЭ-Т. Однако предположим, что вопрос безопасности решен и для подключения требуется пройти аутентификацию, не допускающую злонамеренных вторжений в сеть, к примеру на основе технологии токенов с ограниченным временем жизни.

Рассмотрим возможные состояния устройства на протяжении жизненного цикла этого устройства в качестве fog-узла. На рисунке 3 представлена диаграмма состояний устройства.

На рисунке 3 состояние 0 соответствует начальному состоянию узла, именно с него начинается жизненный цикл устройства в роли fog-узла. Далее единственным возможным действием, выполняемым для перехода в состояние 1, является включение режима «поделиться ресурсами». После аутентификации устройства и его подключения к сети, как это было уточнено выше, устройство начинает считаться fog-узлом и может выполнять пересылку пакетов, то есть узел переходит в состояние 1. Перед миграцией микросервиса на устройство анализируется количество предоставленных им ресурсов, и если их достаточно, то решение о миграции и последующем выполнении задач будет принято. Если же ресурсов не хватает для решения задач обработки данных, узел будет оставаться в режиме пересылки сообщений, либо может принять решение о выходе из сети. Предположим, что ресурсов достаточно и микросервис был перенесен на новый fog-узел, после чего введен в активный режим, тогда узел переходит в состояние 2. Стоит отметить, что хоть и принцип распределения ресурсов не должен допускать перегрузки узлов, все же такие ситуации исключать нельзя. Если узел, выполняющий

задачу, теряет свою способность далее обеспечивать работу микросервиса, то может быть принято решение о выходе узла из сети, либо о переходе устройства в состояние 1, то есть в режим пересылки. Тогда устройство выполняет передачу задач иным узлам с помощью миграции микросервисов. Если же было принято решение о выходе из сети, тогда устройство также передает задачи иным узлам, имеющим для этого достаточное количество ресурсов, все микросервисы с устройства перемещаются на другие fog-узлы, после чего исходный узел покидает сеть и возвращается в состояние 0. Состояние 3 соответствует завершению жизненного цикла устройства в роли fog-узла.

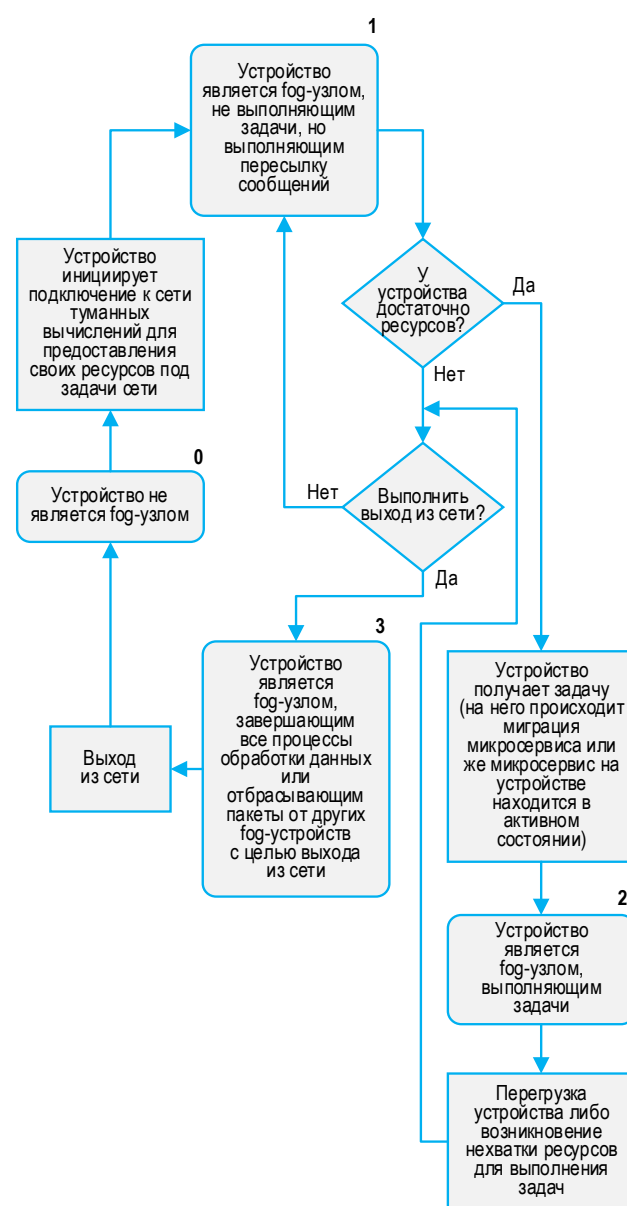


Рис. 3. Диаграмма состояния FD-n

Fig. 3. FD-n States Diagram

Рассмотрим подробнее переход узла из состояния 0 в состояние 1. При подтверждении статуса «сдачи в аренду» своих вычислительных ресурсов устройство проходит авторизацию для подключения к сети. Предлагается предоставлять владельцу устройства выбор из двух вариантов подключения: прямое подключение на базовую станцию (далее – БС) и прямое подключение на БС с поддержкой D2D-коммуникаций по отношению к узлам-соседям. В соответствии с выбранным вариантом подключения осуществляется тот или иной сценарий поведения исходного устройства как нового узла в сети туманных вычислений. Рассмотрим упомянутые сценарии. Стоит также уточнить, что согласно модели, представленной на рисунке 1, оркестратор узла туманных вычислений соответствующего сегмента находится на граничном облаке Edge, которое располагается на самой БС или рядом с ней.

1. Прямое подключение на БС

При прямом подключении на БС, сразу после верификации устройства и одобрения заявки на подключение, исходным узлом формируется сообщение Hello-E, направляемое на БС. Адрес БС сообщается узлу до отправки сообщения после верификации. Так как пользователь сам инициировал подключение к сети, сервисы, осуществляющие это подключение, в том числе отвечающие за первичную установку связи между БС и новым узлом сети, они и формируют сообщение. Сообщение Hello-E необходимо для передачи своего сетевого адреса только что подключенного к сети устройства. БС, при получении от устройства этого сообщения, формирует ответ, подтверждающий получение адреса узла. После получения ответа от БС, устройство воспринимает последнюю как шлюз для обмена данными с другими узлами Fog-сети.

2. Прямое подключение на БС с поддержкой D2D

Этот процесс несколько отличается от первого описанного варианта. Помимо установки связи с БС, устройство отправляет широковещательное сообщение Hello-N, направленное на передачу своего адреса ближайшим узлам сети туманных вычислений – возможным соседям. При этом на полученное сообщение будут реагировать только те из узлов, которые также поддерживают D2D и находятся в ближайшей зоне взаимодействия с исходным узлом – в пределах одного хопа (перехода). Это необходимо не только для того, чтобы поддерживать актуальную топологию сети при подключении к ней новых устройств, но и для возможности быстро устанавливать связь с соседними узлами, если они также поддерживают D2D. Fog-узлы в данном случае будут «общаться» напрямую, поэтому пакеты, отправляемые с устройства, не будут проходить через БС. Предполагается, чтобы через БС осуществлялась связь fog-узлов с другими узлами сети, причем такими, которые находятся на расстоянии

большем, чем 1 прыжок (переход). Устройства, получившие пакет Hello, не пересылают его далее, а записывают адрес нового узла в свои таблицы маршрутизации, после чего каждый из них, в случае, если также поддерживает связь D2D, формирует ответ на сообщение Hello-N с информацией о своем адресе. Новое же устройство, получив ответ, записывает адреса из пришедших пакетов в свои таблицы маршрутизации, устанавливается дуплексная связь между узлами-соседями.

Если ни один узел не ответил на отправленное сообщение Hello-N, устройство в любой ситуации считает шлюзом для обмена данными с остальными узлами сети БС.

После подключения к сети через один из сценариев устройство полноценно становится частью сети туманных вычислений, то есть переходит в состояние 1. Также другие узлы сети получают возможность производить миграцию различных микросервисов на новый узел, в зависимости от его доступных вычислительных ресурсов, а также использовать его как транзитный при передаче пакетов, если это необходимо.

Моделирование и результаты

Для апробации предложенного протокола и его анализа была разработана программная модель и реализованы 2 сценария. Предложенный протокол для динамических туманных вычислений был реализован в среде моделирования NS-3.

Были рассмотрены два основных сценария.

Сценарий 1. Взаимодействие «устройство-Fog» (производилась оценка эффективности предлагаемого протокола, когда устройства выгружают / распределяют свои задачи на FD-n; для этого была рассмотрена топология из пяти устройств UD, подключенных к узлу туманных вычислений).

Сценарий 2. Взаимодействие «Fog-Fog» (рассматривалось десять распределенных узлов туманных вычислений с различной подвижностью, движущихся по квадрату со стороной 10 км).

В таблице 1 представлены параметры, рассматриваемые для настройки моделирования. При выполнении моделирования рассматривались разнообразные данные различных приложений, в частности: датчиков Интернета вещей, обработки изображений, обработки видео и управления дорожным движением и др. Сформированные для моделирования наборы данных были собраны из открытых источников, представленных на ресурсах [7–11].

Задачи были разделены на пять основных категорий по сложности вычислений:

- категория (I), очень простые задачи, включая данные датчиков;
- категория (II), простые задачи, эквивалентные обработке простых изображений и веб-сайтов;

- категория (III), задачи средней сложности, включая 3D-изображения и простые видеоролики;
- категория (IV), сложные задачи, в том числе обработку видео высокого качества;
- категория (V), высокосложные задачи, включая работу с 3D-видео.

ТАБЛИЦА 1. Параметры моделирования

TABLE 1. Simulation Parameters

Параметр	Значение
Сценарий 1: UD-FD	1×1км ²
Сценарий 2: FD-FD	10×10 км ²
Количество устройств UD (N)	5
Количество устройств FD	10
Размещение МЕС	равноудаленное
Скорость перемещения fog-устройств	$V_i \in \{10, 20, 30, 40, 50\}$ км/ч
Дальность передачи	500 м
Максимальный уровень передачи	20 дБм
Память / RAM	2048 Мб
Память / HDD	5 Гб
Процессор / CPU	€ [0.7,2.4] ГГц
Максимальная рабочая нагрузка	50 событий/с
Скорость передачи данных	18 Мб/с

В рамках моделирования задачи случайным образом распределялись между рассматриваемыми устройствами, при этом каждая последующая задача назначалась после выполнения предыдущей на узле.

Для оценки предлагаемого протокола рассматривались три основных показателя производительности:

1) надежность предложенного протокола оценивалась путем измерения количества успешно полученных пакетов; коэффициент доставки пакетов (PDR, аббр. от англ. Packet Delivery Ratio) был рассчитан для предложенного метода и также сравнен с традиционными протоколами;

2) энергоэффективность с точки зрения затрат энергии на задачу;

3) задержка оценивалась через расчет улучшения значения задержки обработки назначенных задач.

Выражение (5) использовалось для расчета процентного улучшения задержки по сравнению с традиционными протоколами.

На рисунке 4 представлены PDR пяти рассматриваемых устройств на разных расстояниях от устройства туманных вычислений. На рисунке

представлены результаты предлагаемого метода и существующего традиционного подхода. PDR отличается от одного устройства к другому из-за их разнородности. Однако все устройства достигли более высокого PDR при использовании предлагаемого протокола. По мере удаления устройства UD от FD fog-узла PDR уменьшается; однако это снижение меньше при использовании предлагаемого протокола.

В результате анализа полученных результатов можно утверждать, что предложенный протокол улучшил PDR в среднем на 34 % по сравнению с существующими традиционными подходами.

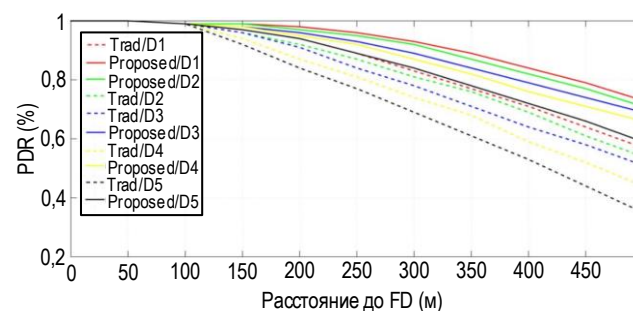


Рис. 4. Значение PDR

Fig. 4. PDR Value

На рисунке 5 представлен средний процент потребляемой энергии конечных устройств (по сравнению с полной батареей) в пяти случаях для трех методов. Три из них являются рассматриваемая топология сети с традиционным протоколом (метод 1), рассматриваемая топология с протоколом, введенным в [12] (метод 2), и топология с предлагаемым протоколом (метод 3).

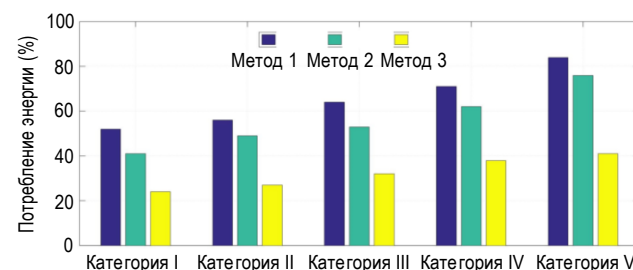


Рис. 5. Потребление энергии

Fig. 5. Energy Consumption

Пять рассмотренных случаев относятся к пяти категориям задач, особенности которых были приведены выше. Во всех пяти случаях предложенный метод позволил потреблять меньше энергии, чем в двух других.

$$\text{Latency improvement (\%)} = \frac{\text{Average latency}_{\text{Traditional}} - \text{Average latency}_{\text{System}(i)}}{\text{Average latency}_{\text{Traditional}}} \quad \forall i \in \mathbb{R}, \quad (5)$$

Стоит отметить, что предложенный протокол значительно снизил потребление энергии при больших рабочих нагрузках. В основном это связано с внедрением микросервисов и их грамотным распределением в туманных вычислениях

В результате анализа полученных данных моделирования можно сделать вывод, что предложенный протокол позволил добиться среднего снижения потребляемой энергии на 41 % по сравнению с методом 1 (традиционный протокол) и на 29 % по сравнению с методом 2 (протокол представлен в работе [12]).

На рисунке 6 представлены значения процента улучшения задержки предлагаемой модели по сравнению с традиционными протоколами для передаваемых данных пяти устройств. Результаты были получены для пяти категорий сервисов (определены выше).

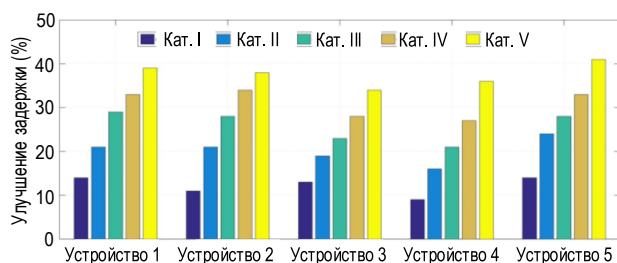


Рис. 6. Процент улучшения задержки

Fig. 6. Percentage of Latency Improvement

Таким образом, предложенный протокол уменьшил задержку для всех типов данных, при этом для

данных микросервиса высококачественного видео, задержка была значительно снижена по сравнению с традиционными протоколами.

В результате проведенного моделирования была доказана эффективность предложенного протокола как с точки зрения уменьшения задержек, так и для сохранения энергоресурсов устройств туманных вычислений.

Заключение

В статье представлены исследования в области динамических туманных вычислений, актуальность которых растет каждый год с учетом требований и характеристик услуг телеприсутствия. Сети IMT-2030 призваны более эффективно работать с сетевыми и вычислительными ресурсами, следовать повестке бережливого производства. Данные цели сформировали целый пласт научных задач в области динамических туманных вычислений.

В работе представлен разработанный метод маршрутизации в сети динамических туманных вычислений, в том числе для задачи миграции микросервисов, позволяющий учитывать особенности и ограничения устройств туманных вычислений. Результаты моделирования нового протокола маршрутизации в среде динамических туманных вычислений показали его эффективность, а именно: снижение потребляемой энергии устройствами туманных вычислений на 41 % и уменьшение доли потерянных пакетов в среднем до 34 %.

Список источников

1. Кучерявый А.Е., Маколкина М.А., Парамонов А.И., Выборнова А.И., Мутханна А.С., Матюхин А.Ю. и др. Модельная сеть для исследований и обучения в области услуг телеприсутствия // Электросвязь. 2022. № 1. С. 14–20. DOI:10.34832/ELSV.2022.26.1.001. EDN:GBQWCV
2. Волков А.Н., Мутханна А.С., Кучерявый А.Е., Бородин А.С., Парамонов А.И., Владимиров С.С. и др. Перспективные исследования сетей и услуг 2030 в лаборатории 6G Meganetlab СПбГУТ // Электросвязь. 2023. № 6. С. 5–14. DOI:10.34832/ELSV.2023.43.6.001. EDN:CJSYLS
3. Volkov A., Muthanna A., Koucheryavy A. AI/machine learning for ultra-reliable low-latency communication // ITU News Magazine. 2020. Iss. 5. PP. 65–68.
4. Лавшук О.А., Листопад Н.И. Метод маршрутизации в сетях PoT с использованием кластеризации для протокола RPL // Проблемы физики, математики и техники. 2023. № 4(57). С. 74–80. DOI:10.54341/20778708_2023_4_57_74. EDN:RVSCGG
5. Mariocco C., Doshi K., Guim Bernat F.G., Smith N.M., Spoczynski M., Verrall T., et al. Methods and apparatus to coordinate edge platforms. Patent US, no. US 2021/0014133 A1, 2021.
6. Тефилова М.Р., Кузьмина Е.А., Волков А.Н. Определение маршрута миграции микросервиса в сети туманных вычислений // Информационные технологии и телекоммуникации. 2023. Т. 11. № 1. С. 50–60. DOI:10.31854/2307-1303-2023-11-1-50-60. EDN:ETUDJB
7. Cruz S., Aguiar A. Cooperative Localization in Vehicular Networks Dataset. IEEEDataPort. 2020. URL: <https://iee-dataport.org/open-access/cooperative-localization-vehicular-networks-dataset> (Accessed 28 June 2024)
8. From Images to 3D Shapes (FI3S). Kaggle. URL: <https://www.kaggle.com/datasets/lehomme/from-images-to-3d-shapesfi3s> (Accessed 20 April 2024)
9. TikTok Dataset // Kaggle. URL: <https://www.kaggle.com/datasets/yasaminjafarian/tiktokdataset> (Accessed 20 April 2024).
10. YouTube Faces With Facial Keypoints. Kaggle. URL: <https://www.kaggle.com/datasets/selfishgene/youtube-faces-with-facial-keypoints> (Accessed 20 April 2024)
11. Google Scraped Image Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/duttadebadri/image-classification> (Accessed 20 April 2024)
12. Ateya A.A., Muthanna A., Gudkova I., Gaidamaka Y., Algarni A.D. Latency and energy-efficient multi-hop routing protocol for unmanned aerial vehicle networks // International Journal of Distributed Sensor Networks. 2019. Vol. 15 Iss. 8. DOI:10.1177/1550147719866392

References

1. Koucheryav A.E., Makolkina M.A., Paramonov A.I., Vybornova A.I., Muthanna A.S., Matyuhin A.Yu., et. al. Model network for research and training in the field of telepresence services. *Electrosvyaz*. 2022;1:14–20. (in Russ.) DOI:10.34832/ELSV.2022.26.1.001. EDN:GBQWCV
2. Volkov A.N., Muthanna A.S.A., Koucheryavy A.E., Borodin A.S., Paramonov A.I., Vladimirov S.S., et. al. Advanced research on networks and services 2030 in the 6G Meganetlab SPbSUT. *Electrosvyaz*. 2023;6:5–14. (in Russ.) DOI:10.34832/ELSV.2023.43.6.001. EDN:CJSYLS
3. Volkov A., Muthanna A., Koucheryavy A. AI/machine learning for ultra-reliable low-latency communication. *ITU News Magazine*. 2020;(5):65–68.
4. Lavshuk O.A., Listopad N.I. Routing method in IIoT networks using clustering for the RPL protocol. *Problems of physics, mathematics and technology*. 2023;4(57):74–80. (in Russ.) DOI:10.54341/20778708_2023_4_57_74. EDN:RVSCGG
5. Mariocco C., Doshi K., Guim Bernat F.G., Smith N.M., Spoczynski M., Verrall T., et all. *Methods and apparatus to coordinate edge platforms*. Patent US, no. US 2021/0014133 A1, 2021.
6. Tefikova M.R., Kuzmina E.A., Volkov A.N. Determination of the route migration for a microservice in a fog computing. *Telecom IT*. 2023;11(1):50–60. (in Russ.) DOI:10.31854/2307-1303-2023-11-1-50-60. EDN:ETUDJB
7. *IEEEDataPort*. Cruz S., Aguiar A. Cooperative Localization in Vehicular Networks Dataset. 2020. URL: <https://iee-dataport.org/open-access/cooperative-localization-vehicular-networks-dataset>. [Accessed 28 June 2024].
8. From Images to 3D Shapes (FI3S). *Kaggle*. URL: <https://www.kaggle.com/datasets/lehomme/from-images-to-3d-shapesfi3s> [Accessed 20 April 2024].
9. *Kaggle*. TikTok Dataset. URL: <https://www.kaggle.com/datasets/yasaminjafarian/tiktokdataset> [Accessed 20 April 2024].
10. *Kaggle*. YouTube Faces With Facial Keypoints. URL: <https://www.kaggle.com/datasets/selfishgene/youtube-faces-with-facial-keypoints> [Accessed 20 April 2024].
11. *Kaggle*. Google Scraped Image Dataset. URL: <https://www.kaggle.com/datasets/duttadebadri/image-classification> [Accessed 20 April 2024].
12. Ateya A.A., Muthanna A., Gudkova I., Gaidamaka Y., Algarni A.D. Latency and energy-efficient multi-hop routing protocol for unmanned aerial vehicle networks. *International Journal of Distributed Sensor Networks*. 2019;15(8). (in Russ.) DOI:10.1177/1550147719866392

Статья поступила в редакцию 10.07.2024; одобрена после рецензирования 09.08.2024; принята к публикации 12.08.2024.

The article was submitted 10.07.2024; approved after reviewing 09.08.2024; accepted for publication 12.08.2024.

Информация об авторе:

ВОЛКОВ
Артем Николаевич

кандидат технических наук, доцент, кафедры сетей связи и передачи данных
Санкт-Петербургского государственного университета телекоммуникаций
им. проф. М.А. Бонч-Бруевича

 <https://orcid.org/0009-0002-4296-1822>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.

Научная статья

УДК 004.942

<https://doi.org/10.31854/1813-324X-2024-10-4-38-47>

Прогнозирование трафика трехмерной сети Интернета вещей высокой плотности как многомерного случайного процесса

Василий Сергеевич Елагин, v.elagin@sut.ru

Александра Андреевна Гребенщикова , grebenshikova.aa@sut.ru

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

Аннотация

Актуальность рассматриваемой в статье тематики состоит в активной трансформации сетей связи и формировании трехмерной сети связи высокой плотности, которая изменяет структуру трафика данных, поэтому и для такого рода сети рассматривается модель трафика как многомерного случайного процесса.

Целью исследования является повышение эффективности прогнозирования сетевого трафика, за счет разработки метода, отличительной особенностью которого является прогнозирование трафика как многомерного случайного процесса с учетом взаимной зависимости отдельных потоков, производимых узлами сети.

Методы. В работе рассмотрен алгоритм обучения искусственной нейронной сети (ИНС) на основе метода уменьшения значения корня средней квадратичной ошибки RMSE, а также предложены методы прогнозирования с использованием ИНС типа LSTM и адаптации параметров модели к изменяющимся условиям функционирования сети. Применение ИНС типа LSTM для прогнозирования многомерного случайного процесса, описывающего трафик в трехмерной сети высокой плотности, может дать лучшие результаты, чем прогнозирование отдельных потоков трафика, как независимых случайных процессов, за счет учета взаимных влияний между различными потоками трафика.

Решение. Построение соответствующей модели, сбор статистики (получение обучающей выборки), обучение ИНС и выполнение прогноза требует использования вычислительных ресурсов. Таким образом эффективность прогнозирования можно определить, как снижение ошибки прогнозирования при сохранении объема используемых ресурсов или уменьшение объема ресурсов при сохранении ошибки прогнозирования. В ходе решения научной задачи выявлены критерии для выбора величины единичного интервала (лага), что в совокупности с интервалом прогнозирования существенно влияет на итоговый сценарий.

Научная новизна работы заключается в оценке изменения ошибки погрешности при прогнозировании трафика трехмерной сети связи высокой плотности как многомерного случайного процесса, по сравнению с представлением прогноза этого же трафика как множества независимых случайных процессов.

Значимость (теоретическая). Эффективность прогнозирования трафика как многомерного случайного процесса в трехмерной сети связи высокой плотности возрастает с увеличением размерности. Таким образом, такой трафик в задачах прогнозирования следует рассматривать как многомерный случайный процесс, размерность которого равна количеству узлов сети, производящих трафик.

Значимость (практическая). Полученные в работе результаты в перспективе могут быть использованы с целью оптимизации функционирования системы управления трафиком.

Ключевые слова: трехмерная сеть связи, сеть связи высокой плотности, трафик, прогнозирование, искусственная нейронная сеть

Ссылка для цитирования: Елагин В.С., Гребенщикова А.А. Прогнозирование трафика трехмерной сети интернета вещей высокой плотности как многомерного случайного процесса // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 38–47. DOI:10.31854/1813-324X-2024-10-4-37-47. EDN:GCJROF

Original research

<https://doi.org/10.31854/1813-324X-2024-10-3-38-47>

Traffic Prediction as a Multidimensional Random Process in a Three-Dimensional High-Density Internet of Things Network

 **Vasiliy S. Elagin**, v.elagin@sut.ru

 **Aleksandra A. Grebenshchikova** , grebenshikova.aa@sut.ru

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

Annotation

Relevance. The relevance of the topic considered in the article lies in the active transformation of communication networks and the formation of a three-dimensional high-density communication network, which changes the structure of data traffic, therefore, for this type of network, a traffic model is considered as a multidimensional random process. **The main purpose** of the study is to improve the efficiency of network traffic forecasting by developing a method, the distinctive feature of which is traffic forecasting as a multidimensional random process, taking into account the mutual dependence of individual flows produced by network nodes.

Methods. The paper considers an algorithm for training an artificial neural network (ANN) based on the method of reducing the root of the mean square error RMSE, and also proposes forecasting methods using LSTM-type ANNs and adapting model parameters to changing network operating conditions. The use of LSTM-type ANN for forecasting a multivariate random process describing traffic in a three-dimensional high-density network can yield better results than forecasting individual traffic flows as independent random processes due to the consideration of mutual influences between different traffic flows.

The results. Building the corresponding model, collecting statistics (obtaining a training sample), training the ANN and performing the forecast require the use of computing resources. Thus, the forecasting efficiency can be defined as a decrease in the forecasting error while maintaining the volume of resources used or a decrease in the volume of resources while maintaining the forecasting error. In the course of solving the scientific problem, criteria were identified for selecting the value of a unit interval (lag), which, together with the forecasting interval, significantly affects the final scenario.

The theoretical significance The scientific novelty of the work lies in the assessment of the change in the error in forecasting the traffic of a three-dimensional high-density communication network as a multivariate random process, compared to presenting the forecast of the same traffic as a set of independent random processes.

Significance (theoretical). The efficiency of traffic forecasting as a multidimensional random process in a three-dimensional high-density communication network increases with increasing dimensionality. Thus, such traffic in forecasting problems should be considered as a multidimensional random process, the dimensionality of which is equal to the number of network nodes producing traffic.

Significance (practical). The results obtained in the work can be used in the future to optimize the functioning of the traffic management system.

Keywords: three-dimensional network, high-density network, traffic prediction, neural networks

For citation: Elagin V.S., Grebenshchikova A.A. Traffic Prediction as a Multidimensional Random Process in a Three-Dimensional High-Density Internet of Things Network. *Proceedings of Telecommunication Universities*. 2024;10(4): 38–47. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-38-47. EDN:GCJROF

Введение

Развитие сетей беспроводной связи и беспроводного доступа приводит к высокой концентрации узлов сети в пространстве. Например, в многоквартир-

тирных зданиях в зоне связи каждого из узлов сети может оказаться до нескольких десятков аналогичных узлов [1]. Таким образом, в подобных сетях создаются условия взаимного влияния между элементами сети. Потоки трафика, производимого

пользователями этих узлов, оказываются зависимыми. Степень этой зависимости определяется различными факторами, одним из которых является расстояние между узлами сети. В трехмерной сети наибольшее влияние оказывают ближайшие соседние узлы [2].

При построении сети с использованием принципов SDN [3], в целях управления потоками трафика и ресурсами сети необходимо получение прогнозируемых значений трафика. Для этого применяются различные методы, основанные на анализе статистики [4]. В большинстве случаев эти методы предполагают рассмотрение потоков, производимых отдельными узлами сети, как независимых случайных процессов [5]. Среди методов прогнозирования хорошо известны инструменты статистического анализа, например, ARIMA [6], а также применение искусственных нейронных сетей (ИНС). Широкое распространение в задачах прогнозирования получили модели ИНС типа LSTM (аббр. от англ. Long Short-Term Memory), которые показали положительные результаты в решении подобных задач [7], благодаря их «способности запоминать» характерные особенности случайных процессов.

Применение ИНС типа LSTM для прогнозирования многомерного случайного процесса, описывающего трафик в трехмерной сети высокой плотности, может дать лучшие результаты, чем прогнозирование отдельных потоков трафика, как независимых случайных процессов, за счет учета взаимных влияний между различными потоками трафика. Эти результаты можно оценить по уменьшению ошибки получаемого прогноза, которое отражается на эффективности функционирования сети (распределения трафика и ресурсов). Поэтому его можно рассматривать как способ повышения эффективности функционирования сети в целом.

В данной работе предлагается модель трафика трехмерной сети высокой плотности как многомерного случайного процесса и метод использования ИНС типа LSTM для его прогнозирования.

Модель трафика

Трафик в сетях, построенных с применением беспроводной связи, создается пользователями, которые случайным образом распределены в зоне обслуживания. В зависимости от способа построения сети, трафик данных передается между пользовательскими терминалами и базовой станцией или точкой доступа, или непосредственно между терминалами пользователей. При этом в каждой точке пространства имеют место сигналы всех узлов сети (базовых станций и терминалов), которые отличаются по уровню мощности. Одни из них могут быть приняты приемными устройствами, а другие имеют слишком малый уровень мощности и воспринима-

ются как дополнительный источник помех (шума) или вообще могут быть проигнорированы, если их уровень мощности слишком мал.

Таким образом, если рассматривать некоторое пространство, обслуживаемое сетью связи, то в каждой точке пространства можно наблюдать потоки трафика между элементами сети. Выбор точки наблюдения определяется целью задачи. Если рассматриваемая сеть построена с использованием точек доступа (базовых станций), то особую роль играют именно точки их размещения, т. к. в такой сети имеют место только потоки трафика между пользователями и точками доступа.

В трехмерной сети точки доступа распределены в пространстве, например, как это показано на рисунке 1.

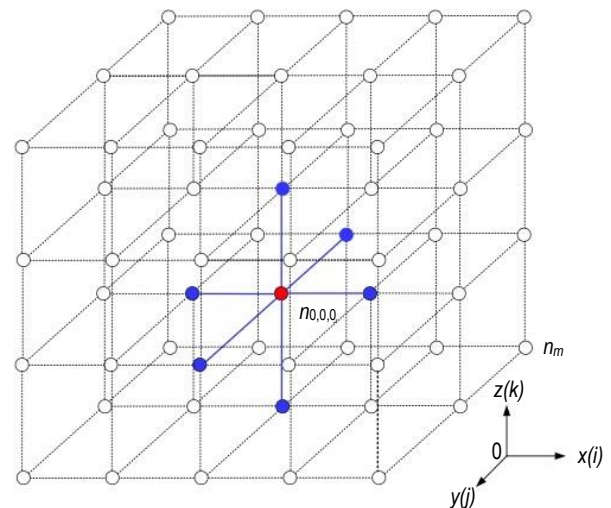


Рис. 1. Пример размещения точек доступа сети в пространстве

Fig. 1. An Example of the Placement of Network Access Points

Трафик в такой сети – это трафик, производимый пользователями, подключенными к ее точкам доступа. В каждой из этих точек трафик представляет собой случайный поток пакетов (кадров, пакетов или сегментов, в зависимости от уровня рассмотрения, для общности будем называть пакетами). В общем случае эти потоки зависимы, так как передаваемые сигналы влияют на процесс передачи в соседних узлах, если используется один частотный канал.

Если две точки доступа оказываются в зоне связи друг друга, то они вынуждены распределять время своей активности так, чтобы интервалы их активности не совпадали и не перекрывались. Это приводит к возникновению зависимостей между потоками трафика в таких точках доступа. Аналогичная задача решается и когда точек доступа больше, чем две. В рассматриваемой модели их может быть гораздо больше. Для примера, в сетях семейства стандартов IEEE 802.11 [8] в многоквартирном жилом здании количество точек доступа,

сигнал от которых принимается с достаточно высоким уровнем мощности, может исчисляться несколькими десятками.

Поэтому трафик в трехмерной сети высокой плотности можно рассматривать как многомерный случайный процесс [9] в виде:

$$X^{(k)} = \{X_1(t), \dots, X_k(t)\}, \quad (1)$$

описываемый совместными функциями распределения:

$$F^{(k)} = \{F_1(x_1, x_2 \dots x_k, t), \dots, F_k(x_1, x_2 \dots x_k, t)\}, \quad (2)$$

где $X_1(t), \dots, X_k(t)$ – случайные процессы в выбранных точках (точках доступа) сети, которые представляют собой количество данных (пакетов), переданных за заданный интервал времени (лаг, от англ. Lag – запаздывание).

В случае независимости $X_1(t), \dots, X_k(t)$ каждый из процессов может быть описан своей функцией распределения, например, распределением Пуассона. Однако в общем случае эти процессы зависимы. Эта зависимость тем выше, чем ближе рассматриваемые точки. Иными словами, в конкретной точке сети трафик в наибольшей степени зависит от трафика ближайших соседей (точек доступа) и в меньшей степени – от удаленных. Таким образом, общая картина трафика в сети может быть описана многомерным процессом (1) и его функциями распределения (2).

При построении SDN для управления потоками трафика требуется информация о трафике и его прогнозируемых значениях. Для решения данной задачи предлагается использование ИНС типа LSTM. Применение ИНС для прогнозирования множества зависимых случайных процессов позволяет учесть, как влияние времени, так и взаимные влияния этих процессов друг на друга. Согласно [10, 11], ИНС типа LSTM используются для прогнозирования трафика в сотовых сетях и в транспортных системах. Также в исследовании [12] отмечается специализация LSTM в прогнозировании временных рядов благодаря ячеистой архитектуре. В таблице 1 представлены преимущества и недостатки моделей прогнозирования для временных рядов.

Таким образом, использование иных методов для решения задачи прогнозирования многомерного процесса либо не эффективно, либо слишком сложно [13]. Традиционные методы прогнозирования с использованием ARIMA, SARIMA и подобные, позволяют получать удовлетворительные прогнозы по отношению лишь к фактору времени, при наличии взаимной зависимости процессов прогноз оказывается менее эффективным.

Процесс прогнозирования трафика в 3D-сети высокой плотности можно описать последовательностью обработки трафика, как показано на рисунке 2.

ТАБЛИЦА 1. Недостатки и преимущества моделей прогнозирования

TABLE 1. The Advantages and Disadvantages of Forecasting Models

Модель прогнозирования	Недостатки	Преимущества
SARIMA ARIMA ARMA	Прогнозирование данных основано исключительно на прошлых значениях	Высокая интерпретируемость параметров
NN RNN LSTM GRU CNN GNN	– низкая интерпретируемость параметров; – требуется большое количество гиперпараметров; – высокие вычислительные затраты на этапе обучения	Согласно научной литературе, демонстрирует лучшие результаты по сравнению со статистическими подходами

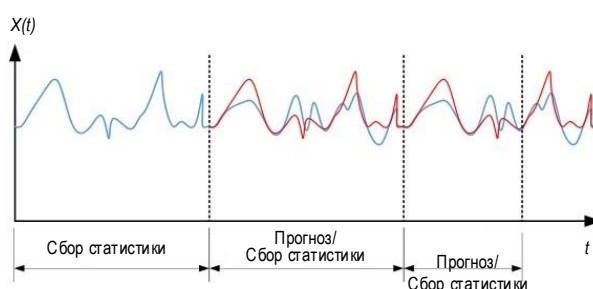


Рис. 2. Сбор статистики и прогнозирование

Fig. 2. Collecting Statistics and Forecasting

Такая последовательность состоит из двух непрерывно повторяющихся фаз: сбора статических данных (результатом которой является выборка значений D_T) и прогнозирования (результатом которого является выборка значений D_F).

Качество функционирования такой системы определяется достоверностью получаемых прогнозов.

Для оценки эффективности полученного прогноза чаще всего применяется средняя разница между значениями, спрогнозированными моделями, и фактическими значениями трафика – корень средней квадратичной ошибки (RMSE) [14]:

$$RMSE_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (x_j^{(i)} - \hat{x}_j^{(i)})^2}, \quad (3)$$

где x_i – значение процесса на i -м отсчете интервала прогнозирования из выборки статистики. $D_T^{(i)} = \{x_1^{(i)} \dots x_n^{(i)}\}$, $x_1^{(i)} \dots x_n^{(i)} \in X_i(t)$, $i = 1 \dots k$, $t \in T$, $\hat{x}_j^{(i)}$ – значение на i -м отсчете интервала прогнозирования из данных прогноза $D_F^{(i)} = \{\hat{x}_1^{(i)} \dots \hat{x}_n^{(i)}\}$, n – количество отсчетов; $D_T^{(i)}$ – выборка значений одного из случайных процессов $X_1(t), \dots, X_k(t)$; $D_F^{(i)}$ – его прогнозируемые значения.

Наряду с RMSE могут использоваться и другие оценки, такие как среднеквадратичная ошибка MSE

(аббр. от англ. Mean Square Error), средняя абсолютная ошибка MAE (аббр. от англ. Mean Absolute Error), коэффициент детерминации R-квадрат. Все эти оценки характеризуют разницу между прогнозируемыми значениями и фактической реализацией процесса, а также между данными обучения и выходными значениями ИНС. Использование различных оценок несколько отражается на получаемом результате, и их выбор является отдельной задачей. В данной работе используется наиболее распространенная оценка RMSE. Хотя следует отметить, что обсуждаемый подход к прогнозированию не зависит принципиально от выбранной оценки.

Оценка RMSE (или любая иная оценка из упомянутых выше) зависит от множества факторов, к которым можно отнести: свойства случайного (прогнозируемого) процесса, продолжительность интервала прогнозирования, количество отсчетов на интервале прогнозирования, продолжительность интервала (интервалов) сбора статистики и количество отсчетов на нем.

Масштаб времени

Выбор продолжительности интервала прогнозирования T_F и единичных интервалов (лагов) существенно влияет на результаты прогнозирования. Величина единичных интервалов времени влияет на оценку свойств потока (процесса). В частности, влияние величины единичных интервалов на свойства случайного процесса характеризуется степенью самоподобия [15]. Чем выше степень самоподобия, тем меньше это влияние. Рассмотрим выбор величины единичного интервала с различных позиций. Во-первых, с точки зрения выбора величины минимального и максимального интервала. А во-вторых, с точки зрения требований со стороны системы управления трафиком.

Минимальная величина единичного интервала определяет детальность представления трафика. Чем меньше эта величина, тем детальнее будет представлен процесс поступления заявок (пакетов, кадров или сегментов). Если устремить эту величину к нулю, то количество пакетов, поступивших за интервал, не превысит единицы (если принимать во внимание момент начала приема / передачи пакета), так как в линии связи пакеты передаются последовательно, а время передачи самого короткого пакета превышает продолжительность этого интервала. В этом случае случайный процесс будет представлен последовательностью нулей и единиц. И в реальном трафике нулей будет гораздо больше, чем единиц.

Таким образом, минимальная величина интервала – это величина, не превышающая время передачи самого короткого пакета:

$$u_{\min} \leq \min \left(\frac{L_{\min}}{b} \right), \quad (5)$$

где $L_{\min}$ – минимальный размер пакета (бит); b – скорость передачи данных (бит/с).

Максимальная величина единичного интервала ограничена продолжительностью наблюдений и требованиями к объему статистического материала. Поскольку для решения задачи прогнозирования требуется набор статистических данных (обучающая выборка), то это требование можно интерпретировать как требование к максимальной величине единичного интервала:

$$u_{\max} = \frac{T_t}{n_t}, \quad (6)$$

где n_t – количество значений в обучающей выборке; T_t – продолжительность наблюдений (продолжительность получения обучающей выборки, сбора статистики).

Интервал получения обучающей выборки T_t должен быть репрезентативен с точки зрения решаемой задачи. Репрезентативность процесса обеспечивается достаточной детализацией процесса изменения трафика. Если дискретный процесс $X(t)$ является представлением некоего непрерывного процесса, то для полного сохранения информации, согласно теореме В.А. Котельникова, величину, обратную единичному интервалу $1/u$, необходимо интерпретировать как удвоенную верхнюю частоту спектра непрерывного процесса. В этом смысле величина единичного интервала не должна ограничивать верхнюю частоту спектра непрерывного процесса, а, следовательно, и его информативность.

Количество данных x_u , переданных за интервал u , случайно и зависит от средней интенсивности трафика. Среднее количество данных будет определяться как:

$$\bar{x}_u = u \bar{a}_u. \quad (7)$$

Очевидно, что значения этой величины могут лежать в пределах от нуля до величины скорости передачи данных в линии связи b :

$$0 < x_u \leq b, \quad (8)$$

где b – скорость передачи данных.

Пусть $f(x_u)$ – плотность вероятности случайной величины.

Рассматривая эту величину с позиции информативности представления данных, количество информации (бит) в случайном процессе можно определить как:

$$h_x = \int_{-\infty}^{\infty} f(x) \log_2 f(x) dx. \quad (9)$$

Согласно [16], связь с единичным интервалом u можно установить через коэффициент Херста:

$$m^{1-H} F^{(m)}(x) \doteq F(x), \quad m \in \mathbb{N}, \quad (10)$$

где H – коэффициент Херста; $H = 1 - \left(\frac{\beta}{2}\right)$; знак $\doteq$ обозначает равенство распределений; $F(x)$ – распределение исходного процесса (полученного на интервалах u); $F^{(m)}(x)$ – распределение процесса, полученного на m интервалах u .

С ростом коэффициента Херста ($0 < H < 1$) распределения вероятностей исходного процесса и процесса, полученного при увеличении интервалов u в m раз, сближаются (10).

Из определения самоподобности второго порядка и коэффициента корреляции $r(k)$ следует, что для $0 < H < 1$ при $H \neq 0,5$:

$$r(k) \sim H(2H - 1)k^{2H-2}, \quad (11)$$

где $k \rightarrow \infty$.

Согласно [17], если $0,5 < H < 1$, $r(k) \sim ck^{-\beta}$, $c > 0$ – константа, $\beta = 2 - 2H$ и $0 < \beta < 1$, тогда коэффициент корреляции медленно затухает, что является основной причиной несуммируемости и исходный процесс является долговременно зависимым:

$$\sum_{k=-\infty}^{\infty} r(k) = \infty. \quad (12)$$

Для равенства (10) при $0 < \beta < 1$ исходный процесс является строго самоподобным в широком смысле и имеет медленно убывающую зависимость [16].

Требования со стороны системы управления трафиком могут иметь первостепенное значение при выборе величины u . Функционирование системы управления предполагает сбор статистики (обучение), получение прогнозов, анализ состояния сети и принятие решения по управлению трафиком. Каждый из этих этапов имеет свои характеристики и требует определенных затрат времени. Это проявляется в определенных характеристиках управления. Тогда величина u должна соответствовать времени реакции системы управления (интервалу управления).

Таким образом, обобщая сказанное, можно сделать вывод, что выбор величины единичного интервала u (лага) зависит от таких факторов, как параметры канала связи, свойства трафика и характеристики системы управления.

Метод использования искусственных нейронных сетей

При использовании ИНС $RMSE$ зависит от объема обучающей выборки D_T . Отрицательное влияние на прогноз может оказать как недостаточное, так и избыточное обучение, поэтому необходимо определить некоторый размер этой выборки, при котором ошибка минимальна [18].

Выбор параметров прогнозирования можно определить, как задачу оптимизации, целевой функцией (%) в которой является:

$$O = \min_{T_t, m} RMSE(T_t, m), \quad (13)$$

где T – продолжительность обучающей выборки (размер выборки); m – количество эпох обучения.

Прогнозирование трафика – это непрерывный процесс, который выполняется в течение всего времени функционирования сети связи. Случайные процессы могут претерпевать как краткосрочные, так и медленные (долговременные) изменения. Для того, чтобы учитывать актуальную ситуацию, обучение ИНС должно производиться непрерывно с учетом требований к продолжительности периода прогнозирования. Соответствующий алгоритм приведен на рисунке 3.

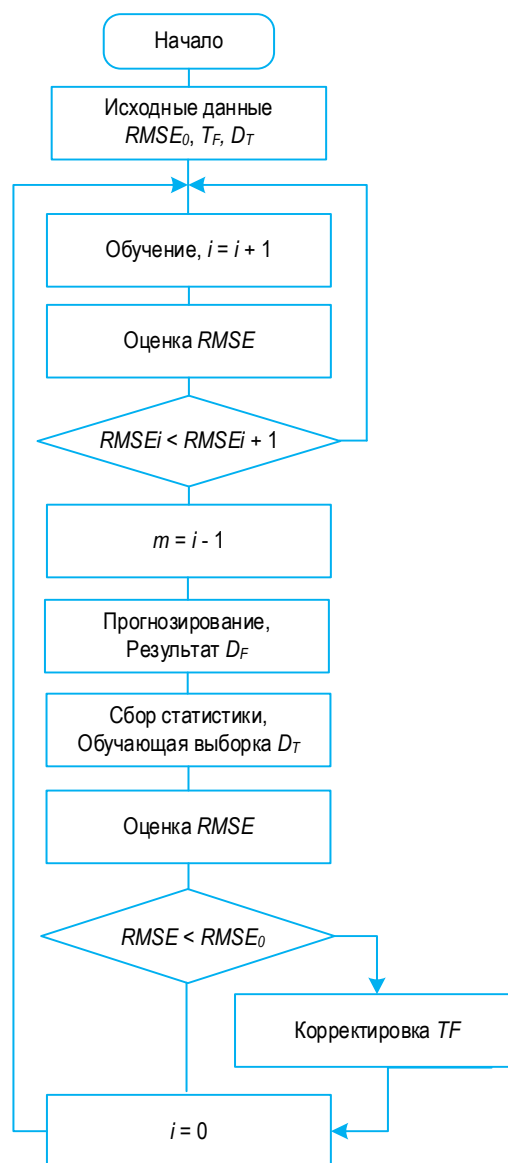


Рис. 3. Алгоритм обучения ИНС

Fig. 3. The Learning Algorithm of Neural Networks

На этапе инициализации может задаваться начальная обучающая выборка D_t продолжительностью T_t . Если такой выборки нет, то для ее получения проводится сбор статистики и непосредственное формирование этой выборки. Далее начинается циклическое обучение ИНС на основе начальной выборки, которое проводится за несколько эпох (циклов обучения). Критерием завершения обучения является отсутствие уменьшения значения $RMSE$. Если после очередной эпохи ошибка не снизилась, то обучение считается оконченным при числе эпох, равном m . Далее ИНС готова выполнить прогноз, что делается на следующем этапе – получение набора прогнозируемых значений D_F .

Полученный прогноз D_F может быть использован как результат в целях управления трафиком, а его качество оценивается сравнением с реальными данными, получаемыми в результате сбора статистики о трафике D_t , на основе которых вычисляется оценка $RMSE$. Если полученная оценка свидетельствует о низком качестве прогнозирования, то производится корректировка интервала обучения T_t .

На рисунке 4 приведен пример прогнозирования случайного процесса, представляющего собой общий трафик в двух взаимозависимых точках доступа в пределах одного дня с 22:22:49 до 22:41:13.

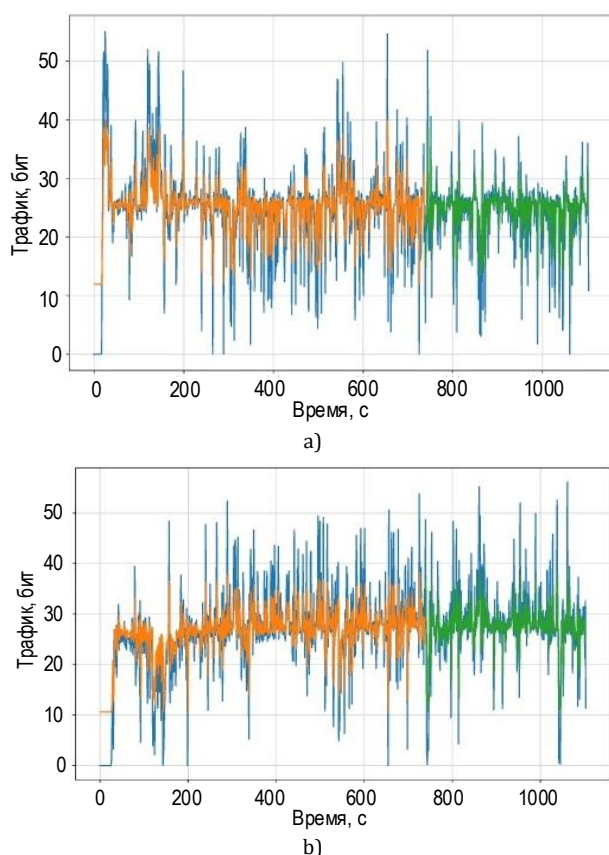


Рис. 4. Пример прогнозирования трафика с использованием ИНС типа LSTM: а) $RMSE = 6,28$; б) $RMSE = 7,29$

Fig. 4. An Example of Traffic Forecasting Using an LSTM: а) $RMSE = 6,28$; б) $RMSE = 7,29$

Процесс (см. рисунок 4) представляет собой набор значений числа бит, переданных за интервалы, равные одной секунде.

На приведенных примерах показан двухмерный случайный процесс, включающий в себя два потока трафика (две точки доступа). Синяя кривая представляет полученные в результате измерений данные, оранжевая – аппроксимация исходного процесса ИНС после обучения, зеленая – прогнозируемый процесс на фоне реализации (синяя кривая). Полученные результаты продемонстрировали высокую точность предсказания на достаточно продолжительном интервале. Объем тренировочной выборки состоял из 740 наблюдений. Продолжительность интервала прогнозирования составила в данном примере 33 % от общей продолжительности наблюдений в 1104 отсчетов. Ошибка составила для процесса на рисунке 4а: $RMSE = 6,28$, а на рисунке 4б: $RMSE = 7,29$. Следует отметить, что из графиков визуально заметна достаточно высокая точность прогнозирования. Полученный процесс повторяет колебания реального процесса, благодаря тому, что ИНС находит сложную закономерность, незаметную на взгляд.

Эффективность метода

Целью прогнозирования случайного процесса является получение значения случайной величины (последовательности или множества случайных величин), которые будут иметь место в будущем. Качество прогнозирования оценивается метрикой, характеризующей ошибку, т. е. разницу между прогнозируемым значением и фактическим значением в момент, на который выполнен прогноз. Для решения задачи прогнозирования с помощью ИНС необходимо построение соответствующей модели, сбор статистики (получение обучающей выборки), обучение ИНС и выполнение прогноза. При этом расходуются вычислительные ресурсы. В таком случае эффективность прогнозирования можно определить, как снижение ошибки прогнозирования при сохранении объема используемых ресурсов или уменьшение объема ресурсов при сохранении ошибки прогнозирования.

В данном случае будем использовать первый подход, т. е. будем оценивать изменение ошибки прогнозирования при прогнозировании трафика трехмерной сети связи высокой плотности, как многомерного случайного процесса по сравнению с прогнозированием этого же трафика как множества независимых случайных процессов.

Будем оценивать ошибку прогнозирования с помощью $RMSE$, вычисляемого согласно (3).

Выражение (9) используется в обоих случаях, т. е. для случая прогнозирования трафика как многомер-

ного случайного процесса при параллельном прогнозировании k временных рядов, так и для случая прогнозирования трафика как k независимых временных рядов. Соответствующие ошибки прогнозирования обозначим как $RMSE_D$ и $RMSE_I$. Тогда эффективность прогнозирования трафика трехмерной сети высокой плотности, как многомерного случайного процесса будем оценивать в %:

$$E = 100 \frac{RMSE_D - RMSE_I}{RMSE_I}. \quad (14)$$

Значение (14) отражает относительное уменьшение ошибки прогнозирования за счет использования «параллельного» прогнозирования многомерного процесса, по сравнению с прогнозированием независимых случайных процессов. На рисунке 5 приведены результаты оценки эффективности прогнозирования трафика как многомерного процесса, где k – размерность.

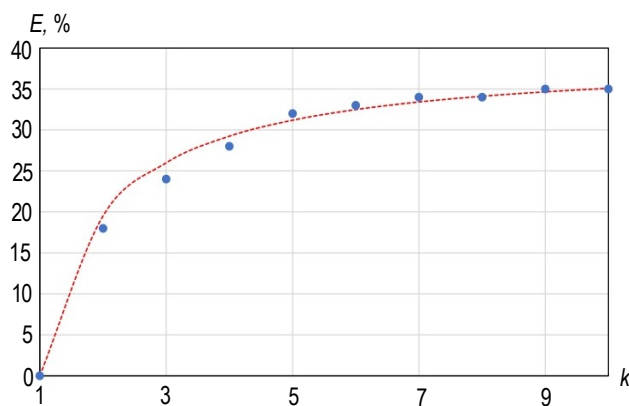


Рис. 5. Оценка эффективности прогнозирования трафика как многомерного случайного процесса, согласно (14)

Fig. 5. The Effectiveness of Traffic Forecasting as a Multidimensional Random Process (14)

Из полученного графика видно, что эффективность прогнозирования трафика как многомерного случайного процесса в трехмерной сети связи высокой плотности возрастает с увеличением размерности. Очевидно, что при $k = 1$ эффективность равна нулю, так как в этом случае процесс является

одномерным. При росте размерности k эффективность метода возрастает, но скорость ее роста снижается. Такая зависимость вполне понятна, так как область взаимных влияний между отдельными потоками трафика в трехмерной сети не бесконечна, как было отмечено выше, она определяется ближайшими соседними узлами. В рассмотренном примере эффективность метода составила 35 % при размерности $k = 10$.

Заключение

Из полученных данных следует, что трафик трехмерной сети высокой плотности в задачах прогнозирования целесообразно рассматривать как многомерный случайный процесс, размерность которого равна количеству узлов сети, производящих трафик. Такой подход позволяет учесть взаимную зависимость потоков трафика, производимых соседними узлами сети. Прогнозирование сетевого трафика как многомерного процесса с помощью ИНС демонстрирует хорошую точность долгосрочного прогноза, в отличие от параметрических моделей типа ARIMA, которые применяются для краткосрочного прогнозирования одномерных процессов.

Качество прогнозирования трафика трехмерной сети высокой плотности как многомерного случайного процесса может быть оценено с помощью суммарной среднеквадратической ошибки, вычисляемой по всем измерениям случайного процесса. Продуктивность использования модели многомерного случайного процесса для описания трафика трехмерной сети высокой плотности может быть оценена относительно изменением суммарной среднеквадратической ошибки. Ее применение позволяет повысить эффективность прогнозирования, согласно приведенному примеру, до 35 % (при размерности $k = 10$).

Эффективность прогнозирования предложенным методом зависит от размерности случайного процесса, возрастая с увеличением размерности; скорость увеличения эффективности зависит от характеристик сети связи, а именно – от количества взаимных влияний между узлами сети.

Список источников

1. 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:10.31854/2307-1303-2022-10-3-1-12. EDN:LHLYEM
2. Бушеленков С.Н., Парамонов А.И. Анализ и формирование структуры сети интернета вещей на основе моделей решеток // Электросвязь. 2021. № 7. С. 23–28. DOI:10.34832/ELSV.2021.20.7.002. EDN:JIBLHT
3. Volkov A., Abdellah A.R., Muthanna A., Makolkina M., Paramonov A., Koucheryavy A. IoT traffic prediction with neural networks learning based on SDN infrastructure // Proceedings of 23rd International Conference on Distributed and Computer and Communication Networks (DCCN 2020, Moscow, Russia, 14–18 September 2020). Lecture Notes in Computer Science. 2020. Vol. 12563. IEEE, 2020. PP. 64–76. DOI:10.1007/978-3-030-66471-8_6. EDN:QRQIPV
4. Абделлах А.Р., Махмуд О.А., Парамонов А.И. Кучерявый А.Е. Прогнозирование задержки в сетях интернета вещей и тактильного интернета с использованием машинного обучения // Электросвязь. 2021. № 1. С. 23–27. DOI:10.34832/ELSV.2021.14.1.002. EDN:IRYPFV

5. Вентцель А.Д. Курс теории случайных процессов. М.: Наука. Физматлит, 1996. 399 с.
6. Joshi M., Hadi T.H. A Review of Network Traffic Analysis and Prediction Techniques. 2015. URL: <https://arxiv.org/abs/1507.05722> (Accessed 25.08.24)
7. Ferreira G.O., Ravazzi C., Dabbene F., Calafiore G.C., Fiore M. Forecasting Network Traffic: A Survey and Tutorial With Open-Source Comparative Evaluation // IEEE Access. 2023. Vol. 11. PP. 6018–6044. DOI:10.1109/ACCESS.2023.3236261
8. Викулов А.С., Парамонов А.И. Постановка задачи замощения плоскости в применении к частотно-территориальному планированию сетей IEEE 802.11 // Радиотехнические и телекоммуникационные системы. 2021. № 1(41). С. 24–32. EDN:HNRRQO
9. Миллер Б.М., Панков А.Р. Теория случайных процессов в примерах и задачах. М.: Физматлит, 2007. 320 с.
10. Manzoor S., Bajwa K.B., Sajid M., Manzoor H., Manzoor M., Ali N., et. al. Modeling of Wireless Traffic Load in Next Generation Wireless Networks // Mathematical Problems in Engineering. 2021. Iss. 1. P. 15. DOI:10.1155/2021/7293093
11. Wang J., Tang J., Xu Z., Wang Y., Xue G., Zhang X., et. al. Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach // Proceeding of IEEE Conference on Computer Communications (Atlanta, USA, 01–04 May 2017). IEEE, 2017. PP. 1–9. DOI:10.1109/INFOCOM.2017.8057090
12. Luo X., Li D., Yang Y., Zhang S. Spatiotemporal traffic flow prediction with KNN and LSTM // Journal of Advanced Transportation. 2019. PP. 1–10. DOI:10.1155/2019/4145353
13. Hochreiter S., Schmidhuber J. Long short-term memory // Neural computation. 1997. Vol. 9. Iss. 8. PP. 1735–80. DOI:10.1162/neco.1997.9.8.1735
14. Kavitha A.K., Praveena S.M. Deep learning model for traffic flow prediction in wireless network // Automatika. 2023. Vol. 64. Iss. 4. PP. 848–857. DOI:10.1080/00051144.2023.2220203
15. Шелухин О.И., Осин А.В., Смольский С.М. Самоподобие и фракталы. Телекоммуникационные приложения. М.: Физматлит, 2008. 368 с. EDN:MVSAB
16. Цыбаков Б.С. Модель телетрафика на основе самоподобного случайного процесса // Радиотехника. 1999. № 5. С. 24–31.
17. Шелухин О.И., Тенякшев А.М., Осин А.В. Фрактальные процессы в телекоммуникациях. М.: Радиотехника, 2003. 480 с. EDN:UKUJXH
18. Sjöberg J., Ljung L. Overtraining, Regularization, and Searching for Minimum in Neural Networks. 1992. URL: https://www.researchgate.net/profile/J-Sjoberg/publication/2738519_Overtraining_Regularization_And_Searching_For_Minimum_With_Application_To_Neural_Networks/links/02e7e519f37e9a2f6e000000/Overtraining-Regularization-And-Searching-For-Minimum-With-Application-To-Neural-Networks.pdf (Accessed 26.08.2024)

References

1. Koucheryavy A., Paramonov A., Makolkina A., Muthanna A. S. A., Vybornova A., Dunaytsev R., et. al. 3 Dimension Multi-layer Heterogenous Ultra Dense Networks. *Telecom IT*. 2022;10(3):1–12. (in Russ.) DOI:10.31854/2307-1303-2022-10-3-1-12. EDN:LHLYEM
2. Bushelenkov S.N., Paramonov A.I. Analysis and formation of the structure of the internet of things network based on lattice models. *Telecommunications*. 2021;7:23–28. (in Russ.) DOI:10.34832/ELSV.2021.20.7.002. EDN:JIBLHT
3. Volkov A., Abdellah A.R., Muthanna A., Makolkina M., Paramonov A., Koucheryavy A. IoT traffic prediction with neural networks learning based on SDN infrastructure. *Proceedings of 23rd International Conference on Distributed and Computer and Communication Networks, DCCN 2020, 14–18 September 2020, Moscow, Russia. Lecture Notes in Computer Science, vol.12563*. IEEE; 2020. p.64–76. DOI:10.1007/978-3-030-66471-8_6. EDN:QRQIPV
4. Abdellah A.R., Mahmud O.A., Paramonov A.I., Koucheryavy A. Delay prediction in IoT and tactile internet using machine learning approach. *Telecommunications*. 2021;1:23–27. (in Russ.) DOI:10.34832/ELSV.2021.14.1.002. EDN:IRYPFV
5. Wentzell A.D. *A Course in the Theory of Stochastic Processes*. Moscow: Nauka Publ.; 1996. 399 p. (in Russ.)
6. Joshi M., Hadi T.H. *A Review of Network Traffic Analysis and Prediction Techniques*. 2015. URL: <https://arxiv.org/abs/1507.05722> [Accessed 25.08.24]
7. Ferreira G.O., Ravazzi C., Dabbene F., Calafiore G.C., Fiore M. Forecasting Network Traffic: A Survey and Tutorial With Open-Source Comparative Evaluation. *IEEE Access*. 2023;11:6018–6044. DOI:10.1109/ACCESS.2023.3236261
8. Vikulov A.S., Paramonov A.I. Problem statement of tiling the plane for frequency and area planning of IEEE 802.11 networks. *Radio Engineering and Telecommunication Systems*. 2021;1(41):24–32. (in Russ.) EDN:HNRRQO
9. Miller B.M., Pankov A.S. *Theory of Random Processes in Examples and Problems*. Moscow: Fizmatlit Publ.; 2007. 320 c. (in Russ.)
10. Manzoor S., Bajwa K.B., Sajid M., Manzoor H., Manzoor M., Ali N., et. al. Modeling of Wireless Traffic Load in Next Generation Wireless Networks. *Mathematical Problems in Engineering*. 2021;1:15. DOI:10.1155/2021/7293093
11. Wang J., Tang J., Xu Z., Wang Y., Xue G., Zhang X., et. al. Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach. *Proceeding of IEEE Conference on Computer Communications, 01–04 May 2017, Atlanta, USA*. IEEE; 2017. p.1–9. DOI:10.1109/INFOCOM.2017.8057090
12. Luo X., Li D., Yang Y., Zhang S. Spatiotemporal traffic flow prediction with KNN and LSTM. *Journal of Advanced Transportation*. 2019;1–10. DOI:10.1155/2019/4145353
13. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural computation*. 1997;9(8):1735–80. DOI:10.1162/neco.1997.9.8.1735
14. Kavitha A.K., Praveena S.M. Deep learning model for traffic flow prediction in wireless network. *Automatika*. 2023;64(4):848–857. DOI:10.1080/00051144.2023.2220203

15. Sheluhin O.I., Smolskiy S.M., Osin A. V. *Self-Similar Processes in Telecommunications*. Moscow: Fizmatlit Publ; 2008. 368 p. (in Russ.) EDN:MVSWAB
16. Tsybakov B.S. Teletraffic model based on self-similar random process. *Radiotekhnica*. 1999;5:24–31. (in Russ.)
17. Sheluhin O.I., Tenjakshev A.M., Osin A.V. *Fractal Processes in Telecommunications*. Moscow: Radiotekhnika Publ.; 2003. 480 p. (in Russ.) EDN:UKUJXH
18. Sjöberg J., Ljung L. *Overtraining, Regularization, and Searching for Minimum in Neural Networks*. 1992. URL: https://www.researchgate.net/profile/J-Sjoberg/publication/2738519_Overtraining-Regularization_And_Searching_For_Minimum_With_Application_To_Neural_Networks/links/02e7e519f37e9a2f6e000000/Overtraining-Regularization-And-Searching-For-Minimum-With-Application-To-Neural-Networks.pdf [Accessed 26.08.2024]

Статья поступила в редакцию 10.07.2024; одобрена после рецензирования 29.07.2024; принята к публикации 05.08.2024.

The article was submitted 10.07.2024; approved after reviewing 29.07.2024; accepted for publication 05.08.2024.

Информация об авторах:

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>

ГРЕБЕНЩИКОВА
Александра Андреевна

ассистент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0008-3118-9957>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

Научная статья

УДК 621.396

<https://doi.org/10.31854/1813-324X-2024-10-4-48-61>

Реализация высококомобильного комплекса пеленгации для БПЛА с применением виртуальных магнитных диполей

Евгений Алексеевич Ищенко✉, kursk1998@yandex.ru

Юрий Геннадьевич Пастернак, pasternakyg@mail.ru

Сергей Михайлович Фёдоров, fedorov_sm@mail.ru

Илья Андреевич Баранников, 8thbar@gmail.com

Воронежский государственный технический университет,
Воронеж, 394006, Российская Федерация

Аннотация

Актуальность. В условиях развития современных широкополосных систем связи особенно актуальной становится задача выполнения пеленгации сигналов с использованием высококомобильных комплексов пеленгации (БПЛА), которые накладывают ограничения на размеры антенных элементов, на расстояние между ними, что приводит к росту математической ошибки пеленга.

Постановка задачи. В статье ставится задача рассмотреть возможность повышения точности пеленгации высококомобильного комплекса путем применения технологии виртуальных магнитных диполей. Особенностью данного метода является минимизация искажений, которые вносятся корпусом-носителем в характеристики измеряемого поля. Для измерения характеристик поля, а также их парциальных компонент применялись векторные антенные элементы.

Целью работы является исследование характеристик комплекса радиопеленгации с использованием методов виртуальных магнитных диполей в условиях искажений, вносимых корпусом-носителем. В качестве примера рассматриваются случаи пеленгации разнополяризованных волн, учета влияния обтекателей антенных элементов, оценка предельной точности разрешения, а также шумовой устойчивости. При моделировании использовался метод конечных элементов, реализованный в DS CST Studio Suite 2024.

Результаты. Во время исследования на БПЛА с комплексом пеленгации падала плоская волна с разными пеленгами, что позволило проводить наиболее точное исследование. К результатам работы также следует отнести создание модели комплекса радиопеленгации, который может быть установлен на малом БПЛА, обеспечивающего высокую точность пеленгации волн в пассивном режиме.

Новизна применяемого метода заключается в формировании виртуальных магнитных диполей на основе измеренных характеристик искаженного электрического поля. Также новизной данной работы являются случаи моделирования, которые связаны с шумовым подавлением, влиянием корпусов антенных элементов, а также оценки предельной разрешающей точности.

Практическая значимость заключается в создании модели комплекса радиопеленгации на основе векторных антенных элементов, которые применяются для измерения характеристик электромагнитного поля с последующей пеленгацией на основе магнитного поля. В процессе исследования были рассмотрены все аспекты и случаи влияния внешних факторов на точность радиопеленгационного комплекса.

Ключевые слова: пеленгация, виртуальная антенная решетка, векторная антенна, обтекатель антенны, шумовая устойчивость, применение БПЛА

Источник финансирования: Работа выполнена при финансовой поддержке Российского научного фонда по гранту № 19-79-10109, <https://rscf.ru/project/19-79-10109>

Ссылка для цитирования: Ищенко Е.А., Пастернак Ю.Г., Фёдоров С.М., Баранников И.А. Реализация высококомобильного комплекса пеленгации для БПЛА с применением виртуальных магнитных диполей // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 48–61. DOI:10.31854/1813-324X-2024-10-4-48-61. EDN:YZJDJC

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-48-61>

Implementation of a Highly Mobile UAVs Direction Finding Complex Using Virtual Magnetic Dipoles

Evgeniy Alekseevich Ishchenko✉, kursk1998@yandex.ru

Yuri Gennadievich Pasternak, pasternaky@mail.ru

Sergey Mikhailovich Fedorov, fedorov_sm@mail.ru

Ilya Andreevich Barannikov, 8thbar@gmail.com

Voronezh State Technical University,
Voronezh, 394006, Russian Federation

Annotation

Relevance. In the context of the development of modern broadband communication systems, the task of performing direction finding of signals using highly mobile direction finding systems (UAVs) is becoming especially urgent. They impose restrictions on the size of antenna elements and the distance between them, which leads to an increase in the mathematical error of the bearing.

Problem statement. The paper sets the task of considering the possibility of increasing the direction finding accuracy of a highly mobile complex by using virtual magnetic dipole technology. A feature of this method is the minimization of distortions introduced by the carrier body into the characteristics of the measured field. To measure the field characteristics, as well as their partial components, vector antenna elements were used.

Goal of the work is to study the characteristics of a radio direction finding complex using virtual magnetic dipole methods in conditions of distortions introduced by the carrier body. As an example, cases of direction finding of differently polarized waves, taking into account the influence of radomes of antenna elements, assessment of the maximum resolution accuracy, as well as noise stability are considered. The finite element **method** implemented in DS CST Studio Suite 2024 was used in the modeling.

Result. During the research, a plane wave with different bearings fell on a UAV with a direction finding complex, which made it possible to carry out the most accurate research. The results of the work also include the creation of a model of a radio direction finding complex, which can be installed on a small UAV, providing high accuracy of wave direction finding in passive mode.

The **novelty** of the method used lies in the application of methods for the formation of virtual magnetic dipoles based on the measured characteristics of the distorted electric field. New to this work are modeling cases that are associated with noise suppression, the influence of antenna element housings, as well as assessment of the maximum resolution accuracy.

Practical significance of the work lies in the creation of a model of a radio location system based on vector antenna elements, which are used to measure the characteristics of the electromagnetic field with subsequent direction finding based on the magnetic field.

Keywords: direction finding, virtual antenna array, vector antenna, antenna radome, noise immunity, UAV application

Funding: The work was supported by the Russian Science Foundation, contract number 19-79-10109, <https://rscf.ru/project/19-79-10109>

For citation: Ishchenko E.A., Pasternak Yu.G., Fedorov S.M., Barannikov I.A. Implementation of a Highly Mobile UAVs Direction Finding Complex Using Virtual Magnetic Dipoles. *Proceedings of Telecommunication Universities*. 2024;10(4):48–61. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-48-61. EDN:KIGLSY

Введение

Применение высокоомобильных пеленгационных комплексов является важной и актуальной

задачей. Особое внимание стоит уделить возможности установки комплексов пеленгации – их антенные блоки на БПЛА. Таким образом, возникает

ограничение, которое вызвано двумя факторами – антенные системы на БПЛА должны быть малыми по весу; число антенн, которые могут быть установлены на БПЛА, ограничено ввиду сохранения электромагнитной совместимости, а также уменьшения помех для объектов. Поэтому можно заметить, что оптимальным вариантом является ситуация, которая связана с применением векторных антенных элементов, которые объединяются в антенную решетку пассивного пеленгатора. Пример использования векторных антенных элементов в роли пеленгационных антенн рассматривается в работе [1]. Авторы отмечают преимущества такого решения: высокая точность измерения характеристик поля, парциальные измерения компонент, возможность применения с использованием эллиптически поляризованных и линейных волн. Однако недостатками данной работы является то, что модель пеленгатора описана только с точки зрения математических моделей, при этом не приводится самой конструкции антенны.

Другим примером реализации векторного сканирования пространства является работа [2]. В процессе исследования авторы предлагают использовать векторное сканирование пространства на основе многолучевой антенны с переключаемыми лучами. Пеленгация базируется на измерении векторных характеристик сигналов, которые возбуждаются от волн, падающих от источника радиоизлучения (ИРИ). Полученная методика позволяет находить пеленг с высокой точностью – 5°, однако выбранная конструкция антенны достаточно велика для БПЛА, а также требует использования дополнительных активных цепей, которые приводят к росту энергопотребления, что, в свою очередь, увеличивает расход аккумулятора, в результате – уменьшается дальность работы, и ограничивается возможность применения БПЛА.

В работе [3] авторы предлагают конструкцию векторных антенн на основе диполей. Преимуществами такого решения является возможность упрощения конструкции обтекателей антенны, а также снижение веса; однако выбранный диапазон частот – VHF – приводит к большим размерам антенны, что не позволяет ее использовать совместно с БПЛА (ширина плеча составляет 350 мм), при этом стоит учитывать, что для комплекса пеленгации требуется установка нескольких антенн, а это не позволяет применять предложенную конструкцию на БПЛА ввиду ее больших размеров. Другим примером конструкции векторных антенн является антенна, которая рассматривается в [4]. Предложенная антенна позволяет обеспечить быстрое измерение характеристик поля, при этом она имеет малые размеры, а также – возможность переключения между портами. Все указанные характе-

ристики позволяют сделать высокоэффективную пеленгационную векторную антенну, однако у приведенной антенны наблюдается нестабильность диаграммы направленности, что может вызывать искажения пеленгационной характеристики, а, следовательно, – дополнительные ошибки определения пеленга. Для решения проблемы искажения характеристик пеленга, которые могут быть вызваны корпусом-носителем, неоднородностью диаграмм направленности на разных частотах, а также шумового воздействия, применяются алгоритмы аппроксимации поля, оптимизации диаграмм направленности, формирования дополнительных отсчетов – все эти методы принято объединять в единую группу формирования виртуальных антенных решеток (ВАР). В работе [5] приводится пример формирования ВАР на основе цилиндрической оптимизации и аппроксимации поля, при этом измерения базовых характеристик реальной антенной решетки (РАР) выполняются на основе дипольных антенных элементов. Предложенная методика позволила значительно снизить боковые лепестки, а, следовательно, уменьшить уровень шумового воздействия на характеристики, измеряемые антенной. Другим примером построения ВАР является работа [6], как показывают полученные в ней результаты, использование технологии аппроксимации поля позволяет значительно повысить точность пеленгации, при этом обеспечить рост шумовой стабильности. При этом авторами показано, что для обеспечения оптимального функционирования ВАР требуется использовать определенное количество виртуальных антенных элементов, так как слишком большое их число, как и недостаточное, приводит к росту ошибок пеленга ввиду возникающих ошибок при решении систем уравнений (переопределенность системы). В работе [7] авторы рассматривают возможность модификации одного из наиболее распространенных методов формирования ВАР – MUSIC, ESPRIT [8], root-MUSIC [9]. Выбранный алгоритм прост в реализации и надежен, однако требует большой статистической выборки для точной настройки определения пеленга, что является его недостатком. Таким образом, отчетливо видно, что использование ВАР значительно улучшает точность пеленгации, однако ввиду малых размеров БПЛА необходимо использовать новые алгоритмы их формирования, вызванные ограничениями размеров РАР [10]. Для решения проблемы расположения применяется множество методов, например, использование роя БПЛА, каждый из которых может нести отдельную пеленгационную антенну [11–14], однако применение такого метода определения характеристик целей приводит к ошибкам, которые могут быть вызваны расположением, разностью высот, а также необходимостью применения нескольких БПЛА.

В данной работе предлагается методика реализации виртуальных магнитных диполей (ВМД) для пеленгационного комплекса, который применяется на БПЛА, использующий в роли пеленгаторных антенн векторные антенны. Предложенный метод позволяет обеспечить высокоэффективное сканирование пространства путем применения одного БПЛА, что повышает надежность и точность определения пеленга. В работе приводится анализ стабильности работы предложенного метода, а также его математическое описание с проведением электродинамического моделирования в программном обеспечении DS CST Studio Suite 2024.

Описание исследуемой задачи

В процессе исследования производилось моделирование комплекса пеленгации, который формировался из кубических векторных антенных элементов, конструкция которых схожа с работой [15]. Так, каждый векторный антенный элемент имел кубическую форму, а характеристики поля измерялись с использованием 12 симметричных вибраторов длиной по 40 мм (размер грани куба 50 мм, по 5 мм оставлено для изоляции проводников, а также возможности их крепления). Каждый антенный элемент был нагружен на эквивалентную схему малошумящего усилителя, которая представляла из себя последовательное соединение резистора в 50 Ом и емкости 1,5 пФ. Благодаря выбранной кубической форме антенный элемент способен выполнять парциальные измерения всех компонент электрического и магнитного поля по x , y , z , а также способен выполнять пеленгацию при всех типах поляризации электромагнитных волн. В роли БПЛА-носителя был выбран квадрокоптер, к которому крепились четыре векторных антенных элемента, что позволяет обеспечить повышение точности пеленгации электромагнитных волн, при этом обеспечить высокую аэродинамическую эффективность и малый вес комплекса пеленгации. На рисунке 1 приводится вид векторного антенного элемента, а также комплекс пеленгации из четырех векторных антенных элементов, установленный на БПЛА.

В процессе исследования у БПЛА были указаны материалы изготовления, чтобы учитывать возможные потери и искажения характеристик поля, вызванные корпусом. Верх корпуса был выполнен из алюминия, низ и винты – из поликарбоната, двигатели – из меди и алюминия. Как видно, корпус БПЛА имеет сложную геометрическую форму, поэтому стоит ожидать, что из-за переотражений волн будут возникать искажения характеристик электрического поля, так как комплекс пеленгации и антенные элементы выступают рассеивателями электрического типа.

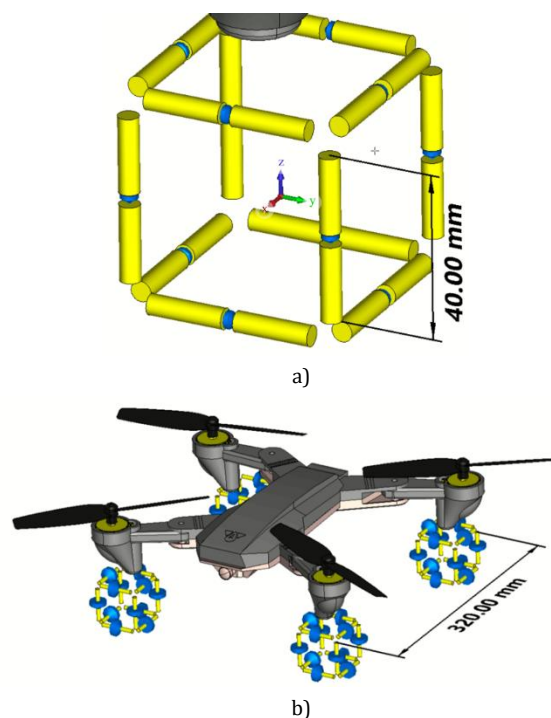


Рис. 1. Исследуемый комплекс пеленгации на основе векторных антенн: а) векторный антенный элемент; б) комплекс пеленгации, установленный на БПЛА

Fig. 1. The Investigated Direction-Finding Complex Based on Vector Antennas: a) Vector Antenna Element; b) Direction Finding Complex Installed on a UAV

Для подтверждения данной теории можно провести электродинамическое моделирование полной конструкции, при этом на окружности вокруг модели – установить пробники электрического и магнитного полей. В результате чего можно будет выполнить измерения и определить характеристическое волновое сопротивление, которое имеет вид:

$$Z = \frac{E}{H}, \quad (1)$$

где Z – волновое сопротивление (Ом); E – напряженность электрического поля (В/м); H – напряженность магнитного поля (А/м).

Известно, что волновое сопротивление свободного пространства является постоянной величиной и составляет 120π или примерно 377 Ом. Именно от величины волнового сопротивления и будет зависеть тип рассеяния волн корпусом-носителем БПЛА, так если значение волнового сопротивления в ближней зоне превышает 120π, то корпус является рассеивателем электрического типа, а следовательно, будут возникать большие искажения электрических компонент поля, что способно вызвать большие ошибки определения пеленга электромагнитных волн. Для определения типа рассеивателя корпуса-носителя выполнялось электродинамическое моделирование исследуемой задачи, так, вокруг БПЛА на уровне расположения антенн формировались пробники поля, которые измеряли E - и

H -поля, после чего определялось волновое сопротивление, причем расчет выполнялся на разных расстояниях от БПЛА – радиус пробников 325, 500, 650, 1000, 1300, 1500 мм. На рисунке 2 приводится вид БПЛА с установленными вокруг пробниками, а также значения волнового сопротивления.

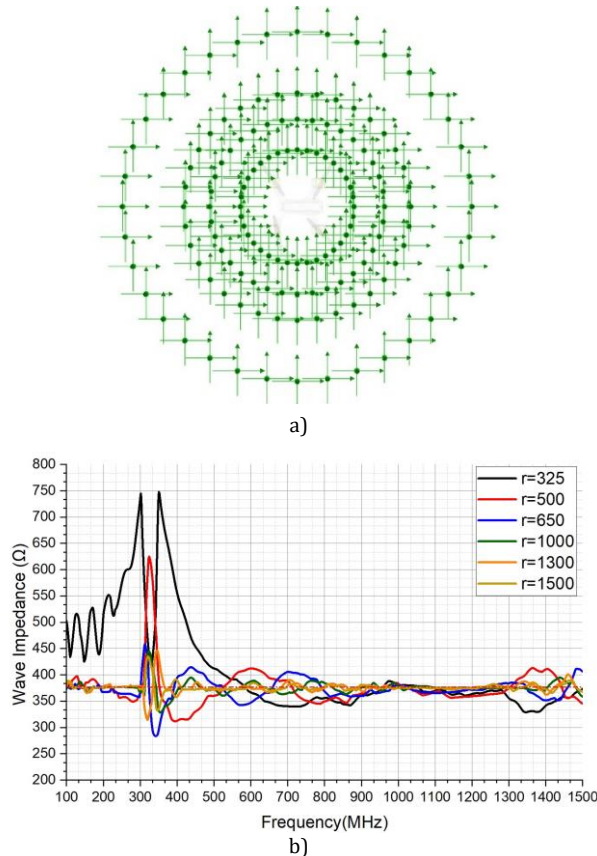


Рис. 2. Оценка характеристик отношения компонент векторов электрического поля к магнитному в ближней зоне рассеяния волн: а) расположение пробников поля в ближней зоне; б) характеристическое волновое сопротивление на разном расстоянии от БПЛА

Fig. 2. Evaluation of the Characteristics of the Ratio of the Components of the Electric Field Vectors to the Magnetic Field in the Near Wave Scattering Zone: a) Location of Field Probes in the Near Zone; b) Characteristic Wave Impedance at Different Distances from the UAV

Как показывают полученные результаты, в ближней зоне корпуса БПЛА отчетливо наблюдается рост волнового сопротивления, что приводит к значительному искажению электрического поля ввиду его большего рассеяния. График волнового сопротивления в ближней зоне можно разделить на ряд основных участков: до 400 МГц – максимально нестабильный резонансный участок, когда наблюдаются перепады сопротивлений, что вызывает серьезные искажения полей; диапазон от 400 до 600 МГц – преобладание электрического типа рассеяния волн, следовательно – большие искажения E -компонент электромагнитного поля; от 600 до 1500 МГц – стабилизация волнового сопротивления, которое становится максимально близким к 120π . Таким образом, появляется возможность вы-

двинуть гипотезу, что пеленгация на основе электрических компонент поля будет приводить к большим ошибкам, однако их измерение намного проще магнитных. Рассмотрим методику формирования ВАР на основе ВМД.

Математическое описание формирования виртуальных магнитных диполей

С учетом полученных картин волновых сопротивлений можно сделать вывод, что при использовании для пеленгования электрического поля стоит ожидать грубых ошибок, которые вызваны его излишним зашумлением и искажением. Для решения данной проблемы предлагается применять магнитное поле, которое будут измерять ВМД, которые, в свою очередь, будут определять характеристики магнитного поля путем восстановления их из электрического, так как его проще измерить.

Известно, что электрическое и магнитное поле связаны между собой на основе уравнений Максвелла; так, на основании уравнений в интегральной форме:

$$\oint_L \vec{E} \cdot d\vec{l} = -\frac{d}{dt} \int_S \vec{B} \cdot d\vec{s}, \quad (2)$$

где $d\vec{l}$ – единичный векторный элемент контура интегрирования L , который ориентирован против часовой стрелки; $d\vec{s}$ – единичный элемент площади, ориентированный по нормали к поверхности S , опирающейся на контур L ; $\vec{B}$ – вектор магнитной индукции, пронизывающей поверхность S .

Путем интегрирования и парциальных преобразований вид магнитного поля на основе напряжений, возбужденных в нагрузках, будет иметь вид:

$$H = -\frac{\sum_{n=1}^4 u_n^{(+,-)}}{i\omega\mu_a a}, \quad (3)$$

где $u_n^{(+,-)}$ – напряжения в нагрузках, измеренные векторным антенным элементом; a – длина диполя; μ_a – абсолютная магнитная проницаемость.

Благодаря использованию векторных антенных элементов появляется возможность измерять парциальные измерения компонент поля, а также осуществлять пеленгацию по ним. Для определения пеленга предлагается использовать корреляционно-интерферометрический метод, который базируется на анализе диаграммы направленности пеленгатора (пеленгационной характеристики) с определением ее максимума.

Пеленгационная характеристика имеет вид:

$$D = \left| \sum_{n=1}^N F_{p_n} \cdot \exp(-jk_0(x_n \cos(\psi) + y_n \sin(\psi))) \right|, \quad (4)$$

где F_{p_n} – поле, на основе которого выполняется пеленгация; p – парциальная компонента (x, y, z); k_0 – волновое число; x_n, y_n – координаты антенных элементов.

Благодаря выбранному методу появляются следующие возможности:

- пеленгация по E - или H -компонентам поля, соответственно используя электрические диполи, или ВМД;
- парциальная пеленгация – как по одной компоненте x, y, z , так и по парным или всем суммарным – наиболее точный способ, так как учитывает все измеренные характеристики;
- управление количеством антенных элементов для выполнения пеленгации (увеличение или уменьшение).

На основе разработанной математической модели проводились дальнейшие исследования эффективности работы разработанной методики повышения точности пеленгации на основе ВМД.

Оценка эффективности пеленгования с учетом разнополяризованных волн

Применение технологии ВМД напрямую связано с пеленгацией разнополяризованных волн. С точки зрения законов электродинамики существуют волны с эллиптической и линейной поляризацией. При этом важно отслеживать возможность верного определения пеленгов во всех случаях. При этом стоит заранее отметить, что не во всех типах волн присутствуют каждая парциальная компонента. Например, в линейно поляризованной волне отсутствует z -компонента поля, что не позволит выполнить верную пеленгацию; если же рассматривать эллиптически поляризованную волну, то оно может выполнять пеленгацию как на основе суммарных полей, так и по z -компоненте, так как она наименее искажена. Это особенно видно по диаграмме обратного рассеяния при падении волны на БПЛА с комплексом пеленгации (рисунок 3).

В процессе определения пеленга также важно рассчитывать статистические показатели пеленгации во всем диапазоне частот. На практике ИРИ передают сигналы в определенной полосе, что позволяет определить его пеленг путем расчета математического ожидания в сканируемом диапазоне частот и даст возможность повысить точность пеленгации.

С точки зрения статистической теории основные показатели имеют следующий вид:

- математическое ожидание:

$$m_1 = \frac{\sum_{n=1}^N x_n}{N}, \quad (5)$$

- среднеквадратическое отклонение пеленга:

$$\sigma = \sqrt{\frac{\sum_{n=1}^N (x_n - tg)^2}{N}}, \quad (6)$$

- абсолютная ошибка пеленга:

$$err = \frac{\sum_{n=1}^N |x_n - tg|}{N}, \quad (7)$$

где m_1 – математическое ожидание пеленга; N – число отсчетов пеленга в диапазоне расчетов; x_n – мгновенное значение пеленга (градусов); tg – ожидаемое значение пеленга; σ – среднеквадратическое отклонение пеленга в диапазоне (градус); err – абсолютная ошибка пеленга (градусы).

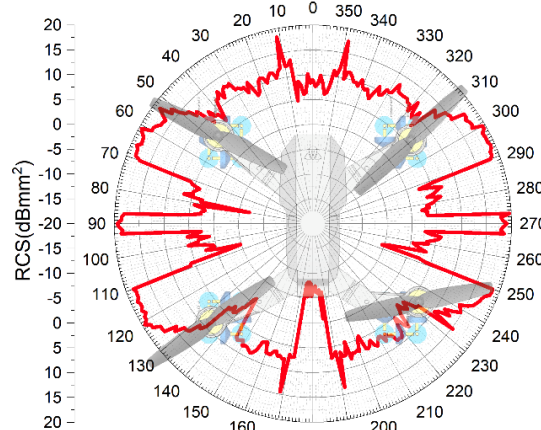


Рис. 3. Диаграмма обратного рассеяния БПЛА с комплексом пеленгации при облучении эллиптически поляризованной волны на частоте 500 МГц

Fig. 3. Radar Cross Section of a UAV with a Direction-Finding Complex When Irradiated by an Elliptically Polarized Wave at a Frequency of 500 MHz

Как уже было описано выше, в процессе моделирования на комплекс пеленгации падали электромагнитные волны с эллиптической (осевое отношение 10 к 1) и линейной поляризацией, после чего определялся пеленг электромагнитной волны. Так, в процессе пеленгации рассматривалась пеленгация по суммарным компонентам поля и по z -компонентам. Однако вторая ситуация может быть эффективно применена только в случае априорного наличия этой компоненты в волне, иначе следует использовать только суммарные компоненты электромагнитного поля. Полученные характеристики поля были использованы для расчета пеленга с использованием корреляционно-интерферометрического метода пеленгации, когда анализируется и обнаруживается пик пеленгационной характеристики. На рисунке 4 приводится вид пеленгационных характеристик, которые строятся с использованием суммарных компонент E - и H -поля при углах падения эллиптически поляризованной волны в 20 и 120 градусов на частоте 450 МГц, на которой как раз и стоит ожидать наибольших искажений характеристик волн.

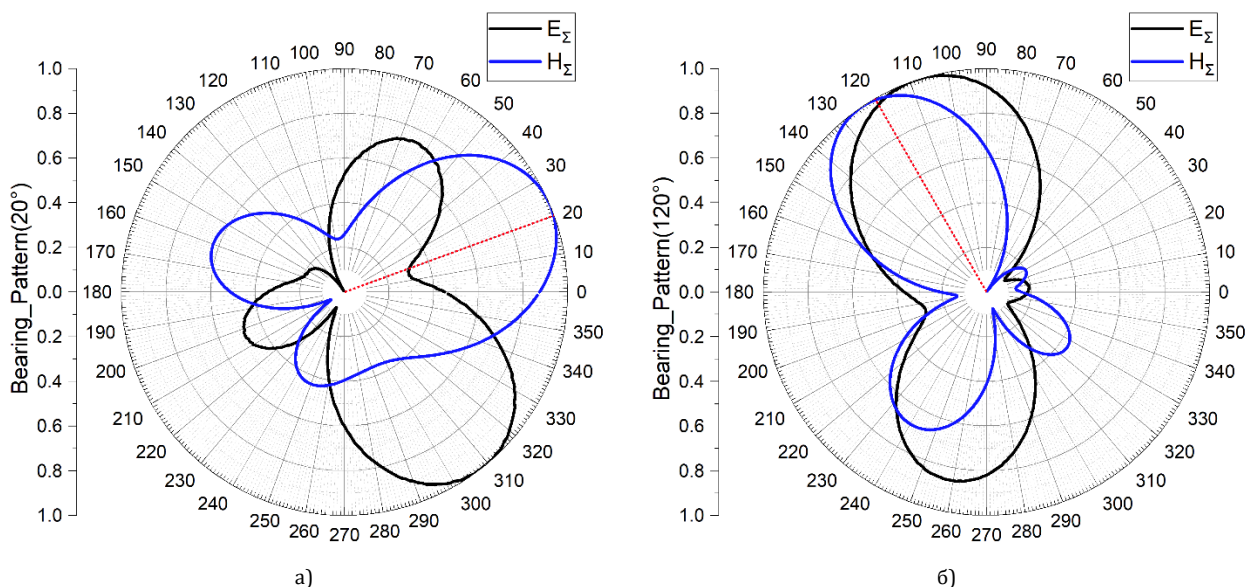


Рис. 4. Пеленгационная характеристика при пеленгации с использованием электрического и магнитного поля по суммарным компонентам: а) ИРИ = 20 °; б) ИРИ = 120 °

Fig. 4. Direction-Finding Characteristics During Direction-Finding Using Electric and Magnetic Fields Based on Total Components: а) True Bearing 20 °; б) True Bearing 120 °

Полученные картины отчетливо показывают, что при пеленгации с использованием ВМД удается достичь высокой точности определения пеленга электромагнитной волны, так как пик пеленгационной характеристики максимально близок к истинному углу падения волны. На рисунке 5 приводится зависимость пеленга электромагнитной волны от частоты при пеленгации по суммарным компонентам поля при облучении волнами с эллиптической и линейной поляризацией, при этом важно отметить, что в последующих исследовани-

ях также будет рассматриваться и пеленгация по z-компоненте поля, так как для усложнения задачи использовалась эллиптически поляризованная волна, которая имеет все три компонента поля. Приведенные пеленгационные характеристики были построены для углов падения в 20 ° и 120 ° в диапазоне частот от 100 до 1000 МГц, а в таблице 1 (значения пеленга приводятся с шагом в 20 °) указаны статистические характеристики пеленгации для диапазона частот 400–800 МГц, так как в нем обеспечивается наиболее стабильный пеленг.

ТАБЛИЦА 1. Характеристики пеленгации разнополяризованных волн

TABLE. 1. Characteristics of Direction Finding of Differently Polarized Waves

Пеленг, °	Линейная поляризация				Эллиптическая поляризация			
	E_z		H_z		E_z		H_z	
	m_1	σ	m_1	σ	m_1	σ	m_1	σ
0	2,90	3,10	2,40	2,30	333,65	345,13	1,38	1,59
20	31,86	10,91	15,55	4,48	312,62	292,63	23,55	3,61
40	73,02	35,04	40,00	0,00	245,94	218,38	41,92	2,55
60	59,63	15,35	62,29	5,73	113,32	80,67	60,85	3,26
80	108,51	75,86	79,63	1,37	127,39	99,43	86,29	6,48
100	120,65	56,04	98,47	1,64	101,91	29,97	109,91	10,24
120	131,76	34,62	117,65	2,48	129,56	43,98	124,88	4,94
140	140,42	0,67	137,00	3,00	130,7	17,35	142,06	2,91
160	156,29	20,99	156,91	3,18	163,32	3,45	165,65	6,18
180	180,00	0,00	179,56	0,69	185,87	5,9	176,32	3,91
Среднее	–	25,26	–	2,49	–	113,69	–	4,57

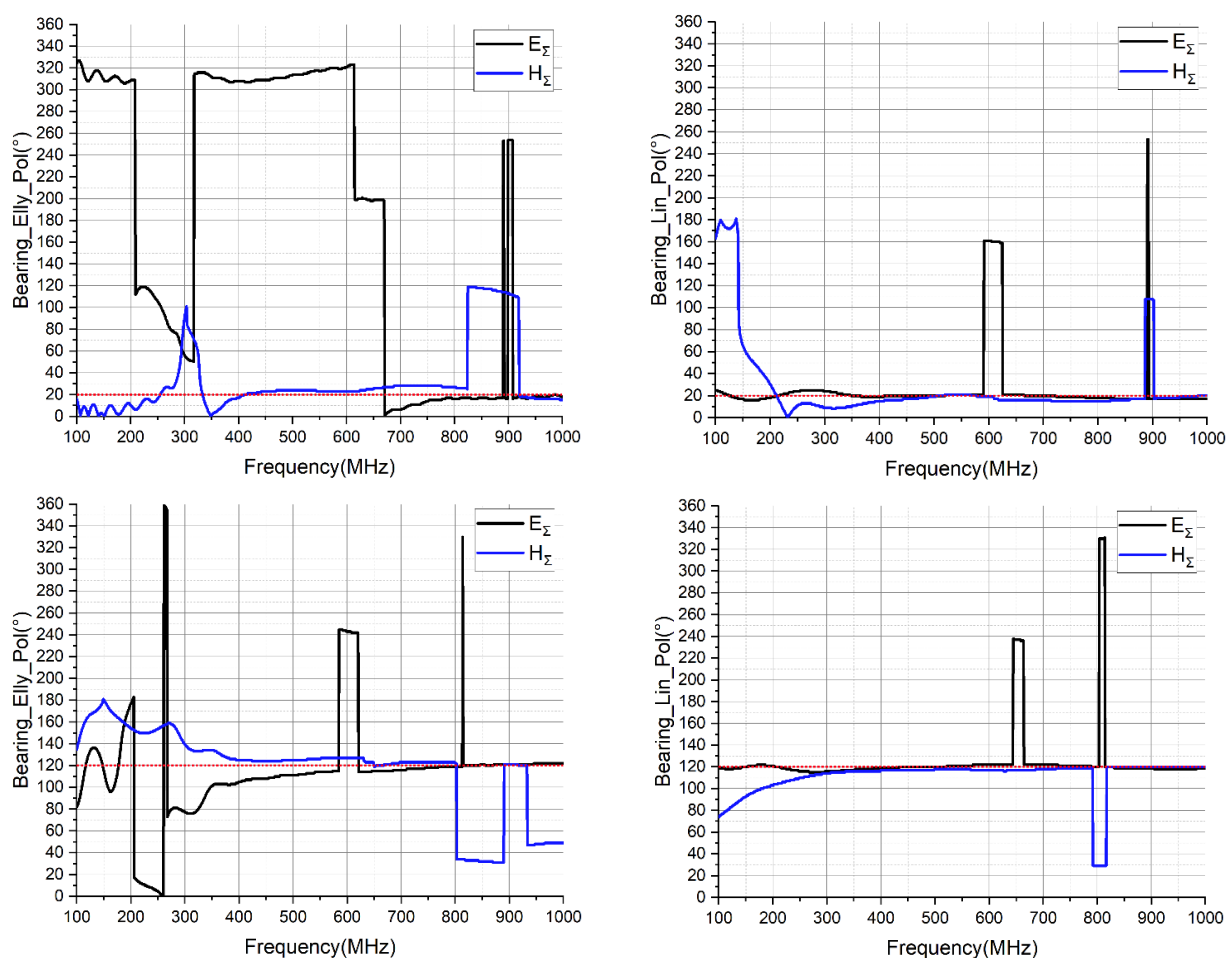


Рис. 5. Пеленгационные характеристики при эллиптической (слева) и линейной (справа) поляризации пеленгации разнополяризованных волн: истинный пеленг 20° (вверху) и 120° (внизу)

Fig. 5. Direction-Finding Characteristics for Elliptical (Left) and Linear (Right) Polarization of Direction-Finding of Differently Polarized Waves: True Bearing 20° (Top) and 120° (Bottom)

Полученные результаты показывают, что предлагаемый метод ВМД обеспечивает высокую точность пеленгации как при линейно, так и при эллиптически поляризованных волнах. Базовый алгоритм пеленгации на основе электрического поля и всех трех парциальных компонент обеспечивает допустимую точность только при облучении линейно (горизонтально) поляризованными волнами; вместе с тем при воздействии эллиптически поляризованных волн возникает грубая ошибка, которая не случается при применении пеленгации на основе ВМД. Для дальнейшей оценки эффективности предлагаемого метода рассмотрим его предельные разрешающие способности при воздействии эллиптически поляризованной волны и при углах пеленга от 0° до 5°.

Определение предельных разрешающих способностей при использовании предложенного метода

Как показали исследования, использование ВМД обеспечивает высокую точность пеленгации, особенно в случае пеленгации эллиптически по-

ляризованных волн. Для того, чтобы определить предельную чувствительность пеленгационного комплекса, на него подавались волны с углами падения от 0° до 5° с шагом в 1°, при этом волна имела эллиптическую поляризацию с осевым отношением 10 к 1. На основе этого выполнялся расчет пеленгов волны с применением суммарных компонент и отдельной z-компоненты, так как у эллиптически поляризованной волны она присутствует. На рисунке 6 приводится вид картин пеленгов при пеленгации с использованием E- и H-компонент поля, а в таблице 2 приводится вид статистических показателей пеленгации при использовании ВМД, так как E-компонентное пеленгование дает грубые промахи. Пеленг рассчитывался в наиболее стабильном диапазоне частот от 400 до 800 МГц. Полученные результаты показывают, что применение предложенного пеленгатора позволяет различать даже малые изменения угла падения волн, однако стоит учитывать ошибку, которая сохраняется ввиду неполного исправления ошибок измерения, которые вызваны искажением характеристик поля.

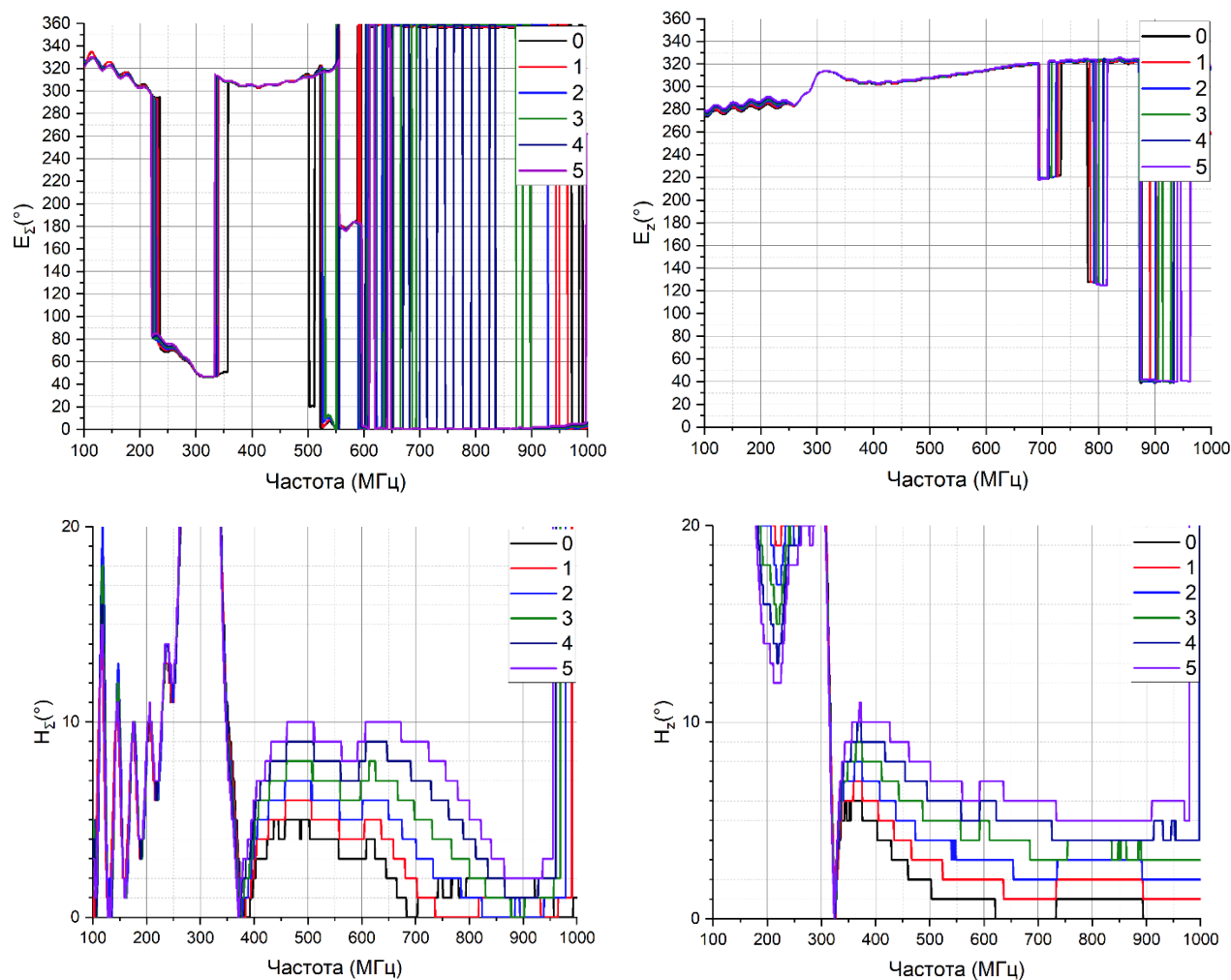


Рис. 6. Оценка предельного разрешения радиопеленгатора: пеленгация с использованием суммарного (слева) и z-компонент (справа) электрического (вверху) и магнитного (внизу) поля

Fig. 6. Estimating the Ultimate Resolution of a Direction Finder: Direction Finding Using the Sum (Left) and z-Components (Right) of the Electric (Top) and Magnetic (Bottom) Fields

ТАБЛИЦА 2. Статистические показатели предельного разрешения пеленгатора в диапазоне от 0° до 5°

TABLE 2. Statistical Indicators of the Maximum Resolution of the Direction Finder in the Range from 0° to 5°

Истинный пеленг, °	H_z			H_z		
	m_1	σ	err	m_1	σ	err
0	1,38	1,59	1,60	0,31	0,56	0,31
1	1,61	1,53	1,19	1,31	0,56	0,31
2	2,46	1,79	1,46	1,40	0,49	0,40
3	5,23	2,54	2,23	3,49	0,70	0,49
4	6,98	3,17	2,98	4,63	0,78	0,63
5	8,72	3,85	3,72	5,69	0,83	0,69
Среднее	–	2,41	2,20	–	0,65	0,47

Также при пеленгации эллиптически поляризованных волн отчетливо заметно, что z-компоненты поля наименее искажены, что приводит к минимальной ошибке (среднеквадратиче-

ское отклонение менее 1°). Таким образом, можно отметить высокую эффективность определения пеленга даже в случаях, когда требуется небольшая разрешающая способность радиопеленгационного комплекса. При этом можно отметить также тот момент, что применение пеленгации по z-компонентам поля расширяет рабочий диапазон частот до 950 МГц. Таким образом, отчетливо видно, что применение ВМД обеспечивает наибольшую эффективность пеленгации с использованием малого комплекса пеленгации, который установлен на БПЛА.

В рассмотренных случаях векторные антенные элементы помещались в свободном пространстве, но на практике такая реализация невозможна ввиду следующих причин:

– во-первых, не обеспечивается достаточная жесткость конструкции, что не позволит оптимально крепить антенные элементы;

– во-вторых, нарушается аэродинамический профиль БПЛА ввиду появления новых вихревых потоков, что уменьшит дальность полетов ввиду роста энергопотребления;

– в-третьих, не обеспечивается центровка векторных антенн, что может привести к возникновению дополнительных ошибок.

Для решения данной проблемы оптимально использовать обтекатель, в который помещается антенный элемент, однако изготовление обтекателя – достаточно трудоемкий процесс, так как требуется использование прозрачных материалов. В нашем случае для формирования неискаженного магнитного поля нас интересует магнитная проницаемость, которая в большинстве материалов равна одному. Рассмотрим случай, когда обтекатель будет выполнен из диэлектрически непрозрачного материала, что позволит повысить экономическую эффективность разработки корпуса устройства (например, путем применения аддитивных технологий).

Определение влияния корпуса обтекателя на точность пеленгации волн

В процессе исследования влияния обтекателя на характеристики радиопеленгационного комплекса рассматривались три случая: материал обтекателя радиопрозрачен ($\epsilon_r = 1$, $\mu_r = 1$); материал близок к полистиролу или большинству 3Д-печатных филаментов ($\epsilon_r = 2$, $\mu_r = 1$); материал с высоким уровнем диэлектрической проницаемости ($\epsilon_r = 5$, $\mu_r = 1$). Вид БПЛА с пеленгационными антеннами, помещенными в обтекатель, приводится на рисунке 7.

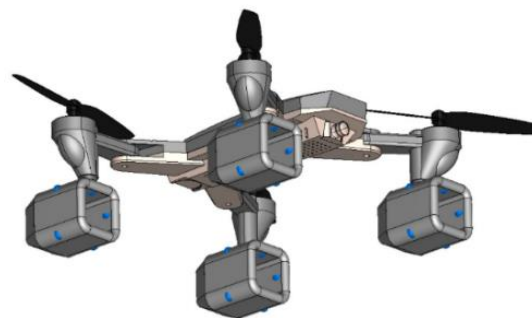


Рис. 7. Модель БПЛА с антеннами пеленгационного комплекса, помещенными в обтекатель

Fig. 7. Model of a UAV with Direction-Finding Antennas Placed in the Radome

Как видно по конструкции обтекателя, он представляет из себя аэродинамически профилированный корпус, который нацелен на стабилизацию воздушных потоков в нижней полусфере БПЛА. В процессе моделирования на БПЛА падала эллиптически поляризованная электромагнитная волна с осевым отношением 10 к 1. На рисунке 8 приводится вид пеленгационной характеристики для угла падения 60° , а в таблице 3 – статистические показатели пеленгации при углах падения от 0° до 180° с шагом в 20° .

Статистические показатели приводятся для суммарных компонент поля, так как пеленгация по ним возможная вне зависимости от типов падающих волн. Полученные результаты показывают, что применение диэлектрически непрозрачных обтекателей не приводит к росту ошибки пеленга, которая при использовании ВМД практически не изменяется.

ТАБЛИЦА 3. Статистические показатели пеленгации при использовании диэлектрически непрозрачного обтекателя

TABLE 3. Statistical Indicators of Direction Finding When Using a Dielectrically Opaque Radome

Пеленг, °	$\epsilon_r = 1$				$\epsilon_r = 2$				$\epsilon_r = 5$			
	m_1		σ		m_1		σ		m_1		σ	
	E_Σ	H_Σ	E_Σ	H_Σ	E_Σ	H_Σ	E_Σ	H_Σ	E_Σ	H_Σ	E_Σ	H_Σ
0	333	1,4	345	1,6	307	4,4	307	4,4	312	4,3	312	4,3
20	274	24	267	4,2	59	27	91	7,4	313	24	292	4,6
40	266	41	236	2,7	326	40	286	1,0	327	40	287	0,9
60	53	59	52	4,6	83	56	74,5	4,5	20	56	41	4,7
80	110	77	87	5,4	71	74	9,2	7,2	100	88	45	9,6
100	102	105	36	6,3	102	102	39	3,5	105	101	44	3,2
120	129	123	45	2,9	110	124	10	4,4	101	125	11	5,1
140	131	141	11	3,6	126	141	15	2,4	126	143	14	3,1
160	101	161	84	7,3	209	157	76	9,9	141	163	22	3,1
180	191	177	49	6,5	176	179	35	3,6	170	178	38	2,2
среднее	–	–	121,2	4,51	–	–	94,3	4,83	–	–	110,6	4,08

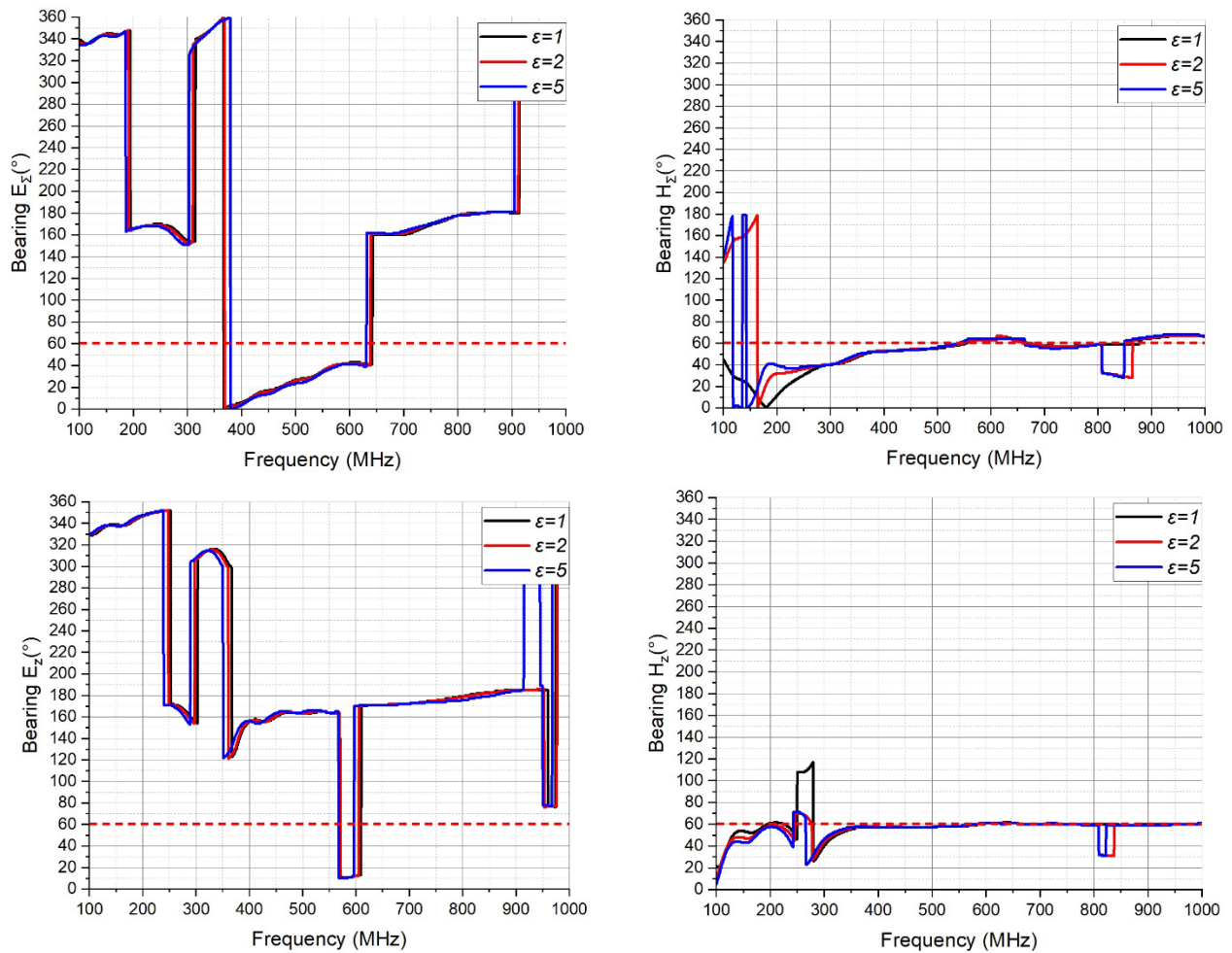


Рис. 8. Пеленгационные характеристики при установке векторных антенн в обтекатель: пеленгация на основе суммарного (вверху) электрического (слева) и z-компонент (внизу) магнитного (справа) поля

Fig. 8. Direction Finding Characteristics When Installing Vector Antennas in a Radome: Direction Finding Based on the Total (Top) Electric (Left) and z-Component (Bottom) Magnetic (Right) Field

Благодаря полученным результатам появляется возможность использовать аддитивные технологии для изготовления обтекателей, что значительно упрощает процесс производства и снижает стоимость радиопеленгационного комплекса. Таким образом, можно отметить, что применение предлагаемого алгоритма формирования ВМД обеспечивает высокую точность пеленгации при разных типах поляризации волн, обладает высокой разрешающей способностью и чувствительностью к углам падения электромагнитных волн, а также сохраняет стабильность характеристик при установке непрозрачного обтекателя на антенный элемент. Однако помимо влияния типов поляризации, корпуса носителя и обтекателя, возможно возникновение шумового воздействия на измеряемые характеристики электромагнитных волн. Рассмотрим эффективность пеленгации в условиях шумового воздействия на пеленгационный комплекс.

Оценка эффективности работы метода ВМД в условиях шумового воздействия

Как было описано выше, пеленгационные комплексы подвержены множеству влияний окружающих объектов. Однако полученные результаты моделирования показывают, что использование ВМД позволяет минимизировать искажения характеристик поля, что позволяет значительно повысить точность пеленгации. Рассмотрим случай, когда на комплекс пеленгации падает эллиптическая волна, которая возбуждает в нагрузках зашумленные напряжения. Для учета шумового влияния использовался комплексный аддитивный Гауссовский шум вида:

$$Noise = u_{ReGN} + i \cdot u_{ImGN} err = \frac{\sum_{n=1}^N |x_n - tg|}{N}. \quad (8)$$

В процессе исследования шумовой устойчивости радиопеленгатора производилось обучение БПЛА эллиптически поляризованной электромаг-

нитной волной, после чего рассчитывались пеленги сигналов при разных отношениях сигнал/шум (ОСШ) в диапазоне частот от 400 до 800 МГц. Для оценки устойчивости пеленга были построены зависимости среднеквадратического отклонения пеленга от ОСШ для пеленгации на основе ВМД (так как при пеленгации на основе электрического

поля априори возникают грубые ошибки пеленга). На рисунке 9 приводятся зависимости среднеквадратического отклонения от уровня ОСШ для углов пеленгации в 20°, 60°, 120°, 160° на основе суммарного магнитного поля и его z-компонент (так как волна эллиптически поляризована, то эти компоненты присутствуют).

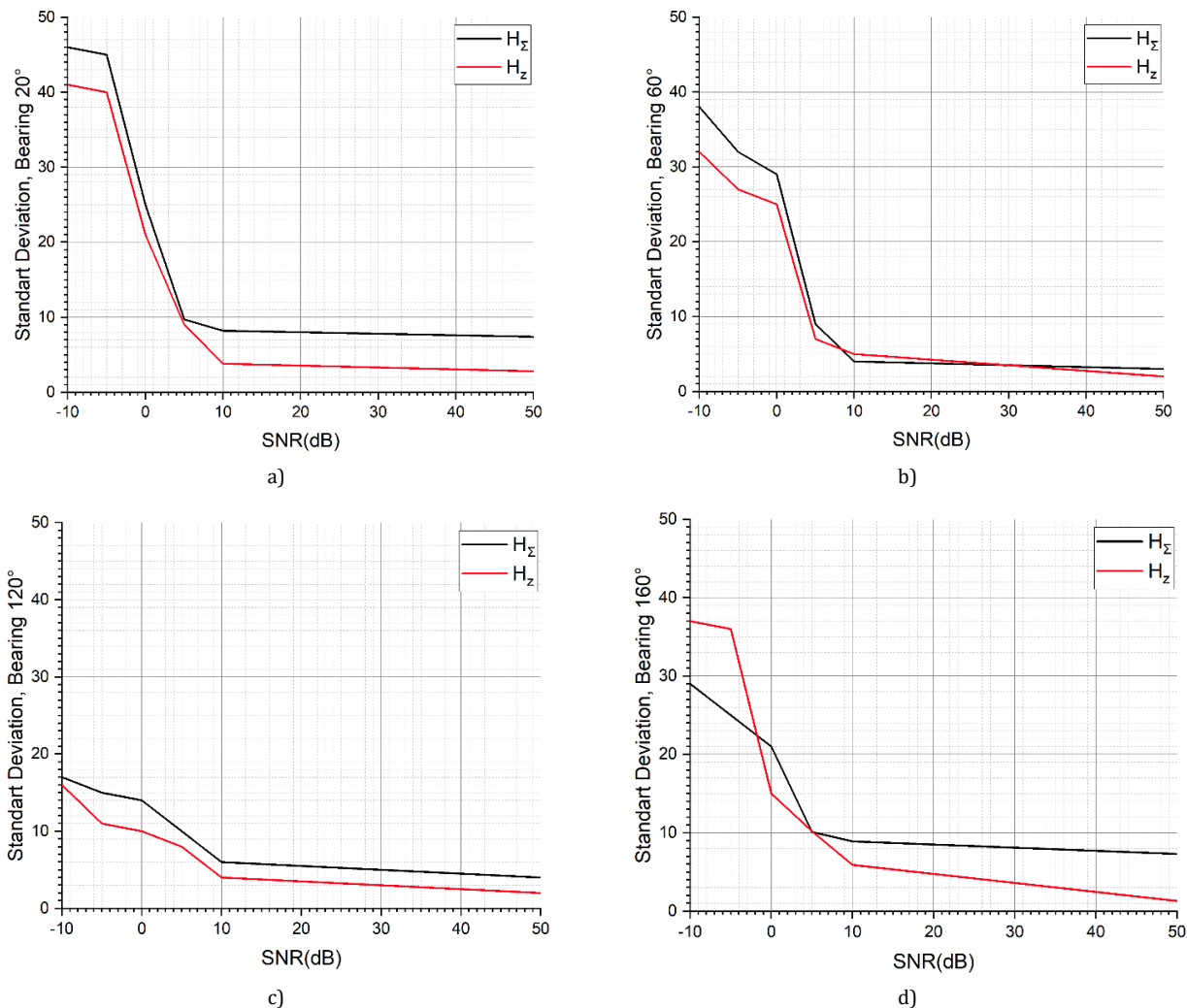


Рис. 9. Зависимость уровня среднеквадратического отклонения от уровня ОСШ при различных значениях пеленга: а) 20°; б) 60°; в) 120°; г) 160°

Fig. 9. Dependence of the Standard Deviation Level on the SNR Level: a) Bearing 20°; b) Bearing 60°; c) Bearing 120°; d) Bearing 160°

Полученные результаты показывают, что применение метода ВМД обеспечивает высокую эффективность пеленгации: до уровня ОСШ в 5 дБ. При таком ОСШ среднеквадратическое отклонение пеленга в диапазоне от 400 до 800 МГц приближается к 10°, что уже можно считать порогом ошибки пеленгации, при этом все углы прихода волн сохраняют данный порог, что свидетельствует о стабильности работы предложенного метода; при этом пеленгация по z-компоненте поля обеспечивает меньшую ошибку определения пеленга благодаря большей помехозащищенности вертикальной компоненты магнитного поля.

Заключение

В работе была предложена методика повышения точности радиопеленгации за счет использования ВМД. Пеленгационный комплекс был установлен на малом высокомобильном БПЛА, что позволяет расширить скорость и область применения комплекса пеленгации. Были исследованы все основные характеристики предлагаемого комплекса радиопеленгации. Так, было получено, что предлагаемая методика формирования ВМД обеспечивает высокую точность пеленгации волн с разной поляризацией, обеспечивает высокую точность и чувствительность при малых изменениях угла, сохра-

няет высокую точность при размещении антенных элементов внутри обтекателя, а также в условиях шумового воздействия. Было показано, что корпус БПЛА является рассеивателем электрического типа в ближней зоне, что вызывает большие искажения

характеристик электрического поля, однако измерение электрического поля намного проще, поэтому был предложен метод преобразования характеристик измеренного электрического поля в магнитное путем формирования ВМД.

Список источников

1. Sun B., Tan B., Ashraf M., Valkama M., Lohan E.S. Embedding the Localization and Imaging Functions in Mobile Systems: An Airport Surveillance Use Case // *IEEE Open Journal of the Communications Society*. 2022. Vol. 3. PP. 1656–1671. DOI:10.1109/OJCOMS.2022.3208945
2. Tarkowski M., Kulas L. RSS-Based DoA Estimation for ESPAR Antennas Using Support Vector Machine // *IEEE Antennas and Wireless Propagation Letters*. 2019. Vol. 18. Iss. 4. PP. 561–565. DOI:10.1109/LAWP.2019.2891021
3. Musicant A., Almog B., Oxenfeld N., Shavit R. Vector Sensor Antenna Design for VHF Band // *IEEE Antennas and Wireless Propagation Letters*. 2015. Vol. 14. PP. 1404–1407. DOI:10.1109/LAWP.2015.2409315
4. Duploux J., Morlaas C., Aubert H., Potier P., Pouliguen P., Djoma C. Reconfigurable Grounded Vector Antenna for 3-D Electromagnetic Direction-Finding Applications // *IEEE Antennas and Wireless Propagation Letters*. 2018. Vol. 17. Iss. 2. PP. 197–200. DOI:10.1109/LAWP.2017.2779878
5. Dawood H.S., El-Khobby H.A., Elnaby M.M.A., Hussein A.H. Optimized VAA Based Synthesis of Elliptical Cylindrical Antenna Array for SLL Reduction and Beam Thinning Using Minimum Number of Elements // *IEEE Access*. 2021. Vol. 9. PP. 50949–50960. DOI:10.1109/ACCESS.2021.3069795
6. Yahia A.A., Elkamchouchi H.M. Design of Virtual Antenna Array for Direction of Arrival Estimation Using Real Antenna Array System // *Proceedings of the 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC, Jeju, South Korea, 23–26 June 2019)*. IEEE, 2019. DOI:10.1109/ITC-CSCC.2019.8793364
7. Zuo M., Xie S., Li Y., Zhang C. Joint Estimation Method for Frequency and DOA of Virtual Antenna Array in Space-Time Domain // *Proceedings of the International Applied Computational Electromagnetics Society Symposium – China (ACES, Nanjing, China, 08–11 August 2019)*. IEEE, 2019. DOI:10.23919/ACES48530.2019.9060452
8. Fayad Y., Wang C., Cao Q., Hafez A.E.-D. Improved ESPRIT algorithm used in spatial subspace for NULA // *Proceedings of the 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST, Islamabad, Pakistan, 12–16 January 2016)*. IEEE, 2016. PP. 620–623. DOI:10.1109/IBCAST.2016.7429943
9. Ateşavcı C.S., Bahadırlar Y., Aldırmaz-Çolak S. DoA Estimation in the Presence of Mutual Coupling Using Root-MUSIC Algorithm // *Proceedings of the 8th International Conference on Electrical and Electronics Engineering (ICEEE, Antalya, Turkey, 09–11 April 2021)*. IEEE, 2021. PP. 292–298. DOI:10.1109/ICEEE52452.2021.9415938
10. Lim J.-s., Song J., Sung K.-M. Forward-backward time varying forgetting factor Kalman filter based DOA estimation algorithm for UAV (Unmanned Aerial Vehicle) autoland // *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 2002. PP. IV-3964–IV-3967. DOI:10.1109/ICASSP.2002.5745525
11. Keskin F., Filik T. Isotropic and Directional DOA Estimation of the Target by UAV Swarm-based 3-D Antenna Array // *Proceedings of the 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT, Istanbul, Turkey, 22–24 October 2020)*. IEEE, 2020. DOI:10.1109/ISMSIT50672.2020.9254341
12. Cheng J., Guan K., Qutin F. A High-Accuracy DOA-Based Localization Method: UAV Virtual Multiantenna Array // *Proceedings of the Seventeenth International Conference on Wireless and Mobile Communications (ICWMC 2021)*. 2021. PP. 10–15.
13. Abeywickrama S., Jayasinghe L., Fu H., Nissanka S., Yuen C. RF-based Direction Finding of UAVs Using DNN // *Proceedings of the IEEE International Conference on Communication Systems (ICCS, Chengdu, China, 19–21 December 2018)*. IEEE, 2018. PP. 157–161. DOI:10.1109/ICCS.2018.8689177
14. Павлов И.Д. Исследование влияния антенных элементов миллиметрового диапазона на ошибки пеленгации фазовым методом // *Известия высших учебных заведений России. Радиоэлектроника*. 2023. Т. 26. № 2. С. 25–36. DOI:10.32603/1993-8985-2023-26-2-25-36. EDN:FRDWEB
15. Ashikhmin A.V., Ishchenko E.A., Pasternak Y.G., Pershin P.V., Sivash M.A., Fedorov S.M. Design of Virtual Magnetic Dipole Antenna Array to Reduce the Systematic Bearing Error Caused by Wave Diffraction on the Antenna System and its Carrier Body // *Proceedings of the 7th All-Russian Microwave Conference (RMC, Moscow, Russia, 25–27 November 2020)*. IEEE, 2020. PP. 111–114. DOI:10.1109/RMC50626.2020.9312312

References

1. Sun B., Tan B., Ashraf M., Valkama M., Lohan E.S. Embedding the Localization and Imaging Functions in Mobile Systems: An Airport Surveillance Use Case. *IEEE Open Journal of the Communications Society*. 2022;3:1656–1671. DOI:10.1109/OJCOMS.2022.3208945
2. Tarkowski M., Kulas L. RSS-Based DoA Estimation for ESPAR Antennas Using Support Vector Machine. *IEEE Antennas and Wireless Propagation Letters*. 2019;18(4):561–565. DOI:10.1109/LAWP.2019.2891021
3. Musicant A., Almog B., Oxenfeld N., Shavit R. Vector Sensor Antenna Design for VHF Band. *IEEE Antennas and Wireless Propagation Letters*. 2015;14:1404–1407. DOI:10.1109/LAWP.2015.2409315
4. Duploux J., Morlaas C., Aubert H., Potier P., Pouliguen P., Djoma C. Reconfigurable Grounded Vector Antenna for 3-D Electromagnetic Direction-Finding Applications // *IEEE Antennas and Wireless Propagation Letters*. 2018;17(2):197–200. DOI:10.1109/LAWP.2017.2779878

5. Dawood H.S., El-Khobby H.A., Elnaby M.M.A., Hussein A.H. Optimized VAA Based Synthesis of Elliptical Cylindrical Antenna Array for SLL Reduction and Beam Thinning Using Minimum Number of Elements. *IEEE Access*. 2021;9: 50949–50960. DOI:10.1109/ACCESS.2021.3069795
6. Yahia A.A., Elkamchouchi H.M. Design of Virtual Antenna Array for Direction of Arrival Estimation Using Real Antenna Array System. *Proceedings of the 34th International Technical Conference on Circuits/Systems, Computers and Communications, ITC-CSCC, 23–26 June 2019, Jeju, South Korea*. IEEE; 2019. DOI:10.1109/ITC-CSCC.2019.8793364
7. Zuo M., Xie S., Li Y., Zhang C. Joint Estimation Method for Frequency and DOA of Virtual Antenna Array in Space-Time Domain. *Proceedings of the International Applied Computational Electromagnetics Society Symposium – China (ACES), 08–11 August 2019, Nanjing, China*. IEEE; 2019. DOI:10.23919/ACES48530.2019.9060452
8. Fayad Y., Wang C., Cao Q., Hafez A.E.-D. Improved ESPRIT algorithm used in spatial subspace for NULA. *Proceedings of the 13th International Bhurban Conference on Applied Sciences and Technology, IBCAST, 12–16 January 2016, Islamabad, Pakistan*. IEEE; 2016. p.620–623. DOI:10.1109/IBCAST.2016.7429943
9. Ateşavcı C.S., Bahadirlar Y., Aldırmaz-Çolak S. DoA Estimation in the Presence of Mutual Coupling Using Root-MUSIC Algorithm. *Proceedings of the 8th International Conference on Electrical and Electronics Engineering, ICEEE, 09–11 April 2021, Antalya, Turkey*. IEEE; 2021. p.292–298. DOI:10.1109/ICEEE52452.2021.9415938
10. Lim J.-s., Song J., Sung K.-M. Forward-backward time varying forgetting factor Kalman filter based DOA estimation algorithm for UAV (Unmanned Aerial Vehicle) autoland. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE; 2002. PP. IV-3964–IV-3967. DOI:10.1109/ICASSP.2002.5745525
11. Keskin F., Filik T. Isotropic and Directional DOA Estimation of the Target by UAV Swarm-based 3-D Antenna Array. *Proceedings of the 4th International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT, 22–24 October 2020, 22–24 October 2020, Istanbul, Turkey*. IEEE; 2020. DOI:10.1109/ISMSIT50672.2020.9254341
12. Cheng J., Guan K., Qutin F. A High-Accuracy DOA-Based Localization Method: UAV Virtual Multiantenna Array. *Proceedings of the Seventeenth International Conference on Wireless and Mobile Communications, ICWMC 2021*. 2021. p.10–15.
13. Abeywickrama S., Jayasinghe L., Fu H., Nissanka S., Yuen C. RF-based Direction Finding of UAVs Using DNN. *Proceedings of the IEEE International Conference on Communication Systems, ICCS, 19–21 December 2018, Chengdu, China*. IEEE; 2018. p.157–161. DOI:10.1109/ICCS.2018.8689177
14. Pavlov I.D. Research into the Influence of Millimeter-Wave Antennas on Direction Finding Errors by Phase Method. *Journal of the Russian Universities. Radioelectronics*. 2023;26(2):25–36. (In Russ.) DOI:10.32603/1993-8985-2023-26-2-25-36. EDN:FRDWEB
15. Ashikhmin A.V., Ishchenko E.A., Pasternak Y.G., Pershin P.V., Sivash M.A., Fedorov S.M. Design of Virtual Magnetic Dipole Antenna Array to Reduce the Systematic Bearing Error Caused by Wave Diffraction on the Antenna System and its Carrier Body. *Proceedings of the 7th All-Russian Microwave Conference, RMC, 25–27 November 2020, Moscow, Russia*. IEEE; 2020. p.111–114. DOI:10.1109/RMC50626.2020.9312312

Статья поступила в редакцию 25.05.2024; одобрена после рецензирования 10.06.2024; принята к публикации 13.06.2024.

The article was submitted 25.05.2024; approved after reviewing 10.06.2024; accepted for publication 13.06.2024.

Информация об авторах:

**Ищенко
Евгений Алексеевич**

аспирант кафедры радиоэлектронных устройств и систем Воронежского государственного технического университета
 <https://orcid.org/0000-0002-5270-0792>

**ПАСТЕРНАК
Юрий Геннадьевич**

доктор технических наук, профессор, профессор кафедры радиоэлектронных устройств и систем Воронежского государственного технического университета
 <https://orcid.org/0000-0002-2031-5531>

**ФЁДОРОВ
Сергей Михайлович**

кандидат технических наук, доцент, доцент кафедры радиоэлектронных устройств и систем Воронежского государственного технического университета
 <https://orcid.org/0000-0001-9027-6163>

**БАРАННИКОВ
Илья Андреевич**

аспирант кафедры радиоэлектронных устройств и систем Воронежского государственного технического университета
 <https://orcid.org/0000-0003-3093-0455>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

Научная статья

УДК 621.391

<https://doi.org/10.31854/1813-324X-2024-10-4-62-72>

Модель и методы маршрутизации трафика в сети связи с использованием БПЛА

Константин Алексеевич Кузнецов, kuznetsov.sut@gmail.com

Александр Иванович Парамонов, paramonov@sut.ru

Аммар Салех Али Мутханна✉, muthanna.asa@sut.ru

Андрей Евгеньевич Кучерявый, akouch@sut.ru

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

Аннотация

Актуальность. Развитие сетей 5G и последующих поколений сопровождается развитием новых услуг, в частности, услуг виртуальной, дополненной реальности, а также телеприсутствия и сетей радиодоступа. В частности, происходит повышение рабочих частот, что ставит дополнительные задачи по организации сети, способной обеспечить требования к качеству обслуживания трафика со стороны новых услуг и обеспечению доступности связи пользователям. Эти задачи могут решаться различными способами размещения точек доступа, в том числе с использованием БПЛА. Такой подход обеспечивает оперативность построения и гибкость структуры сети доступа, но также требует применения методов размещения точек доступа по отношению к пользователям и другим элементам сети связи.

Постановка задачи: разработка методов размещения маршрутизаторов в рое БПЛА и выбора маршрутов пропуска трафика при организации сети доступа, в целях повышения эффективности функционирования сети связи.

Цель работы: повышение эффективности построения сети доступа с использованием БПЛА за счет разработки методов кластеризации и распределения маршрутизаторов в рое БПЛА.

Используемые методы. Исследования проводились с использованием положений теории информации, математических методов оптимизации, методов теории графов и методов кластеризации. Численные результаты получены методом численного моделирования на языке python.

Результат. Разработанная модель и методы позволяют произвести распределение маршрутизаторов (точек доступа) сети, размещенных на БПЛА с учетом качества обслуживания и обеспечения построения связной mesh-сети и ее связи с сетью подвижной связи, которые могут быть использованы как в современных, так и перспективных сетях связи.

Новизна: разработан модельно-методический аппарат, позволяющий повысить эффективность построения сетей беспроводного доступа с применением БПЛА, в частности позволяющий выбирать позиции размещения маршрутизаторов в рое БПЛА и логическую структуру сети. Разработанный модельно-методический аппарат решает задачу маршрутизации трафика с учетом качества его обслуживания.

Практическая значимость: предложенные модель и методы могут быть использованы при организации обслуживания в сетях 5G и последующих поколений. В частности, они позволяют обеспечить доступность связи и оперативность организации сети в случаях недостаточного покрытия, а также в случаях выхода из строя отдельных элементов сети. Возможность выгрузки трафика в локальную сеть позволяет повысить качество обслуживания трафика в сети оператора связи.

Ключевые слова: сети 5G и последующих поколений, кластеризация, БПЛА, трафик, маршрутизация, сеть доступа, качество обслуживания

Ссылка для цитирования: 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

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-62-72>

Model and Methods of Traffic Routing in a Communication Network Using UAVs

 **Konstantin A. Kuznetsov**, kuznetsov.sut@gmail.com

 **Alexander I. Paramonov**, paramonov@sut.ru

 **Ammar S.A. Muthanna** , muthanna.asa@sut.ru

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

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

Annotation

Relevance. The development of 5G networks and subsequent generations is accompanied by the development of new services, in particular, virtual, augmented reality services, as well as telepresence, as well as radio access networks. In particular, there is an increase in operating frequencies, which poses additional challenges for organizing a network that can meet the requirements for the quality of traffic service from new services and ensure the availability of communication to users. These problems can be solved by various methods of placing access points, including using UAVs. This approach ensures the efficiency of construction and flexibility of the access network structure, but also requires the use of methods for placing access points in relation to users and other elements of the communication network.

Problem statement: development of methods for placing routers in a UAV swarm and selecting traffic routes when organizing an access network, in order to improve the efficiency of the communication network.

Purpose of the work: improving the efficiency of building an access network using UAVs through the development of clustering methods and distributing routers in a UAV swarm.

Methods. The studies were carried out using the provisions of information theory, mathematical optimization methods, graph theory methods and clustering methods. The numerical results were obtained using the numerical simulation method in Python.

Result. The developed model and methods allow for the distribution of network routers (access points) located on UAVs taking into account the quality of service and ensuring the construction of a connected mesh network and its connection with the mobile network, which can be used in both modern and future communication networks.

Novelty: a modeling and methodological apparatus has been developed that allows for increasing the efficiency of building wireless access networks using UAVs, in particular, allowing for selecting the placement positions of routers in a UAV swarm and the logical structure of the network. The developed modeling and methodological apparatus solves the problem of traffic routing taking into account the quality of its service.

Practical significance: the proposed model and methods can be used to organize service in 5G networks and subsequent generations. In particular, they allow for ensuring the availability of communication and the efficiency of network organization in cases of insufficient coverage, as well as in cases of failure of individual network elements. The ability to unload traffic to a local network allows for improving the quality of traffic service in the operator's network.

Keywords: 5G and next generation networks, clustering, UAVs, routing, access network, quality of service

For citation: Kuznetsov K.A., Paramonov A.I., Muthanna A.S.A., Kucheryavy 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

Введение

Развитие сетей связи направлено на повышение основных показателей качества, таких как достижимая скорость передачи данных, задержка до-

ставки, емкость сети [1–5]. Сети 5G и последующих поколений ориентированы на предоставление услуг, чувствительных к задержкам и скорости передачи данных. Рост проникновения услуг виртуальной и дополненной реальности, а в послед-

ствии и услуг телеприсутствия [6] приводит к необходимости обеспечения высокого качества обслуживания трафика в сети.

В частности, возникает необходимость передачи трафика голографических изображений [7], а также трафика, производимого костюмами телеприсутствия, обеспечивающими интерфейс взаимодействия с метавселенной [8]. Метавселенная направлена на создание виртуального мира для всех видов деятельности, включая образование, торговлю и игры, и считается следующим поколением Интернета. При поддержке AR/VR-приложений онлайн-пользователям предоставляются услуги, схожие с очными. Работа приложений метавселенной обеспечивается синхронизацией данных и широким покрытием беспроводной сети – две практические проблемы, которые необходимо решить, поскольку в сервисах телеприсутствия обычно задействованы носимые беспроводные устройства. Ввиду повышения частотного диапазона, используемого в сетях 5G и выделяемого для 6G, стремительно возрастает затухание сигнала. Таким образом, развертывание базовых станций в пригородах с низкой плотностью населения становится экономически неэффективно [9]. Беспилотные летательные аппараты (БПЛА) являются экономически более выгодным решением для организации покрытия сети и синхронизации данных метавселенных в пригородной зоне.

Применение костюмов телеприсутствия или подобных им устройств, обеспечивающих получение информации о динамике движения людей или иных объектов, существенно отличается от сбора информации с помощью изображений и в значительной степени дополняет «картину» состояния и изменения окружающего нас мира. Трафик, производимый упомянутыми услугами, должен обслуживаться с тем уровнем качества, который требуется для их реализации. Одним из методов решения данной задачи является применение выгрузки трафика или выделения дополнительных ресурсов для его обслуживания.

Повышение основных показателей качества достигается внедрением новых технических решений, а также новых моделей и методов организации сетей связи. Одним из «узких мест» для решения задачи повышения показателей качества являются сети доступа. В сетях подвижной связи они ограничены из-за конечного объема используемого радиочастотного ресурса. Наиболее эффективным методом является повышение достижимой скорости передачи данных. В ходе эволюции сетей подвижной связи можно наблюдать процесс повышения полосы частот, занимаемых каналом связи и повышения рабочих частот. Это естественный процесс, который происходит из-за необходимости повышения пропускной способности.

Согласно известному выражению К. Шеннона [10], достижимая скорость передачи определяется по выражению: $b = w \log_2(1 + SNR)$ бит/с, где w – полоса пропускания канала; SNR – отношение сигнал/шум. Несмотря на некоторое уменьшение SNR с увеличением полосы пропускания, это является наиболее «простым» методом увеличения скорости. Слово «простым» взято в кавычки потому, что увеличение полосы пропускания связано с повышением рабочей частоты и необходимостью решения многих задач, связанных с этим.

Однако каждое последующее поколение сетей связи предполагает повышение полосы частот и увеличение рабочих частот. Так, в сетях 6G уже предполагается использование субмиллиметрового диапазона и рабочих частот вплоть до 1 ТГц. Особенности распространения сигналов в этом диапазоне волн таковы, что эффективно они могут быть использованы только на относительно малых расстояниях между приемником и передатчиком и в зоне прямой видимости между ними. Свойства сигналов в этом диапазоне волн приближаются к свойствам распространения света.

В настоящее время являются открытыми задачи создания электронных приборов, позволяющих реализовать устройства для генерации, модуляции, приема и обработки таких сигналов. Также открыты вопросы создания антенных устройств в этом диапазоне частот. Тем не менее, исследования [11, 12] позволяют ожидать, что такие элементы будут созданы и они позволят использовать особенности данного диапазона для повышения эффективности сети связи.

Впрочем, задачи, связанные с особенностями распространения сигналов в субмиллиметровом диапазоне волн, требуют решения на уровне организации сети доступа. В частности, «затенение» окружающими объектами, в том числе человеческим телом, может приводить к полной потере связи. Это требует использование большего числа базовых станций (антенн) и соответствующего управления ими.

Одним из подходов к решению задачи организации покрытия в сетях 5G и 6G может быть использование мобильных маршрутизаторов. Уже сейчас производители оборудования выпускают 5G-маршрутизаторы с автономным питанием [13], которые позволяют организовать локальную сеть стандарта IEEE 802.11. Основной задачей применения такого устройства является обеспечение достаточно высокой скорости передачи на уровне подключения к сети 5G.

При работе в перспективных сетях на частотах субмиллиметрового диапазона нужно обеспечить достаточно малое расстояние и прямую видимость между маршрутизатором и антенной базовой стан-

ции. Одним из способов достижения этого является применение БПЛА [14–16]. В данной работе предлагается метод применения БПЛА или группы БПЛА для размещения маршрутизаторов и организации обслуживания пользователей, находящихся в сложных условиях для приема сигналов сети.

Модель и постановка задачи

В общем случае, при размещении маршрутизатора на БПЛА необходимо решить задачу выбора точки его размещения. При этом возможны варианты, когда БПЛА является неподвижным, например, привязным [17], и когда БПЛА может перемещаться как относительно пользователей, так и относительно других БПЛА. Эти варианты отличаются характером движения БПЛА, если для первого варианта структуру сети перестраивать не требуется из-за изменения позиций узлов сети, то во втором варианте такая потребность может иметь место. Однако, и первый и второй варианты требуют решения двух основных задач:

- выбора позиций БПЛА относительно пользователей и базовых станций сети связи;
- выбора логической структуры сети, т. е. маршрутов пропуска трафика в mesh-сети.

Условная модель сети приведена на рисунке 1. На рисунке показаны одиночный БПЛА и группа (рой) БПЛА. Маршрутизатор, размещенный на одиночном БПЛА, выполняет роль точки доступа, обслуживая пользователей в своей зоне связи. Маршрутизаторы, размещенные на БПЛА роя, организованы в mesh (ячеистую) сеть. Один или несколько маршрутизаторов из этой группы могут взаимодействовать с базовыми станциями сети подвижной связи. Маршрутизаторы в рое, в зависимости от условий, могут выполнять функции связи с базовой станцией сети подвижной связи, с другими маршрутизаторами группы, с пользователями.

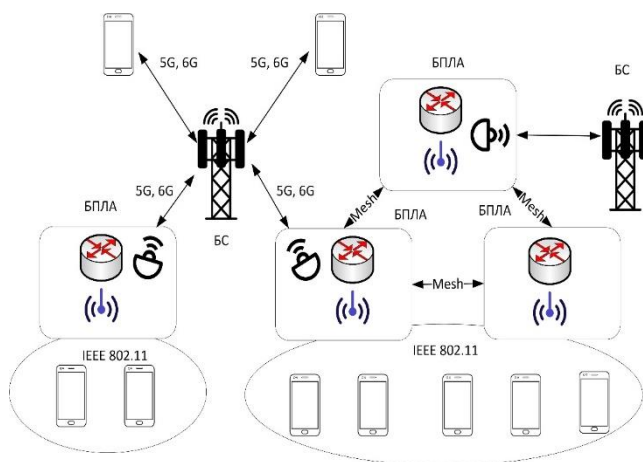


Рис. 1. Модель сети с использованием маршрутизаторов на БПЛА

Fig. 1. Network Model Using Routers on UAVs

Применение группы БПЛА позволяет покрыть связью пользователей, распределенных в некотором пространстве и организовать пропуск трафика между ними и сетью подвижной связи.

Выбор позиций маршрутизаторов

Рассмотрим вариант, когда имеется группа из n маршрутизаторов. В общем случае $n \in N$ и может быть задано или ограничено $n \leq n_{\max}$. Для выбора позиций размещения точек доступа чаще всего применяют различные методы кластеризации [18]. При этом объектами кластеризации являются пользователи, а за позицию размещения точек доступа принимаются центры найденных кластеров.

Задача кластеризации, как правило, решается с помощью таких алгоритмов, как k -средних, DBSCAN, FOREL [19]; в работе [20] построен алгоритм с использованием фрактальных свойств сети. Каждый из этих алгоритмов имеет свои достоинства и недостатки. Например, алгоритм k -средних целесообразно применять при равномерном распределении пользователей и заранее известном числе кластеров (форма кластера при росте количества объектов стремится к форме многогранника). DBSCAN уместен при выраженных скоплениях пользователей, причем плотность пользователей в зоне кластеров должна быть примерно одинаковой (кластер может иметь произвольную форму). FOREL позволяет найти решения как при выраженных скоплениях пользователей, так и при их равномерном распределении (форма кластера стремится к форме круга); алгоритм применим при выраженных скоплениях пользователей.

В данной задаче алгоритм кластеризации должен, наряду с выделением групп пользователей (кластеров), учитывать их взаимное расположение, т. к. маршрутизаторы в центрах кластеров должны образовывать связную сеть, а сама эта сеть должна быть связана с сетью подвижной связи.

Далее предлагается алгоритм, построенный на основе алгоритма кластеризации FOREL [21]. Идея данного алгоритма предпочтительна по ряду причин. Во-первых, в данном методе задается размер кластера R , который фактически является максимально возможным размером выделяемого кластера. В задаче построения сети это позволяет определить размер кластера через достижимую скорость передачи данных b на участке пользователь-точка доступа. Если определена зависимость $b = f(R)$, то:

$$R = \arg\{b_{\min} = f(R)\}, \quad (1)$$

где $b_{\min}$ – минимально допустимая скорость передачи данных, задаваемая в качестве параметра при построении сети.

Функциональная зависимость $f(R)$ специфична для применяемого стандарта или группы стандартов и может быть определена на их основе, например в [22].

Для большей общности эту функцию можно аппроксимировать с помощью известного выражения К. Шеннона, с учетом корректирующего коэффициента α :

$$b(R) = \alpha w \log_2(1 + \gamma(R)), \quad (2)$$

где α – корректирующий коэффициент, учитывающий отличие используемой технологии от аналитической модели (теоретически достижимой скорости); γ – отношение сигнал/шум; w – полоса пропускания (Гц).

В общем случае $0 < \alpha$. Если не применяются методы пространственно-временного уплотнения (ММО) – $0 < \alpha < 1$. При применении таких методов коэффициент может превышать единицу.

Отношение сигнал/шум, как функция расстояния, может быть определено как:

$$\gamma(R) = \frac{s_{RX}}{p_n} = \frac{1}{p_n} a(R) s_{TX}, \quad (3)$$

где $a(R)$ – зависимость затухания сигнала от расстояния, в качестве зависимости может быть принята одна из известных моделей [23].

Зависимость скорости передачи данных от расстояния, согласно стандарту IEEE 802.11ac, при $w = 20$ МГц, приведена на рисунке 1 (синяя линия). Красная кривая получена согласно выражению (2): $w = 20$ МГц, $\alpha = 0,5$.

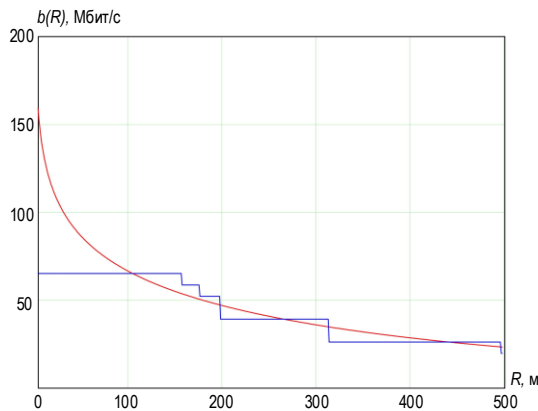


Рис. 2. Зависимость скорости передачи данных от расстояния между оборудованием пользователя и точкой доступа

Fig. 2. Data Transfer Rate Dependence on the Distance Between the User Equipment and the Access Point

Приведенный график показывает достаточно высокую близость модели (2) к зависимости, полученной по данным стандарта. Наибольшая ошибка наблюдается при малых расстояниях, т. е. при высоких значениях отношения сигнал/шум. Это мож-

но объяснить технологическими возможностями, которые учитываются в данном стандарте.

Применение выбранного подхода кластеризации решает задачу минимизации суммы расстояний между элементами и центрами кластеров. При описании системы приведенными выше выражениями целевая функция кластеризации будет выражаться как:

$$O = \max_{c_j, j=1 \dots n} \sum_{j=1}^n \sum_{i=1}^{m_j} b(c_j, u_{ji}), \quad (4)$$

где $b(c_j, u_{ji}) = b(d(c_j, u_{ji}))$ – достижимая скорость передачи между центром j -го кластера и i -м элементом j -го кластера; $d(c_j, u_{ji})$ – расстояние между центром j -го кластера и i -м элементом j -го кластера.

Таким образом, применение данного метода позволяет выбрать позиции центров кластеров (позиций точек доступа) так, чтобы суммарная скорость передачи в сети была максимальна, при этом условие (1) обеспечивает минимально допустимую величину скорости передачи данных.

Решение (4) дает близкие к оптимальным позиции центров кластеров по отношению к элементам, которыми являются пользователи. Однако в данном случае необходимо учитывать положение центров кластеров относительно базовых станций сети подвижной связи и относительно друг друга.

Будем полагать, что R_c – радиус зоны связи кластера, R_m – радиус зоны связи маршрутизатора с базовой станцией сети подвижной связи. Эти параметры должны быть введены в виде ограничений в используемый алгоритм кластеризации.

Особенность данного алгоритма состоит в том, что выбор позиции центра кластера производится с учетом упомянутых ограничений, а именно:

$$\begin{aligned} \exists j \in \{1 \dots n\}, d(c_j, B) \leq R_m \\ \exists(j, i) \in \{1 \dots n\}, j \neq i, d(c_j, c_i) \leq R_c \end{aligned} \quad (5)$$

Эти ограничения фактически определяют то, что как минимум один из маршрутизаторов должен быть в зоне связи базовой станции сети подвижной связи, и каждый из маршрутизаторов должен быть в зоне связи хотя бы одного из соседних маршрутизаторов. Эти ограничения обеспечивают соединения между маршрутизаторами, расположенными на БПЛА, и с сетью подвижной связи.

Эти ограничения могут быть усилены заданием минимального количества связей между маршрутизаторами $r_{\min}$ и с базовыми станциями сети подвижной связи $M_{\min}$ (6).

Псевдокод предлагаемого алгоритма приведен ниже.

$$\begin{aligned} \exists J \in \{1 \dots n\}, \quad |J| \geq M_{\min}, d(c_j \in J, B) \leq R_M \\ \exists L \in \{1 \dots n\}, \quad |L| \geq r_{\min}, d(c_j \in L, c_i \in L) \leq R_C, \quad j \neq i \end{aligned} \quad (6)$$

Исходные данные:

– множество объектов кластеризации (пользователей) $U = \{u_i\}$, $i = 1 \dots m$; каждому из объектов соответствует точка с заданными координатами $u_i \triangleq p(u_i) = (x_i^{(u)}, y_i^{(u)}, z_i^{(u)})$, $i = 1, \dots, m$;

– множество базовых станций сети подвижной связи $V = \{v_i\}$, $i = 1 \dots k$; каждому из объектов соответствует точка с заданными координатами $v_i \triangleq p(v_i) = (x_i^{(k)}, y_i^{(k)}, z_i^{(k)})$, $i = 1, \dots, M$;

– множество центров кластеров $C = \{c_i\}$, $i = 1, \dots, n$

Каждому из центров кластеров c_j при $j = 1, \dots, n$ соответствует точка с заданными координатами $c_i \triangleq p(c_i) = (x_i^{(c)}, y_i^{(c)}, z_i^{(c)})$, $i = 1, \dots, n$.

Алгоритм

```

 $U_0 = \{u_i\}$ ,  $i = 1 \dots m$ ; // Множество некластеризованных объектов
 $C = \emptyset$  // Множество кластеров
 $i=1$ ;
//Случайная точка в зоне обслуживания
 $p(c_i) = \{randomPos(x, y, z) \in S, \quad d(c_i, v_q) \leq R_M, \quad d(c_i, c_j) \leq R_C, \quad v_q \in V\}$ 
 $C_0 = \{u_j \in U_0; \quad d(c_i, u_j) \leq R\}$ ; //Образовать кластер  $C_0$ 
While  $U_0 \neq \emptyset$ 
{
//Вычислить центр кластера
 $cm = S(U, V, R)$ 
//Если центр масс не совпадает с центром кластера и он находится на удалении не превышающим  $R_C$  от центра другого кластера
if  $cm \neq p(c_i)$ 
{
if  $\exists c_j \in C, \quad d(cm, c_j) \leq R_C$ 
 $p(c_i) = cm$  //Перейти в центр масс
else
 $\omega = \omega + \Delta\omega$ ;
}
else
{
 $U_0 = U_0 \setminus C_0$  //Исключить элементы кластера из множества некластеризованных элементов
 $C = C \cup cm$ 
}
} //End while
ReDistributeNones(cond) //Опционально
    
```

Центр кластера (центр масс) вычисляется согласно выражению (7), где K – число элементов кластера; x_0, y_0, z_0 – координаты центра кластера; η_i – весовой коэффициент элемента (пользователя); x_{Cn}, y_{Cn}, z_{Cn} – координаты центра ближайшего соседнего кластера, ω – весовой коэффициент.

Особенность вычисления центра кластера состоит в том, что учитывается расстояние до центра ближайшего кластера. Изменяя величину ω , можно управлять «тяготением» центра искомого кластера к ближайшему соседнему кластеру.

Приведенный выше алгоритм описывает общий трехмерный случай, когда кластер представляет собой сферу. Ввиду недостаточной наглядности представления результатов кластеризации в пространстве на рисунке 3 приведен пример кластеризации на плоскости. В данном примере моделируется кластеризация 1000 пользователей, распределенных на плоскости. Для моделирования используется точечный процесс Гиббса [24].

После выполнения приведенного выше алгоритма будут найдены центры кластеров. На этом алгоритм может быть завершен, а элементы могут быть отнесены к кластерам в том порядке, в котором выделялись кластеры в ходе выполнения алгоритма. Однако такой подход дает хорошие решения в случае наличия скоплений элементов (пользователей), причем размер этих скоплений не превышает размера кластера R . В противном случае размеры кластеров могут сильно отличаться, поэтому целесообразно по завершении алгоритма выполнить перераспределение элементов путем отнесения их к ближайшим центрам кластеров. В приведенном выше алгоритме эта операция выполняется процедурой ReDistributeNones(y/n). Ее аргумент определяет требуется ли выполнение перераспределения элементов.

После операции перераспределения элементов кластеры будут иметь форму многогранников, стороны которых определяются диаграммой Вороного [25], построенной относительно найденных центров кластеров (рисунок 4).

$$S_m(C_0, V, R) = \left\{ \begin{aligned} x_0 &= \frac{1}{K} \left(\sum_{i=1}^K x_i \eta_i + x_{Cn} \omega \right), y_0 = \frac{1}{K} \left(\sum_{i=1}^K y_i \eta_i + y_{Cn} \omega \right), \\ z_0 &= \frac{1}{K} \left(\sum_{i=1}^K z_i \eta_i + z_{Cn} \omega \right) \end{aligned} \right\}, \quad (7)$$

$$K = |C_0|, \quad (x_i, y_i, z_i) = p(u_i), \quad u_i \in C_0, \quad (x_{Cn}, y_{Cn}, z_{Cn}) = p(c_n), \quad c_n \in C, \quad \eta_i, \omega \geq 0.$$

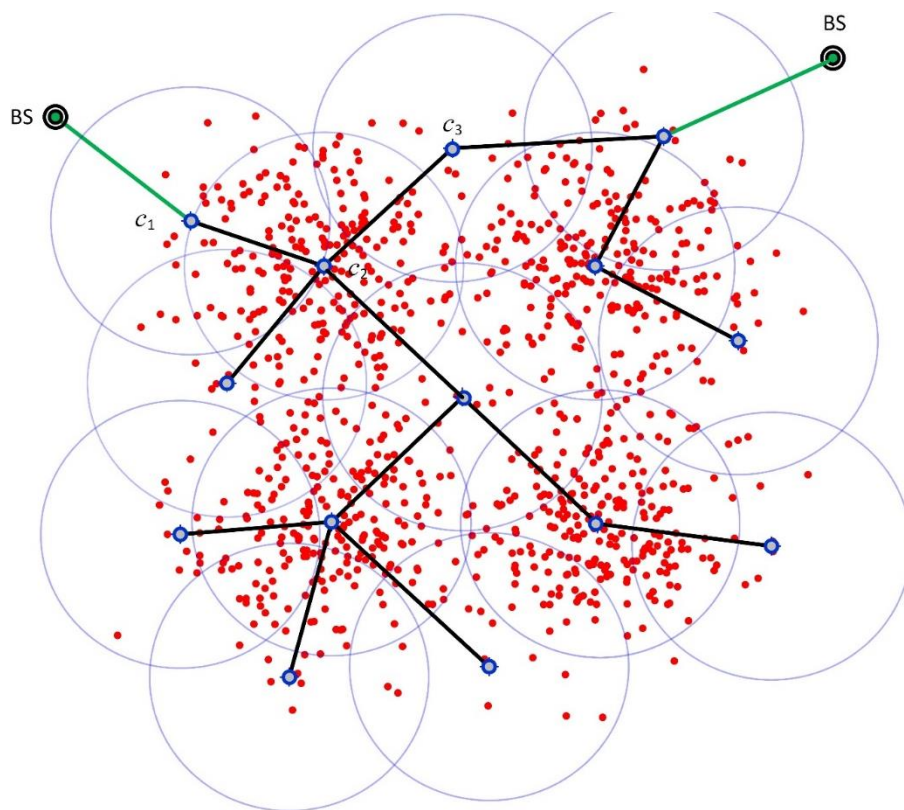


Рис. 3. Результат кластеризации и выбора структуры сети

Fig. 3. Result of Clustering and Selection of Network Structure

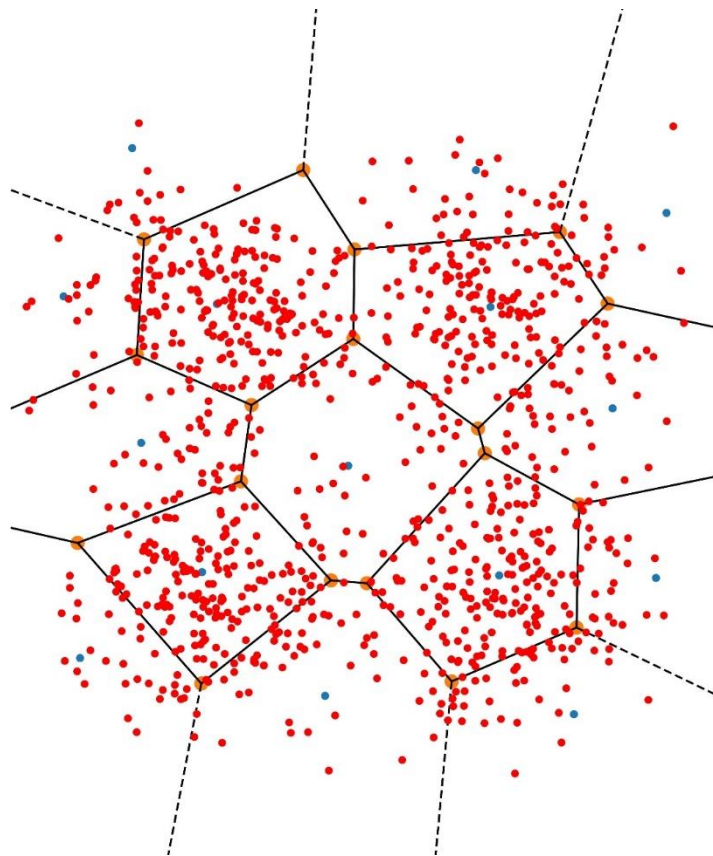


Рис. 4. Результат модификации формы кластеров

Fig. 4. Result of Cluster Shape Modification

В качестве критерия необходимости перераспределения предлагается использовать дисперсию количества элементов в кластерах, определяемую как:

$$\text{cond} = \begin{cases} \text{true} & D_{II}(K) < D_I(K) \\ \text{false} & D_{II}(K) \geq D_I(K) \end{cases} \quad (8)$$

где D_I , D_{II} – дисперсии числа элементов кластеров для алгоритма без перераспределения элементов и с их перераспределением, соответственно.

Дисперсии для обоих случаев могут быть найдены по выражению:

$$D(K) = \frac{1}{n} \sum_{i=1}^n (K_i - E(K))^2, \quad E(K) = \frac{1}{n} \sum_{i=1}^n K_i, \quad (9)$$

где n – число кластеров; K_i – число элементов в i -м кластере.

Предложенный метод и алгоритм позволяют произвести кластеризацию пользователей в заданной зоне обслуживания с учетом допустимых расстояний между центрами кластеров и допустимых расстояний до базовых станций сети подвижной связи. Благодаря этим качествам данный метод позволяет построить связную mesh-сеть на основе маршрутизаторов, размещаемых в центрах кластеров, имеющую связь с базовыми станциями сети подвижной связи. Алгоритм кластеризации обеспечивает выделение кластеров по максимуму качества обслуживания.

Маршрутизация трафика

Найденные позиции маршрутизаторов (центры кластеров) объединены в mesh-сеть, структура которой может быть определена кратчайшими маршрутами между отправителем и получателем. Под кратчайшим будем понимать маршрут, обеспечивающий наибольшее качество обслуживания трафика. Качество обслуживания в этом случае характеризуется такими показателями как достижимая скорость передачи данных, задержка доставки и вероятность потери данных.

Структура mesh-сети определяется матрицей:

$$ST = \begin{bmatrix} l_{11} & \dots & l_{1n} \\ \vdots & & \vdots \\ l_{n1} & \dots & l_{nn} \end{bmatrix}, \quad l_{ij} = \begin{cases} q_{ij} & l_{ij} \leq R_c \\ \infty & l_{ij} > R_c \end{cases} \quad (10)$$

$i, j = 1 \dots n,$

где q_{ij} – показатель качества линии связи между узлами сети; l_{ij} – характеристика линии связи между узлами сети.

Как показатель качества предлагается интегральный показатель, учитывающий скорость передачи данных, нагрузку и вероятность потерь:

$$q_{ij} = -\omega_1 b_{ij} + \omega_2 \rho_{ij} + \omega_3 p_{ij}, \quad (11)$$

где $\omega_1, \omega_2, \omega_3$ – коэффициенты условной стоимости единицы соответствующего показателя.

Интегральный показатель имеет смысл условного выигрыша или расходов, в зависимости от знака. Знак слагаемого в выражении (11) положителен для показателей, рост которых ведет к увеличению условных расходов, в данном случае – к уменьшению качества обслуживания.

В таком случае использование любого метода поиска кратчайшего пути для сети, описываемой матрицей (10), позволяет выбрать маршрут, имеющий минимальное значение суммы коэффициентов для каждого из ребер, входящих в него, формально можно записать как:

$$\Omega_{st} = \arg \left(\Omega_{st} = \Omega_{st} \cup q_{ij} \mid \min_{q_{ij} \in ST} \sum_{r=1}^{L_{st}} q_{ij}, \right. \\ \left. i, j = 1, \dots, n, i \neq j \right), \quad (9)$$

где Ω_{st} – множество вершин (узлов), входящих в кратчайший путь между вершинами s и t .

Для нахождения всех кратчайших путей можно использовать, например, алгоритм Данцига [26].

Эффективность

Для оценки эффективности предложенных методов, в качестве альтернативы, выберем не модифицированный алгоритм FOREL. При его применении распределение маршрутизаторов в зоне обслуживания не гарантирует связность mesh-сети и доступность сети подвижной связи. Позиции маршрутизаторов в такой ситуации случайны и зависят от распределения пользователей, но никак не ограничены в части взаимного расположения; связность образуемой ими сети также является случайной величиной.

Если маршрутизаторы, размещенные в выбранных позициях (центрах кластеров), не образуют связной сети, то пользователи, включенные в эти маршрутизаторы, не имеют доступа к другой части сети и выхода во внешнюю сеть. Таким образом, для сравнения методов наиболее показательной характеристикой является связность сети, образованной при размещении маршрутизаторов.

Связность сети можно описать вероятностью существования маршрута между узлами, которую можно оценить долей возможных маршрутов в сети:

$$\text{con} = \frac{n_e}{(n+1)^2}, \quad (13)$$

где n – количество кластеров (маршрутизаторов); n_e – количество доступных маршрутов.

Величина $(n + 1)^2$ – это максимально возможное количество маршрутов между всеми n узлами сети и одним узлом, описывающим базовую станцию сети подвижной связи, которое имеет место при полной связности сети.

Вероятность связности (13) в случае не модифицированного алгоритма не превышает единицы, а в случае модифицированного алгоритма она всегда равна единице. Величину $1 - \text{con}$ будем рассматривать как выигрыш, полученный при использовании предложенных методов:

$$e = 1 - \text{con}. \quad (14)$$

Эта величина зависит от радиуса зоны связи маршрутизаторов R_c и R_m , полагая, что эти величины равны $R_c = R_m$, можно оценить зависимость $e(R_c/W)$, где W – размер зоны обслуживания.

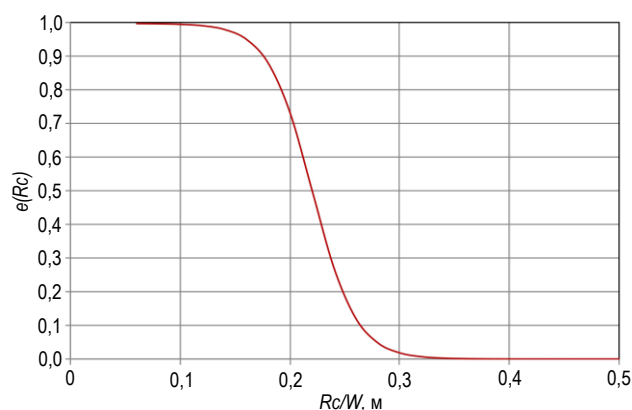


Рис. 5. Зависимость эффективности метода от радиуса зоны связи между кластерами R_c

Fig. 5. Dependence of the Efficiency of the Method on the Radius of the Communication Zone between Clusters R_c

Можно заметить, что применение метода целесообразно, когда радиус зоны связи маршрутизаторов в mesh-сети меньше размера зоны обслуживания. Это естественно, т. к. в противном случае любое расположение маршрутизатора равнозначно с точки зрения обеспечения связности.

В приведенном выше примере, при размере зоны обслуживания 500×500 м и заданном распре-

делении пользователей (размер зоны связи маршрутизатора – около 0,23 размера зоны обслуживания) эффективность предложенного метода составляет около 50 %. Эффективность тем выше, чем меньше отношения радиуса зоны связи маршрутизатора к размеру зоны обслуживания. Таким образом его применение целесообразно при необходимости обслуживания относительно больших площадей или объемов.

Выводы

Во-первых, моделирование линии связи между маршрутизатором и пользователем с помощью модели К. Шеннона с учетом технологических отличий его реализации в действующем стандарте сети беспроводного доступа позволяет решать проблему кластеризации пользователей как задачу максимизации суммарной достижимой скорости передачи данных.

Во-вторых, предложенный алгоритм кластеризации, являющийся модификацией алгоритма FOREL, обеспечивает выбор позиций маршрутизаторов для обслуживания пользователей и организации связанной mesh-сети и связи этой сети с базовой станцией подвижной связи.

В-третьих, модификация алгоритма, обеспечивающая перераспределение элементов, позволяет повысить качество результата путем выравнивания их количества в выделяемых кластерах.

В-четвертых, метод маршрутизации трафика в mesh-сети обеспечивает выбор маршрутов на основе минимизации интегрального показателя качества, учитывающего достижимую скорость передачи данных, величину нагрузки и вероятность потерь.

В-пятых, предложенные методы распределения маршрутизаторов в рое БПЛА или на единичных или привязных БПЛА эффективен в случае, когда размер зоны связи маршрутизатора меньше размера зоны обслуживания.

Список источников

1. Taleb T., Benzaïd C., Lopez M.B., Mikhaylov K., Tarkoma S., Kostakos P., et al. 6G System Architecture: A Service of Services Vision // ITU Journal on Future and Evolving Technologies. 2022. Vol. 3. Iss. 3.
2. Rec. ITU-T Technical Report (01/2020). Network 2030 – Additional Representative Use Cases and Key Network Requirements for Network 2030.
3. Rec. ITU-T Deliverable (10/2019). New Services and Capabilities for Network 2030: Description, Technical Gap and Performance Target Analysis.
4. Li R. Network 2030. A Blueprint of Technology, Applications and Market Drivers Towards the Year 2030 and Beyond. 2019.
5. Rec. ITU-T Technical Specification (06/2020). Network 2030 Architecture Framework.
6. Волков А.Н., Мутханна А.С.А., Кучерявый А.Е., Бородин А.С., Парамонов А.И., Владимиров С.С. и др. Перспективные исследования сетей и услуг 2030 в лаборатории 6G Meganetlab СПбГУТ // Электросвязь. 2023. № 6. С. 5–14. DOI:10.34832/ELSV.2023.43.6.001. EDN:CJSYLS
7. Демидов Н.А. Исследование трафика 3d-видеопотока на имитационной модели // Электросвязь. 2024. № 3. С. 44–48. DOI:10.34832/ELSV.2024.52.3.008. EDN:DNQCWX

8. Rec. ITU Focus Group Technical Specification (12/2023). Definition of metaverse.
9. Mane-Deshmukh P.V. Designing of Wireless Sensor Network to Protect Agricultural Farm from Wild Animals // *i-Manager's Journal on Information Technology*. 2018. Vol. 7. Iss. 4. PP. 30–36.
10. Shannon C.E., Weaver W. *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press, 1964.
11. Akyildiz I.F., Han C., Hu Z., Nie S., Jornet J.M. Terahertz Band Communication: An Old Problem Revisited and Research Directions for the Next Decade // *IEEE Transactions on Communications*. 2022. Vol. 70. Iss. 6. PP. 4250–4285. DOI:10.1109/TCOMM.2022.3171800
12. Petrov V., Pyattaev A., Moltchanov D., Koucheryavy Y. Terahertz band communications: Applications, research challenges, and standardization activities // *Proceedings of the 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT, Lisbon, Portugal, 18–20 October 2016)*. IEEE, 2016. PP. 183–190. DOI:10.1109/ICUMT.2016.7765354
13. 120 results for "5g router" // Amazon. URL: <https://www.amazon.com/5g-router/s?k=5g+router> (Accessed 01.07.2024)
14. Дорохова А.А., Парамонов А.И. Исследование трафика и качества обслуживания в самоорганизующихся сетях на базе БПЛА // *Информационные технологии и телекоммуникации*. 2016. Т. 4. № 2. С. 12–25. EDN:XDCORF
15. Варельджян К.С., Парамонов А.И., Киричек Р.В. Оптимизация траектории движения БПЛА в летающих сенсорных сетях // *Электросвязь*. 2015. № 7. С. 20–25. EDN:UAYFOL
16. Захаров М.В., Киричек Р.В., Парамонов А.И. Задача распределения ресурсов в группах БПЛА // *Информационные технологии и телекоммуникации*. 2015. Т. 3. № 1. С. 62–70. EDN:TUXWKP
17. Вишнеvский В.М. Методы и алгоритмы проектирования и реализации привязных высотных беспилотных телекоммуникационных платформ // XIII Всероссийское совещание по проблемам управления ВСПУ-2019 (Москва, Россия, 17–20 июня 2019 г.). М.: Институт проблем управления им. В.А. Трапезникова РАН, 2019. С. 40–42. DOI:10.25728/vspu.2019.0040. EDN:KFCQMJ
18. Факторный, дискриминантный и кластерный анализ. Пер. с англ. М.: Финансы и статистика, 1989. 215 с.
19. 2.3. Clustering // Scikit-learn. URL: <https://scikit-learn.org/stable/modules/clustering.html> (Accessed 01.07.2024)
20. Марочкина А.В. Моделирование и кластеризация трехмерной сети интернета вещей с применением метода оценки фрактальной размерности // *Электросвязь*. 2023. № 6. С. 60–66. DOI:10.34832/ELSV.2023.43.6.008. EDN:ZBNQKI
21. Загоруйко Н.Г., Ёлкина В.Н., Лбов Г.С. Алгоритмы обнаружения эмпирических закономерностей. Новосибирск: Наука, 1985. 110 с.
22. Викулов А.С., Парамонов А.И. Модель канала OFDM в задаче оценки эффективности сети IEEE 802.11 // *Инфокоммуникационные технологии*. 2018. Т. 16. № 3. С. 290–297. DOI:10.18469/ikt.2018.16.3.06. EDN:EMWAAZ
23. Рекомендация МСЭ-R P.1238-9 (06/2017). Данные о распространении радиоволн и методы прогнозирования для планирования систем радиосвязи внутри помещений и локальных зонных радиосетей в частотном диапазоне 300 МГц–100 ГГц.
24. Daley D.J., Vere-Jones D. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Springer Science & Business Media, 2006. 471 p.
25. Препарата Ф., Шеймос М. Вычислительная геометрия: Введение. Пер. с англ. М.: Мир, 1989. 478 с.
26. Майника Э. Алгоритмы оптимизации на сетях и графах. Пер. с англ. М.: Мир. 1981. 323 с.

References

1. Taleb T., Benzaïd C., Lopez M.B., Mikhaylov K., Tarkoma S., Kostakos P., et al. 6G System Architecture: A Service of Services Vision. *ITU Journal on Future and Evolving Technologies*. 2022;3(3).
2. Rec. ITU-T Technical Report. *Network 2030 – Additional Representative Use Cases and Key Network Requirements for Network 2030*. June 2020.
3. Rec. ITU-T Deliverable. *New Services and Capabilities for Network 2030: Description, Technical Gap and Performance Target Analysis*. October 2019.
4. Li R. *Network 2030. A Blueprint of Technology, Applications and Market Drivers Towards the Year 2030 and Beyond*. 2019.
5. Rec. ITU-T Technical Specification. *Network 2030 Architecture Framework*. June 2020.
6. Volkov A.N., Muthanna A.S.A., Kucheryavy A.E., Borodin A.S., Paramonov A.I., Vladimirov S.S., et al. Perspective research of networks and services 2030 in the laboratory 6G MEGANETLAB SPBSUT. *Electrosvyaz*. 2023;6(5–14). (in Russ.) DOI:10.34832/ELSV.2023.43.6.001. EDN:CJSYLS
7. Demidov N.A. Investigation of 3D video stream traffic on a simulation model. *Electrosvyaz*. 2024;3:44–48. (in Russ.) DOI:10.34832/ELSV.2024.52.3.008. EDN:DNQCWX
8. Rec. ITU Focus Group Technical Specification. *Definition of metaverse*. December 2023.
9. Mane-Deshmukh P.V. Designing of Wireless Sensor Network to Protect Agricultural Farm from Wild Animals. *i-Manager's Journal on Information Technology*. 2018;7(4):30–36.
10. Shannon C.E., Weaver W. *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press, 1964.
11. Akyildiz I.F., Han C., Hu Z., Nie S., Jornet J.M. Terahertz Band Communication: An Old Problem Revisited and Research Directions for the Next Decade. *IEEE Transactions on Communications*. 2022;70(6):4250–4285. DOI:10.1109/TCOMM.2022.3171800
12. Petrov V., Pyattaev A., Moltchanov D., Koucheryavy Y. Terahertz band communications: Applications, research challenges, and standardization activities. *Proceedings of the 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, ICUMT, 18–20 October 2016, Lisbon, Portugal*. IEEE; 2016. p.183–190. DOI:10.1109/ICUMT.2016.7765354
13. Amazon. 120 results for "5g router". URL: <https://www.amazon.com/5g-router/s?k=5g+router> [Accessed 01.07.2024]

14. Dorohova A., Paramonov A. Traffic and Quality of service research in a Flying Ad-Hoc Network. *Telecom IT*. 2016;4(2): 12–25. (in Russ.) EDN:XDCORF
15. Varelzhian K.S., Paramonov A.I., Kirichek R.V. Optimization of the UAV's motion trajectory in flying ubiquitous sensor networks. *Electrosvyaz*. 2015;7:20–25. (in Russ.) EDN:UAYFOL
16. Zaharov M.V., Kirichek R.V., Paramonov A.I. Resource Allocation Problems in Groups UAVs. *Telecom IT*. 2015;3(1):62–70. (in Russ.) EDN:TUXWKP
17. Vishnevskiy V.M. Methods and algorithms of design and realization of tethered high-altitude unmanned telecommunication platforms. *Proceedings of the XIIIth All-Russian Conference on Management Problems VSPU-2019, 17–20 June 2019, Moscow, Russia*. Moscow: Institute of Control Sciences RAS Publ.; 2019. p.40–42. (in Russ.) DOI:10.25728/vspu.2019.0040. EDN:KFCQMJ
18. *Factor, discriminant and cluster analysis*. Moscow: Finance and statistics, 1989. 215 p. (in Russ.)
19. *Scikit-learn*. 2.3. Clustering. URL: <https://scikit-learn.org/stable/modules/clustering.html> [Accessed 01.07.2024]
20. Marochkina A.V. Modeling and clustering a 3d internet of things network using the fractal dimension estimation method. *Electrosvyaz*. 2023;6:60–66. (in Russ.) DOI:10.34832/ELSV.2023.43.6.008. EDN:ZBNQKI
21. Zagoruiko N.G., Yolkina V.N., Lbov G.S. Algorithms for detection of empirical regularities. Novosibirsk: Nauka Publ.; 1985. 110 p. (in Russ.)
22. Vikulov A.S., Paramonov A.I. OFDM channel model in the problem of the IEEE 802.11 network efficiency estimation. *Infocommunication Technologies*. 2018;16(3):290–297. (in Russ.) DOI:10.18469/ikt.2018.16.3.06. EDN:EMWAAZ
23. Rec. ITU-R P.1238-9. *Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 100 GHz*. June 2017.
24. Daley D.J., Vere-Jones D. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Springer Science & Business Media; 2006. 471 p.
25. Preparato F.P., Shamos M.I. *Computational Geometry. An Introduction*. Springer-Verlag; 1985.
26. Minieka E. *Optimization Algorithms for Networks and Graphs*. Marcel Dekker; 1978. 356 p.

Статья поступила в редакцию 01.07.2024; одобрена после рецензирования 09.08.2024; принята к публикации 15.08.2024.

The article was submitted 01.07.2024; approved after reviewing 09.08.2024; accepted for publication 15.08.2024.

Информация об авторах:

**КУЗНЕЦОВ
Константин Алексеевич**

аспирант кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0001-6167-2711>

**ПАРАМОНОВ
Александр Иванович**

доктор технических наук, доцент, профессор кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-4104-3504>

**МУТХАННА
Аммар Салех Али**

доктор технических наук, доцент, профессор кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0003-0213-8145>

**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/0000-0003-0213-8145>

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.

Обзорная статья

УДК 621.396.969

<https://doi.org/10.31854/1813-324X-2024-10-4-73-98>

Комплекс моделей позиционирования устройств в сетях шестого поколения. Часть 1. Обзор методов и постановка задачи

 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. На сегодняшний день терагерцовые радиосистемы рассматриваются как технологическая основа интеграции методов и средств радиосвязи и радиолокации в перспективных сетях шестого поколения. Если в сетях 4G LTE возможности позиционирования пользовательских устройств с использованием инфраструктуры базовых станций рассматривались как вспомогательные опции, то в сетях 5G NR технологии определения местоположения (ОМП) стали полноправными сервисами, требования к которым специфицированы наряду с услугами связи. Новой тенденцией позиционирования в сетях 5G NR, по сравнению с сетями 4G LTE, стала однопозиционная оценка координат и ориентации устройства по сигналам единственной базовой станции с возможностью различать прямые и отраженные сигналы. Сети 6G все еще находятся в стадии становления, однако уже можно констатировать, что они знаменуют собой очередной этап эволюции цифровых экосистем, который характеризуется конвергенцией технологий связи, локализации и зондирования радиоэфира и окружающего пространства радиотехническими средствами.

Цель. Настоящая работа открывает цикл исследований, посвященный обзору моделей, методов и алгоритмов позиционирования устройств в сетях 6G. Целью цикла является поиск и обоснование новых радиотехнических средств достижения дециметровой точности оценок координат устройств 6G. В первой части цикла выполняется обзор методов и формализация модели сбора первичных измерений.

Методом исследования является аналитический обзор состояния проблемы по актуальным научным публикациям, концептуальное моделирование, категориальный подход, экспертное комбинирование, сопоставительный анализ, формализация, математическое и имитационное моделирование.

Решение/результаты. В результате обзора методов позиционирования устройств при переходе к сетям 6G актуализируются ключевые показатели эффективности и сценарии ОМП. В результате сопоставительного анализа сетей 5G и 6G систематизируются новые факторы, достоинства и недостатки технологий позиционирования при переходе от сетей диапазона миллиметровых волн к сетям терагерцового диапазона. Формализованная математическая модель сбора первичных измерений используется в имитационной модели оценки точности позиционирования устройств во второй части цикла.

Новизна. Настоящий цикл является первым подобным исследованием в отечественном научном сегменте по сетевому позиционированию шестого поколения терагерцового диапазона, в котором в авторской редакции выполнен обзор методов и систематизирован комплекс новых факторов ОМП в сетях связи.

Теоретическая значимость обзора-анализа заключается в установлении как технологических препятствий, так и новых возможностей по увеличению точности позиционирования при переходе к сетям 6G.

Практическая значимость формализованной математической модели заключается в ее последующей программной реализации для численного обоснования пределов точности позиционирования в сетях 6G.

Ключевые слова: позиционирование, ориентация, терагерцовый диапазон, сети 6G, реконфигурируемые интеллектуальные поверхности, ближняя зона, составной массив антенной решетки, локальная и глобальная системы координат

Источник финансирования: статья подготовлена в рамках прикладных научных исследований СПбГУТ, регистрационный номер 1023031600087-9-2.2.4;2.2.5;2.2.6;1.2.1;2.2.3 в ЕГИСУ НИОКТР.

Ссылка для цитирования: Фокин Г.А. Комплекс моделей позиционирования устройств в сетях шестого поколения. Часть 1. Обзор методов и постановка задачи // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 73–98. DOI:10.31854/1813-324X-2024-10-4-73-98. EDN:FTOVZJ

Review research

<https://doi.org/10.31854/1813-324X-2024-10-4-73-98>

A Set of Models for Device Positioning in Sixth Generation Networks. Part 1. Methods Survey and Problem Statement

 Grigoriy A. Fokin, fokin.ga@sut.ru

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

Annotation

Relevance. Today, terahertz radio systems are considered as a technological basis for integrating methods and means of radio communication and radar in promising sixth-generation networks. If in 4G LTE networks the capabilities of positioning user equipment using the infrastructure of base stations were considered as auxiliary options, then in 5G NR networks, location determination technologies (LDTs) have become full-fledged services, the requirements for which are specified along with communication services. A new trend in positioning in 5G NR networks, compared to 4G LTE networks, has become a single-position assessment of the coordinates and orientation of the user equipment based on signals from a single base station with the ability to distinguish between direct and reflected signals. 6G networks are still in their infancy, but it can already be stated that they mark the next stage in the evolution of digital ecosystems, which is characterized by the convergence of communication technologies, localization and sensing of radio air and the surrounding space by radio engineering means.

Purpose. This work opens a research cycle devoted to the review of models, methods and algorithms for positioning devices in 6G networks. The goal of the cycle is to find and justify new radio engineering means for achieving decimeter accuracy in 6G device coordinate estimates. The first part of the cycle provides an overview of the methods and formalization of the model for collecting primary measurements.

Method is an analytical review of the state of the problem based on current scientific publications, conceptual modeling, categorical approach, expert combination, comparative analysis, formalization, mathematical and simulation modeling.

Results. As a result of the review of device positioning methods during the transition to 6G networks, key performance indicators and LDT scenarios are updated. As a result of the comparative analysis of 5G and 6G networks, new factors, advantages and disadvantages of positioning technologies during the transition from millimeter wave networks to terahertz networks are systematized. A formalized mathematical model for collecting primary measurements is used in the simulation model for assessing the accuracy of device positioning in the second part of the cycle.

Novelty. This cycle is the first such study in the Russian scientific segment on network positioning of the sixth generation of the terahertz range, in which the author's version provides an overview of methods and a systematization of a set of new factors of the OMP in communication networks.

The theoretical significance of the review-analysis lies in the establishment of both technological obstacles and new opportunities for increasing positioning accuracy during the transition to 6G networks.

The practical significance of the formalized mathematical model lies in its subsequent software implementation for numerical justification of the limits of positioning accuracy in 6G networks.

Keywords: positioning, orientation, terahertz range, 6G networks, reconfigurable intelligent surfaces, near field, antenna array of subarrays, local and global coordinate systems

Funding: The scientific article was prepared within the framework of applied scientific research SPbSUT, registration number 1023031600087-9-2.2.4;2.2.5;2.2.6;1.2.1;2.2.3 in the information system (<https://www.rosrid.ru/information>).

For citation: Fokin G.A. A set of Models for Device Positioning in Sixth Generation Networks. Part 1. Methods Survey and Problem Statement. *Proceedings of Telecommunication Universities*. 2024;10(4):73–98. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-73-98. EDN:VYMCTD

1. Введение

Впервые идея использования инфраструктуры базовых станций (BS, *аббр. от англ.* Base Station) и радиосигналов сотовых сетей подвижной радиосвязи для решения задач определения местоположения (ОМП) пользовательских устройств (UE, *аббр. от англ.* User Equipment) была оформлена Федеральной комиссией по электросвязи США (Federal Communications Commission) в середине 90-х гг. в виде требований к позиционированию устройства, совершающего экстренный вызов [1]. Несмотря на достаточно скромную по сегодняшним меркам точность позиционирования (до 100 м), данное решение зарубежного регулятора послужило мощным импульсом к развитию технологий сетевого позиционирования в сотовых системах подвижной радиосвязи [2]. В сетях второго поколения 2G GSM (*аббр. от англ.* Global System for Mobile) изначально не было предусмотрено каких-либо специальных сигналов и процедур сетевого позиционирования, поэтому точность ОМП радиотехническими методами ограничивалась разрешением по времени и определялась длительностью передачи одного бита 3,69 мкс, что при синхронизации в радиолинии BS-UE по времени оборота (RTT, *аббр. от англ.* Round Trip Time) позволяло достигнуть точности первичных дальномерных (ДМ) измерений $d = RTT/2 \cdot c \sim 550$ м [3]. В сетях третьего поколения 3G UMTS (*аббр. от англ.* Universal Mobile Telecommunications System) поддержка сетевого позиционирования UE на уровне радиоинтерфейса так и не появилась, однако при использовании полосы в 5 МГц длительность элементарного временного символа 0,26 мкс позволила повысить разрешение ДМ измерений почти на порядок – до 78 м [4]. В сетях четвертого поколения LTE (*аббр. от англ.* Long-Term Evolution) впервые были специфицированы специальные опорные сигналы позиционирования (PRS, *аббр. от англ.* Positioning Reference Signals), при использовании которых в полосе 20 МГц с длительностью элементарного временного интервала 32,552 нс достигалась точность ДМ разрешения до 9,7656 м, что в очередной раз почти на порядок превышало точность по сравнению с предыдущим поколением [5]. В сетях пятого поколения 5G NR (*аббр. от англ.* New Radio) при надлежащей конфигурации [6] и обработке сигналов PRS [7] в канале с шириной полосы частот 400 МГц и длительностью элементарного временного интервала 2,035 нс, что соответствует частоте дискретизации 491,52 МГц, разрешение ДМ измерений повышалось более, чем на порядок – до значения 0,61 м. Точность первичных измерений по сигналам PRS [8] вместе с геометрическим фактором топологии BS [9] непосредственно определяет итоговую точность оценок координат (ОК) [10], поэтому в сетях пятого поколения сетевое позиционирование метровой точности

стало не только новой технологической возможностью, предоставляемой сетями связи [11], но также и новым инструментом, использование которого позволяет повысить эффективность построения и функционирования самих сетей связи [12].

Помимо традиционных ДМ [13] и разностно-дальномерных (РДМ) методов [14] позиционирования в сотовых сетях подвижной радиосвязи актуален и востребован также и угломерный (УМ) метод ОМП [15]. В сетях LTE данный метод реализуется преимущественно на BS алгоритмами пространственной обработки сигналов (ПОС) [1] посредством многоэлементных антенных решеток (АР) [2]. В сетях 5G NR диапазона миллиметровых волн (ММВ) сбор первичных УМ измерений, в связи с уменьшением длины волны, может осуществляться уже не только на BS, но и на UE, оборудованных многоэлементными АР [16].

Отличительной особенностью корректной работоспособности большинства методов многопозиционной оценки координат в сетях 2G–4G диапазона дециметровых волн (ДМВ) является необходимость предварительного извлечения компоненты луча прямой видимости (LOS, *аббр. от англ.* Line Of Sight) в радиолиниях BS-UE [17]. Данное требование объясняется тем, что в диапазоне ДМВ, по сравнению с диапазоном ММВ сетей 5G, в условиях отсутствия прямой видимости (NLOS, *аббр. от англ.* Non Los) в радиолиниях BS-UE наблюдается достаточное количество (десятки) многолучевых компонент (МЛК), полученных в условиях однократных и многократных переотражений [18], которые затруднительно отличить от лучей LOS, и которые существенно снижают точность ОМП геометрическими методами.

Диапазон ММВ вместе с многоэлементными АР на BS и UE сделали возможной в сетях 5G NR однопозиционную оценку координат и ориентации UE по сигналам единственной радиолинии BS-UE [19]. Высокие потери при распространении радиоволн (РРВ) в диапазоне ММВ в совокупности с использованием высоконаправленных радиолиний BS-UE при работе узкими лучами на передачу и прием привели к тому, что радиоканал BS-UE приобрел новое свойство так называемой разреженности (*от англ.* Sparse Channel), когда число различных МЛК (единицы), по сравнению с диапазоном ДМВ, становится незначительным [20]. В таком сценарии становится возможным различать прямой луч и отдельные однократно отраженные МЛК и обрабатывать их по отдельности [21]; например, при увеличении ширины полосы частот до 400 МГц разрешаемая разность хода лучей достигает 0,75 м. Результаты имитационного моделирования по вторичной обработке первичных ДМ и УМ измерений в канале «вверх» (UE→BS) и «вниз» (BS→UE) [22] подтвердили возможность достижения точности

ОК и ориентации до одного метра и одного градуса, соответственно. В [23] делается интересный вывод о том, что точность позиционирования устройств в диапазоне ММВ определяется, в первую очередь, именно УМ измерениями на BS.

Подтверждением тенденции в пользу однопозиционной оценки координат в сетях 5G является цикл работ [24–26]. В работе [24] впервые вводится понятие *персонального радара* диапазона ММВ. В работе [25] авторы предлагают новые метрики для характеристики точности оценок координат РЕВ (*аббр. от англ. Position Error Bound*) и ориентации ОЕВ (*аббр. от англ. Orientation Error Bound*), полученные из нижней границы Крамера – Рао (НГКР) при условии работы многоэлементными АР на передаче и приеме; устанавливаются их нижние пределы; показано достижение точности до одного метра и до одного градуса, соответственно. В [26] представлены результаты исследования работы передатчика и приемника в режиме пространственного разнесения и диаграммообразования (ДО); в частности, показано, что с точки зрения позиционирования предпочтительным является первый режим работы.

Обзор тенденций развития радиотехнологий в сетях шестого поколения 6G [27–29], существенных для позиционирования UE, позволяет выделить следующие новые факторы:

1) переход из ММВ диапазона в диапазон субмиллиметровых или терагерцовых волн и использование сверхширокополосных радиоканалов с шириной полосы до единиц гигагерц;

2) переход от систем MIMO (*аббр. от англ. Multiple-Input Multiple-Output*) из десятков элементов в составе единой АР к многоантенным системам сверхбольшой размерности UM-MIMO (*аббр. от англ. Ultra Massive MIMO*) из сотен и тысяч элементов с организацией их в АР из составных подмассивов AOSA (*аббр. от англ. Array of Sub Array*);

3) использование реконфигурируемых интеллектуальных поверхностей (RIS, *аббр. от англ. Reconfigurable Intelligent Surface*) в радиолиниях между BS и UE для управления отраженными МЛК в предсказуемом порядке;

4) учет модели кривизны фронта волны (SWM, *аббр. от англ. Spherical Wave Model*) и эффекта ближней зоны (NF, *аббр. от англ. Near Field*) в радиолиниях BS–UE, BS–RIS, RIS–UE.

Развитие технологий радиотехнического позиционирования в сетях B5G (*аббр. от англ. Beyond 5G*) примечательно тем, что если применительно к сетям 5G диапазона ММВ появился термин *персонального радара* [24], то применительно к сетям 6G терагерцового диапазона уже говорят о *конвергенции технологий связи, локализации и зондирования радиозфира* и окружающего пространства радио-

техническими методами и средствами [30]. Появился и новый термин для обозначения таких сетей – интегрированные сети зондирования и связи (ISAC, *аббр. от англ. Integrated Sensing and Communications*) [31]. Для достижения требований к услугам связи, локализации и радиотехнического зондирования в сетях шестого поколения 6G обособливается использование полос частот с шириной до 10 ГГц в диапазоне выше 100 ГГц [32]. Интеграция технологий связи и зондирования в сетях шестого поколения может ожидаемо привести к переосмыслению в академическом и производственном сообществе сценариев использования таких экосистем 6G далеко за пределы их традиционного восприятия [33].

Под руководством зарубежных основоположников сетевого позиционирования Н. Wymeersch и G. Seco-Granados закладываются теоретические основы [34] и сценарии практического применения [35] моделей и методов локализации и зондирования радиотехническими средствами в сетях 6G [36]; формулируются новые направления научных исследований на ближайшую декаду [37].

В связи с чрезвычайно высокими потерями при РРВ в терагерцовом диапазоне и необходимости использования исключительно узких лучей с высокими коэффициентами усиления (КУ) особого внимания заслуживает проблема радиопокрытия в сетях 6G [38, 39]. Новым примером использования инфраструктуры BS сетей шестого поколения является многопозиционное зондирование объектов методами радиотехнического позиционирования [40]. При невозможности построения достоверного геометрического сценария взаимного расположения объекта и субъекта позиционирования предпочтению в сети ISAC отдается метод и средствам машинного и сквозного обучения (*от англ. End-to-End Learning*) [41, 42].

Отдельным научным направлением в рамках построения и функционирования сетей 6G ISAC является так называемая голографическая локализация, суть которой сводится к использованию RIS и эффекта NF для прогноза и частичного управления условиями РРВ с целью повышения точности позиционирования по предсказуемо отраженным МЛК [43–46].

Настоящее исследование посвящено разработке комплекса моделей и методов позиционирования устройств в сетях шестого поколения 6G. В первой части цикла выполняется анализ состояния проблемы и формализуется необходимая математическая модель, на основе которой во второй части выполняется имитационное моделирование и обобщение численных результатов.

Объектом исследования являются интегрированные сети локализации, зондирования и связи

шестого поколения 6G терагерцового диапазона. *Предмет* исследования – модели, методы и алгоритмы позиционирования устройств в сетях шестого поколения 6G терагерцового диапазона. *Методы* исследования: аналитический обзор состояния проблемы по актуальным научным публикациям, концептуальное моделирование, категориальный подход, экспертное комбинирование, формализация, математическое и имитационное моделирование.

Научной проблемой цикла исследований является поиск и обоснование новых радиотехнических средств достижения дециметровой и сантиметровой точности ОК устройств в сетях шестого поколения 6G терагерцового диапазона.

Задачи цикла исследований:

1) обзор методов позиционирования UE при переходе к сетям шестого поколения терагерцового диапазона;

2) формализация математической модели сбора первичных измерений для позиционирования UE в сетях 6G;

3) обзор алгоритмов позиционирования UE при переходе к сетям шестого поколения терагерцового диапазона;

4) формализация математической модели оценки точности позиционирования UE в сетях 6G;

5) оценка точности позиционирования UE в сетях 6G средствами имитационного моделирования, анализ численных результатов;

6) научное обоснование новых радиотехнических средств достижения дециметровой и сантиметровой точности ОК устройств в сетях шестого поколения.

В первой части цикла решаются первые две задачи. Материал исследования организован далее следующим образом. В разделе 2 выполнен анализ состояния проблемы позиционирования в сетях 6G, в частности: формализованы критерии и категории позиционирования в сетях 6G, систематизированы геометрические методы оценки координат и ориентации UE, актуализированы показатели и сценарии позиционирования в сетях 6G, обозначены тенденции и новые технологические особенности позиционирования устройств при переходе от сетей пятого поколения 5G диапазона ММВ к сетям шестого поколения 6G диапазона терагерцовых волн. В разделе 3 представлена математическая модель сбора первичных измерений для позиционирования UE с использованием одной BS и одной RIS по прямым LOS- и отраженным NLOS-сигналам, в частности: формализован геометрический сценарий позиционирования в локальной и глобальной системах координат (СК), систематизирована модель канала и приема MIMO-системы в дальней зоне с единым массивом и составными массивами AP AOSA, предложена модель канала и приема MIMO-системы в

ближней зоне, учитывающая кривизну прихода (COA, *аббр. от англ.* Curvature of Arrival) SWM в условиях NF, эффект расщепления луча (BSE, *аббр. от англ.* Beam Split Effect), а также факторы несовершенства аппаратного обеспечения. В заключении приводится постановка задачи на разработку и программную реализацию имитационной модели позиционирования UE в сетях шестого поколения 6G.

2. Анализ состояния проблемы позиционирования в сетях 6G

В настоящем разделе выполняется анализ состояния проблемы позиционирования устройств в сетях шестого поколения 6G, описываются известные подходы ОМП радиотехническими средствами с акцентом на геометрические модели и методы. Актуализируются ключевые показатели эффективности (KPI, *аббр. от англ.* Key Performance Indicators) и сценарии позиционирования в сетях шестого поколения 6G. Выполняется сопоставительный анализ систем позиционирования по диапазонам частот, а также систематизируются новые технологии позиционирования при переходе от сетей пятого поколения 5G диапазона ММВ к сетям шестого поколения 6G диапазона терагерцовых волн.

2.1. Укрупненные признаки систем позиционирования

Под позиционированием или локализацией обычно понимают ОК и ориентации (для UE с AP) подвижных устройств с помощью одной или нескольких стационарных опорных BS с известными координатами и ориентацией. Позиционирование радиотехническими методами осуществляется по первичным измерениям на основе известных пилотных или опорных сигналов в канале «вверх» радиолинии UE→BS и в канале «вниз» радиолинии BS→UE; в первом случае прием и обработка пилотных сигналов осуществляется на BS, а во втором случае – на UE.

В процессе PPB в радиолинии между UE и BS переданные пилотные или опорные сигналы могут претерпевать потери, отражения, дифракцию и рассеяние. Выполняя прием и обработку пилотных/опорных сигналов известной формы, можно извлечь следующие первичные радиотехнические измерения: время прихода (TOA, *аббр. от англ.* Time of Arrival), разность времен прихода (TDOA, *аббр. от англ.* Time Difference of Arrival), угол прихода/ухода (AOA / AOD, *аббр. от англ.* Angle of Arrival / Departure), разность углов ухода (ADOD, *аббр. от англ.* Angle Difference of Departure). Каждое первичное измерение характеризуется своей геометрической линией положения на плоскости (рисунок 1) и поверхностью положения в пространстве.

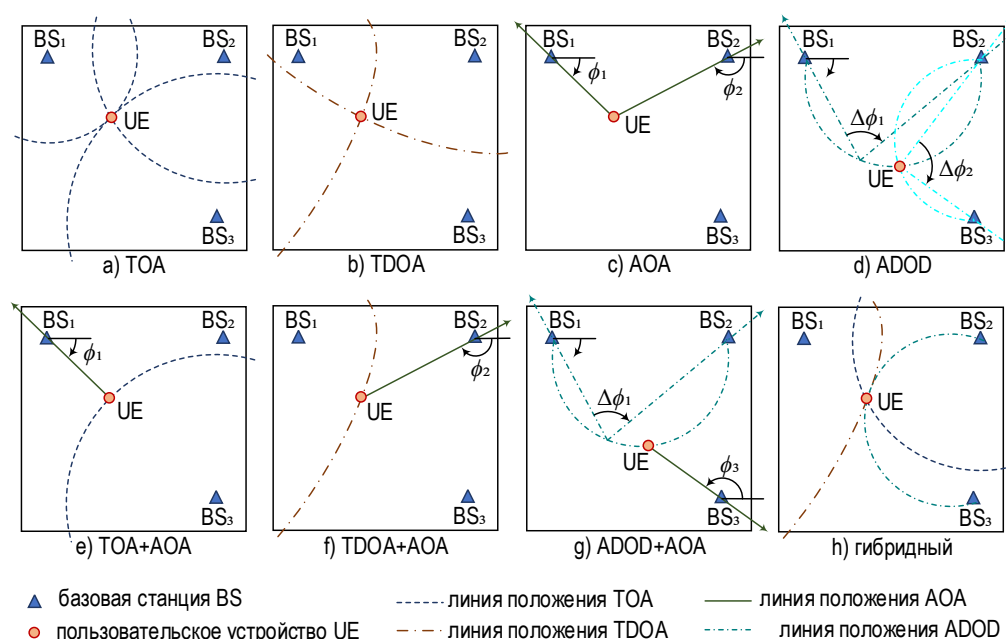


Рис. 1. Геометрические методы позиционирования на плоскости

Fig. 1. Geometric Methods of Positioning on a Plane

Используя вторичную обработку первичных измерений, известные координаты опорных BS и геометрические соотношения между взаимным расположением UE и BS, можно выполнить ОК и ориентации UE. В отличие от моделей и методов позиционирования в сетях предыдущих поколений, в сетях шестого поколения 6G, помимо радиолиний, между BS и UE используются первичные измерения, полученные в радиолиниях с участием RIS, координаты которых также известны. Далее без утраты общности в настоящей работе будем полагать, что сбор и обработка первичных измерений осуществляется в канале «вверх» радиолинии UE→BS, когда UE выступает устройством формирования и передачи, а BS – устройством приема и обработки опорных сигналов, по которым осуществляется сбор первичных ДМ и УМ измерений радиотехническими методами. Таблица 1 содержит укрупненные признаки классификации и соответствующие им примеры систем позиционирования.

2.2. Анализ методов позиционирования

2.2.1. Геометрическое позиционирование и локализация с обучением

Геометрическое позиционирование широко используется в современных сетях связи и навигации и основано на геометрических соотношениях между взаимным расположением UE и BS. Собранные на BS первичные измерения используются как параметры при вторичной обработке. В основе вторичной обработки лежат нелинейные соотношения, связывающие геометрические линии положе-

ния первичных измерений, известные координаты / ориентацию опорных BS и неизвестные координаты / ориентацию UE.

ТАБЛИЦА 1. Укрупненные признаки и примеры систем позиционирования

TABLE 1. General Features and Examples of Positioning Systems

Признак	Пример
Сценарий позиционирования	снаружи помещений, внутри помещений
Радиотехническая система	ГНСС, сотовые сети 2G-5G, B5G/6G, сети радиодоступа WiFi
Метод позиционирования	на основе геометрических соотношений, на основе машинного обучения
Первичные сигналы	радиоизлучение, оптическое (видимое) излучение
Тип радиолокации	активная (передача и прием), пассивная (прием)
Архитектура системы	централизованная, распределенная, кластерная
Оцениваемые координаты	абсолютные, относительные
Взаимодействие субъектов	кооперативное / некооперативное позиционирование

Рисунок 1 иллюстрирует известные геометрические методы позиционирования UE радиотехническими методами с использованием опорных BS с известными координатами. Нелинейное соотношение геометрического метода позиционирования формализуется целевой функцией, оптимизируя которую, оценивают неизвестные координаты и ориентацию UE на плоскости или в пространстве.

Данные методы основаны преимущественно на предположении о сборе первичных ДМ и УМ измерений в условиях LOS в радиоперелиниях между BS, UE и RIS. Преимуществом геометрических методов позиционирования является возможность установления теоретических пределов точности их работы аналитически.

В более сложных условиях PPB при NLOS в радиоперелиниях между BS, UE и RIS смоделировать геометрический сценарий взаимного расположения субъектов BS, RIS и объекта позиционирования UE становится уже затруднительно. В таких сценариях находит применение позиционирование или локализация с машинным обучением (ML, аббр. от англ. Machine Learning). Алгоритмы ML основаны на предварительном оффлайн-обучении. Примером использования данного подхода в задачах позиционирования является обработка уровней принимаемых сигналов (RSSI, аббр. от англ. Received Signal Strength Indicator) от точек доступа беспроводных локальных сетей (WLAN, аббр. от англ. Wireless Local Area Network) и построение баз данных радиокарт [47]. Данный этап измерений выполняется заблаговременно (оффлайн); при работе онлайн выполняется измерение текущих принимаемых значений RSSI и сопоставление их с заранее записанными в базе данных, на основе чего выполняется ОМП по принципу наибольшего сходства.

В сетях шестого поколения 6G терагерцового диапазона радиоканал вследствие высоких потерь PPB и использования высоконаправленных радиоперелиний на передаче и приеме характеризуется разряженностью, когда различными оказываются несколько однократно отраженных МЛК в условиях NLOS, а также – LOS. Поэтому в сетях 6G наибольший интерес представляют модели и методы геометрического позиционирования по прямым и однократно отраженным лучам.

2.2.2. Оценка координат и ориентации геометрическими методами

Различают ОК на плоскости (позиционирование в 2D) и в пространстве (позиционирование в 3D). Рисунок 1 иллюстрирует геометрические методы позиционирования на плоскости при сборе измерений в канале «вверх» (UL, аббр. от англ. Uplink), когда UE работает на передачу, а BS работают на прием. С точки зрения вторичной обработки первичных измерений времени и угла прихода сигнала различают четыре геометрических метода позиционирования: дальномерный TOA (см. рисунок 1a), разностно-дальномерный TDOA (см. рисунок 1b), угломерный AOA (см. рисунок 1c), разностно-угломерный ADOD (см. рисунок 1d). Остальные геометрические методы представляют их комбинации:

- ДМ–УМ (TOA + AOA) (см. рисунок 1e);
- РДМ–УМ (TDOA + AOA) (см. рисунок 1f);
- РУМ–УМ (ADOD + AOA) (см. рисунок 1g);
- гибридный (TOA + TDOA + ADOD) (рисунок 1h).

ДМ и РДМ методы основаны на измерении задержки (от англ. Laturation) и реализуются алгоритмами трилатерации, а УМ методы основаны на измерении углов (от англ. Angulation) и реализуются алгоритмами триангуляции.

Для практической реализации ДМ метода TOA (см. рисунок 1a) требуется синхронизация передатчика и приемника. При отсутствии синхронизации BS–UE задержку передачи-приема TOA можно оценить альтернативным подходом – измерением РТТ.

Для практической реализации РДМ метода TDOA (см. рисунок 1b) измеряются разности времен прихода сигналов, поэтому достаточно синхронизировать только BS.

Для практической реализации УМ метода AOA (см. рисунок 1c) требуется оборудовать BS антенными решетками, на которых возможно измерение угла прихода сигнала по линии пеленга AOA в направлении на UE. При этом различают одномерное пеленгование 1D AOA на плоскости, когда измеряется только азимут, и двумерное пеленгование 2D AOA в пространстве, когда измеряется азимут и угол места. В обоих случаях линиями положения являются линии пеленга, на пересечении которых находится позиционируемое UE.

Для практической реализации УМ метода AOD требуется оборудовать UE антенными решетками, на которых возможно измерение угла ухода сигнала по пеленгу AOD в направлении заданной BS. Однако вследствие неизвестной ориентации UE на плоскости и в пространстве для ОК интерес представляют не абсолютные, а относительные измерения углов ухода ADOD (см. рисунок 1d); при этом измерения AOD служат для оценки ориентации UE.

В идеализированном сценарии при отсутствии погрешностей первичных измерений и неточностей их вторичной обработки в результате пересечения линий или поверхностей положения получается точка, которая является искомыми координатами позиционируемого UE. Однако на практике первичные измерения характеризуются погрешностями различного происхождения, поэтому в результате пересечения линий или поверхностей положения получается не отдельная точка, а область неопределенности на плоскости или в пространстве, внутри которой робастные алгоритмы вторичной обработки должны находить ОК устройств с наименьшей погрешностью.

Таблица 2 систематизирует требования к геометрическим методам позиционирования.

ТАБЛИЦА 2. Требования к геометрическим методам позиционирования

TABLE 2. Requirements for Geometric Positioning Methods

Геометрические методы позиционирования / требования		TOA	TDOA	AOA	ADOD
Требуемое число BS на одно измерение	–	1	2	1	2
Геометрическая линия или поверхность положения	2D	окружность	гипербола	линия	дуга
	3D	сфера	гиперболоид	линия	поверхность вращения дуги
Требуемое число измерений для оценки координат	2D	3	3	2	3
	3D	4	4	2	4
Требуемое число измерений AOD для оценки ориентации	1D	1, необходима оценка координат в 2D			
	2D	1, необходима оценка координат в 3D			
	3D	2, необходима оценка координат в 2D			
Системные требования	–	синхронизация BS–UE, RTT	синхронизация BS–BS	антенная решетка на BS	антенная решетка на UE

Оценка ориентации UE представляет интерес для решения задач диаграммообразования AP устройства при работе на передачу и прием. Требованием для оценки ориентации UE является измерение углов ухода AOD на стороне UE. По числу измеряемых углов Эйлера выделяют оценку ориентации в сценариях 1D, 2D и 3D, где измеряются углы $[\alpha]$, $[\alpha, \beta]$ и $[\alpha, \beta, \gamma]$, соответственно. Соотношения между углами Эйлера α, β, γ и матрицами поворота в локальной и глобальной СК описывается далее.

Оценка ориентации в 1D (оценка угла Эйлера $[\alpha]$) выполняется при позиционировании на плоскости с использованием линейной эквидистантной AP (ULA, аббр. от англ. Uniform Linear Array) при известных координатах в 2D. Например, объект, осуществляющий движение на плоскости, может оценить свою ориентацию в 1D вместе с оценкой координат посредством измерения углов ухода AOD.

Оценка ориентации в 2D (оценка углов Эйлера $[\alpha, \beta]$) выполняется при позиционировании в пространстве при известных координатах в 3D и требует измерения по меньшей мере одной пары углов ухода AOD (азимута и угла места).

Оценка ориентации в 3D (оценка углов Эйлера $[\alpha, \beta, \gamma]$) выполняется при позиционировании в пространстве при известных координатах в 3D и требует измерения по меньшей мере двух пар углов ухода AOD (азимута и угла места). Вычисление углов Эйлера $[\alpha, \beta, \gamma]$ осуществляется методами оптимизации, например, методом наименьших квадратов (МНК) или методом максимального правдоподобия.

2.3. Анализ показателей и сценариев позиционирования

2.3.1. Показатели эффективности

При построении и функционировании систем, а также при планировании услуг сетевого позиционирования следует учитывать следующие KPI технологий ОМП в сетях 6G [1–3].

Точность оценок координат и ориентации UE является ключевым показателем технологий сетевого позиционирования 6G. Измеряется корнем из среднеквадратической ошибки ОК или интегральной функцией распределения ошибок ОК. При заданной модели сигнала и отношении сигнал/шум (ОСШ) общепринятой оценкой пределов точности позиционирования является НГКР дисперсии погрешности ОК. Аналитическая оценка НГКР является дополнительным источником информации для поиска и обоснования наиболее эффективного в заданных условиях алгоритма позиционирования.

Радиопокрытие заданной области радионавигационным полем из опорных и/или других сигналов сетей 6G, по которым решается задача ОМП, также является ключевым показателем технологий сетевого позиционирования 6G. Вследствие высоких потерь радиосигналов в диапазоне ММВ сетей 5G и терагерцовых волн сетей 6G по сравнению с потерями в диапазоне ДМВ сетей 4G, зоны радиопокрытия шестого поколения для решения задач позиционирования с заданной точностью следует планировать заблаговременно, например, с использованием цифровых моделей местности. Радиопокрытие измеряется обычно территорией на плоскости в 2D или объемом пространства в 3D.

Задержка в предоставлении услуги позиционирования определяется как интервал времени между отправкой UE запроса на ОМП и получением от сети ответа с ОК. При использовании сети PRS задержка определяется периодичностью их передачи и временем их обработки [6, 7].

Обновление оценок координат в сценариях позиционирования подвижных UE или транспортных средств определяется как величина, обратная времени задержки получения и обновления ОК в сценариях отслеживания подвижных UE и измеряется частотой обновления [48, 49].

Стабильность предоставления услуги позиционирования можно оценить как снижение точности

ОК на интервале времени так называемой «глухоты», которая является следствием кратковременной утраты взаимной ориентации лучей в высоконаправленных радиополосах между стационарной BS и подвижным UE [50–52].

Масштабируемость предоставления услуги позиционирования определяется как способность системы сетевого позиционирования 6G адаптироваться к решению задачи ОМП с использованием инфраструктуры BS при увеличении числа UE.

Мобильность позиционируемых подвижных UE ограничивается влиянием эффекта Доплера и определяется допустимой скоростью устройств, при которой услуга позиционирования оказывается с требуемыми KPI.

Сложность системы сетевого позиционирования 6G определяется объемом необходимого для реализации услуг ОМП аппаратного и программного обеспечения. Под аппаратным обеспечением

понимают, в первую очередь, оборудование инфраструктуры BS, RIS и UE [53]. Настоящее исследование посвящено особенностям программной реализации алгоритмов ОК и ориентации UE в заданных сценариях.

2.3.2. Сценарии позиционирования

В сетях шестого поколения 6G выделяют следующие сценарии: телехирургия, голография, связь транспортных средств, цифровые двойники, кооперативные роботы, локализация угроз. Во всех перечисленных сценариях востребовано высокоточное радиотехническое UE. Таблица 3 систематизирует ожидаемые KPI и сценарии позиционирования в сетях 6G [53].

Анализ показателей говорит о стремлении к достижению дециметровой и сантиметровой точности ОК в сетях 5G и 6G, соответственно.

ТАБЛИЦА 3. Показатели эффективности и сценарии позиционирования в сетях 6G

TABLE 3. Performance Indicators and Positioning Scenarios in 6G Networks

Показатели эффективности / сценарии позиционирования 6G	Точность ОК	Точность ориентации, °	Радиопокрытие, м	Задержка, мс	Частота обновления, Гц	Масштабируемость, 1/км ²	Мобильность, км/ч
5G KPI	10 см	10	–	1	–	10 ⁶	500
6G KPI	1 см	1	–	~ 0,1	–	10 ⁷	1000
Телехирургия	< 1 мм	< 1	–	–	≥ 10	–	–
Голография	< 1 см	1	10	5	≥ 100	–	–
Связь ТС	< 1 м	–	> 100	1	1–10	–	–
Цифровые двойники	1 см	–	~ 130	0,1–1	≥ 10	–	–
Кооперативные роботы	< 1 см	< 1	–	1	≥ 10	–	–
Локализация угроз	1 см	–	~ 130	1	≥ 10	–	–

2.4. Анализ технологий позиционирования

2.4.1. Анализ технологий позиционирования по диапазонам частот

Формализуем технологии и средства радиотехнического позиционирования в сетях пятого 5G и шестого 6G поколений. В зависимости от используемого спектра электромагнитного излучения различают технологии сетевого позиционирования, работающие в четырех укрупненных диапазонах (таблица 4).

Диапазон ультракоротких волн (УКВ) – до 30 ГГц, традиционный для подвижной связи CRF (Conventional Radio Frequency), включает ДМВ и сантиметровые (СМВ) волны.

Диапазон миллиметровых волн – от 30 до 100 ГГц получил распространение в сетях 5G.

Диапазон субмиллиметровых или терагерцевых волн – от 0,3–3 ТГц получил распространение в сетях шестого поколения 6G.

Диапазон видимого света (VL, аббр. от англ. Visible Light) – электромагнитное излучение, воспринимаемое человеческим глазом от 400 до 770 ТГц.

ТАБЛИЦА 4. Категории систем позиционирования по диапазону частот

TABLE 4. Categories of Positioning Systems by Frequency Range

Обозначение	Диапазон	Системы	Особенности
УКВ (CRF)	0,3 – 30 ГГц	GNSS, 2G GSM, 3G UMTS, 4G LTE, WiFi	многопозиционная ОК преимущественно в условиях LOS
ММВ (mmWave)	30 – 300 ГГц	5G NR	однопозиционная ОК в условиях LOS/NLOS
ТГц (THz)	0,3 – 3 ТГц	6G	однопозиционная ОК в условиях LOS/NLOS с RIS в ближней зоне
Свет (VL)	400 – 790 ТГц	LIDAR	обособленные приборы, не являющиеся составной частью системы связи

В диапазоне УКВ работают глобальные навигационные спутниковые системы ГНСС и ГЛОНАСС. Спутниковые системы решают задачи позиционирования преимущественно снаружи помещений и могут достигать метровой точности при передаче ассистирующей информации для приемника ГНСС по протоколу SUPL (аббр. от англ. Secure User Plane Location) в сетях подвижной радиосвязи, начиная с LTE [1–3]. Внутри помещений условия LOS с UE нарушаются, и системы ГНСС перестают работать с ожидаемой точностью. Для позиционирования внутри помещений получили распространение беспроводные локальные сети семейства стандартов WiFi.

В диапазоне ММВ работают сети пятого поколения 5G NR, использующие многоантенные системы MIMO с гибридной архитектурой. Увеличение несущей частоты вместе с ростом ширины полосы сигнала с десятков (LTE) до сотен (NR) мегагерц позволяет выполнять разрешение МЛК и, таким образом, повышать точность позиционирования в условиях NLOS. Также отличительной особенностью ИМП устройства в сетях 5G является возможность решения задачи позиционирования однопозиционным методом с использованием одной BS.

В диапазоне VL работают сверхширокополосные системы связи VLC (аббр. от англ. Visible Light Communication). В качестве приемопередатчиков используются лазерные диоды и светодиоды. Лазерные диоды обеспечивают работу в широкой полосе частот на достаточно больших расстояниях; препятствием в их использовании является необходимость точного выравнивания направленный линий связи. Пример практического использования – лидары (LIDAR, аббр. от англ. Light Detection and Ranging), которые, однако, работают как обособленные приборы и не являются составной частью какой-либо системы связи. Светодиоды используются для освещения и обеспечивают достаточное световое покрытие, но чрезмерно чувствительны к окружающей среде [53].

2.4.2. Анализ особенностей позиционирования

Технологии сетевого позиционирования шестого поколения 6G терагерцового диапазона можно рассматривать как развитие и дополнение технологий сетевого позиционирования 5G миллиметрового диапазона. Выполним далее сравнение технологий позиционирования в сетях 5G и 6G. Таблица 5 систематизирует факторы, достоинства и недостатки технологий позиционирования, согласно обзору [53].

Накладные расходы на установление и поддержание направленного соединения обычно выше, чем для соединения устройств, которые работают с ненаправленными антеннами. Под накладными

расходами понимаются кадры, передача и прием которых не связаны с полезным информационным обменом. Применительно к управлению лучами – это передача и прием кадров при изменении ориентации лучей на передаче и приеме до тех пор, пока они не будут выровнены по критерию максимального уровня принимаемого сигнала [50–52]. Установление двунаправленного соединения реализуется сканированием лучами на передаче и приеме по азимуту и углу места, в результате чего лучи в радиолинии оказываются сориентированными друг в друга. При реализации процедуры сканирования следует соблюдать компромисс между скоростью и точностью выравнивания лучей: чем шире лучи на передаче и приеме, тем быстрее они окажутся сориентированными друг в друга; соответственно, чем уже лучи, тем дольше длится перебор пространства по азимуту и углу места и тем выше накладные расходы на установление такого двунаправленного соединения. В отечественной терминологии такой подход иногда называют амплитудной пеленгацией.

Конфигурация сигнала позиционирования в OFDM-радиоинтерфейсе задается целым набором параметров даже внутри кадра: число поднесущих в домене частот, число временных интервалов в домене времени. Таковы, например, выделенные PRS, подробно исследованные в [6, 7]. Чем больше PRS занимают частотно-временного ресурса в OFDM-сигнале, тем меньше ресурсных блоков остается на передачу полезной информации и, соответственно, выше накладные расходы. Помимо конфигурации внутри кадра, PRS характеризуются и периодичностью передачи в последовательности кадров, что также влияет и на задержку передачи данных, и на частоту обновления ОК.

Переход из традиционного для сетей подвижной радиосвязи диапазона УКВ в ММВ и ТГц-диапазон знаменует существенным увеличением несущей частоты f , а также занимаемой ширины полосы W радиосигнала. Рост несущей частоты приводит к пропорциональному уменьшению длины волны, что, в свою очередь, позволяет: снизить геометрические размеры АР с сохранением числа ее элементов, или увеличить размерность решетки с сохранением ее геометрических размеров. С одной стороны, недостатком увеличения несущей частоты при переходе из ММВ в ТГц-диапазон является резкое увеличение потерь при РРВ. С другой стороны, преимуществом увеличения потерь при РРВ является снижение числа разрешимых в домене задержки МЛК, которые обычно снижают точность позиционирования. Сама возможность разрешения МЛК возрастает с увеличением ширины полосы сигнала. Так, например, при ширине полосы 100 МГц разрешение по расстоянию составит 3 м, а при ширине полосы 1 ГГц – уже 0,3 м.

ТАБЛИЦА 5. Анализ факторов и особенностей позиционирования в сетях 5G и 6G

TABLE 5. Analysis of Factors and Features of Positioning in 5G and 6G Networks

Фактор	5G, ММВ	Достоинство	Недостаток	6G, ТГц	Достоинство	Недостаток
Несущая частота, длина волны	30–100 ГГц, 3–10 мм	небольшое число МЛК	потери РРВ пропорциональны $10^4 d$	0,1–10 ТГц, 0,03–3 мм	миниатюризация	потери РРВ пропорциональны $10^5 - 10^7 d$
Ширина полосы частот	400 МГц	ДМ разрешение $\sim 0,75$ м	высокая частота дискретизации	1–10 ГГц	ДМ разрешение 3–30 см	высокое энергопотребление АЦП, большой объем данных
Размер AP	10×10 элементов	угловое разрешение $\sim 10^\circ$	средние накладные расходы на управление лучом	100×100 элементов	угловое разрешение $< 1^\circ$, без МЛК в луче	высокие накладные расходы на управление лучом
Тип AP	гибридная планарная решетка UPA	компромиссное построение тракта TX / RX, измерение азимута и угла места	скорость сканирования пропорционально ширине луча	составной массив решетки AOSA	приемлемая реализация, измерение азимута и угла места, УМ ОМП без ДМ	скорость сканирования пропорциональна ширине луча, нужна калибровка AP
Несовершенство аппаратуры	Квантование в PS, АЦП	снижение сложности аппаратной реализации и энергопотребления	потеря точности ОК и ориентации	IQI, PN, PA, ADC	возможность использовать эффекты для уточнения ОК и ориентации	ограничение мощности, невозможность использования типовых DSP
Требования к синхронизации	< 10 нс	требуется для ДМ и РДМ методов	сложно поддерживать	< 1 нс	требуется для ДМ и РДМ методов	чрезвычайно сложно реализовать и поддерживать
Форма используемого сигнала	OFDM	компенсация МЛК, структурированность сигнала	нелинейные искажения вследствие высокого пик-фактора	–	форма сигнала не принципиальна	DSP и определяется формой сигнала
Эффекты РРВ	небольшое число кластеров	небольшое число NLOS-компонент, возможность извлечь луч LOS	снижение точности при блокировании луча LOS	небольшое число компонент, расщепление луча BSE, ближняя зона NF	малое число компонент, возможность использования эффектов BSE и NF	использование эффектов BSE и NF потребует новых DSP
Метод позиционирования	многопозиционная и однопозиционная ОК	простота и масштабируемость	нужна синхронизация и координация нескольких BS	многопозиционная и однопозиционная ОК с участием RIS	снижение требований к синхронизации за счет только УМ	новая инфраструктура RIS, требуется калибровка RIS
Конфигурация сигнала позиционирования	DL-PRS, AOD	широковещательные сигналы, комбинирование ДМ и УМ измерений	измерение углов по признаку мощности	–	повышенная точность ОК и ориентации	увеличение накладных расходов и задержки

Примечание: АЦП – аналого-цифровой преобразователь; UPA (аббр. от англ. Uniform Planar Array) – планарная AP

Увеличение f и W усложняют аппаратную реализацию приемопередатчиков и приводят к необходимости учета несовершенства аппаратного обеспечения, влияющего на точность позиционирования. Более того, отработанная в сетях LTE и NR форма OFDM-сигнала может оказаться неподходящей для широкополосных сигналов терагерцового диапазона вследствие известной проблемы высокого пик-фактора. Альтернативным вариантом яв-

ляются OFDM-сигналы с дискретным преобразованием Фурье, используя которые передатчик модулирует все поднесущие одинаковыми данными, что снижает пик-фактор, сохраняя устойчивость к обработке МЛК. При построении математической и имитационной моделей радиоканала MIMO терагерцового диапазона необходимо учитывать эффекты SWM в условиях NF, а также эффект расщепления луча BSE при широкой полосе частот W в сравнении с несущей частотой f .

В системах радиосвязи использование реконфигурируемых интеллектуальных поверхностей RIS позволяет повысить отношение сигнал/шум и, таким образом, увеличить пропускную способность. В задачах ОМП RIS позволяет повысить точность ОК и ориентации UE за счет того, что является собой (помимо BS в однопозиционном сценарии) дополнительный опорный пассивный субъект геометрического позиционирования с известными координатами, который реализует сбор вспомогательных первичных измерений об отраженных компонентах, что оказывается особенно информативно при учете кривизны прихода фронта проходящей волны COA в условиях NF. Функционирование RIS можно рассматривать как работу пассивной опорной BS, выполняющей переотражение МЛК предсказуемым образом. RIS не требуют синхронизации, удовлетворяют требованиям по минимальному энергопотреблению на управление отражающими свойствами и при достаточно большом по сравнению с длиной волны геометрическом размере могут формировать условия NF.

В отличие технологий многопозиционного сетевого позиционирования в сетях четвертого поколения 4G LTE диапазона УКВ, основанных преимущественно на обработке лучей LOS с участием не менее четырех BS, в сетях пятого 5G и последующих B5G поколений диапазона ММВ и терагерцовых волн получает распространение однопозиционная ОК устройства с участием единственной BS. Это становится возможным за счет свойства разряжен-

ности радиоканала, учета однократно отраженных МЛК в условиях NLOS, а также при содействии RIS. В частности, разрешение отдельных МЛК в домене времени становится возможным за счет увеличения ширины полосы сигнала до единиц ГГц.

3. Математическая модель сбора первичных измерений для позиционирования

В настоящем разделе выполняется математическая формализация геометрического сценария позиционирования в сетях 6G, описывающая взаимное расположение объекта и субъектов позиционирования, а также соответствующие им пространственные вектора в глобальной и локальной СК. Затем приводится обобщенная модель канала и приема МIMO-системы в дальней зоне, включающая условия наличия и отсутствия LOS, а также радиолинию с участием RIS. В заключении выполняется уточнение модели канала и приема МIMO-системы в NF с учетом эффекта расщепления луча и несовершенства аппаратного обеспечения. При подготовке раздела использованы материалы из [53].

3.1. Геометрический сценарий позиционирования

3.1.1. Параметры и обозначения

Формализуем математическую модель сбора первичных ДМ и УМ измерений при позиционировании UE с использованием BS и RIS. Рисунок 2 иллюстрирует сценарий позиционирования.

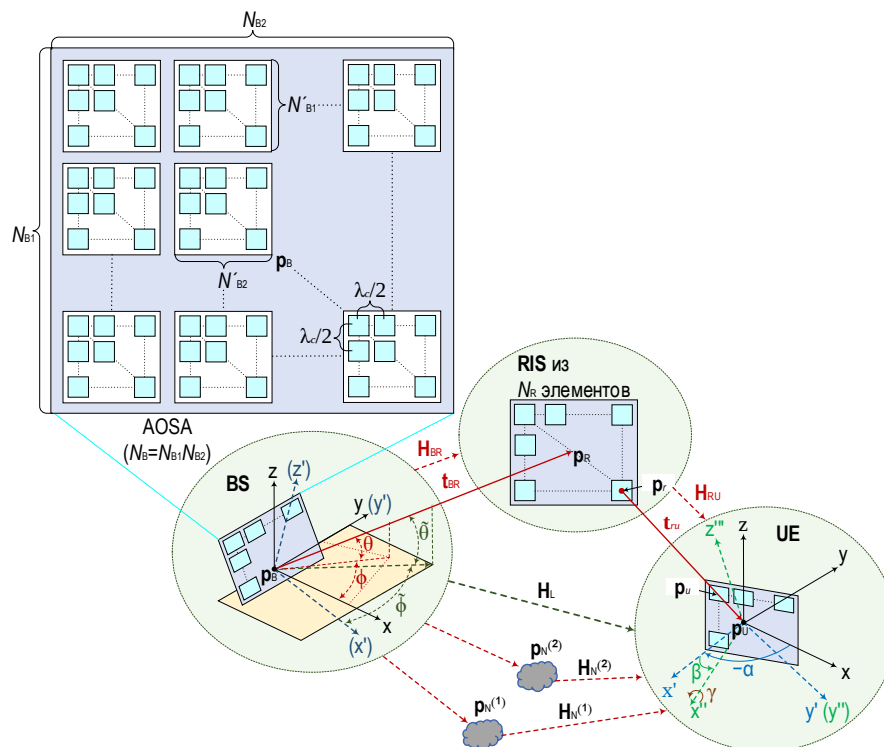


Рис. 2. Сценарий позиционирования в сетях 6G

Fig. 2. Positioning Scenario in 6G Networks

Отличительной особенностью позиционирования в сетях 6G является использование на BS составных массивов SA (аббр. от англ. Sub Array) AP, получившей название AOSA. Сценарий позиционирования 6G на рисунке 2 можно рассматривать как многоантенную систему сверхвысокой размерности UM-MIMO, где BS, UE и RIS оборудованы решетками. Под глобальными координатами BS, UE и RIS далее понимается геометрический центр их AP. Таблица 6 содержит параметры математической модели позиционирования UE в сетях 6G [53].

ТАБЛИЦА 6. Параметры математической модели позиционирования в сетях 6G

TABLE 6. Parameters of the Mathematical Model of Positioning in 6G Networks

Обозначение	Описание
$\mathbf{p}_Q$	координаты BS, RIS, UE, NLOS, $Q \in \{B, R, U, N\}$
$\mathbf{p}_q$	глобальные координаты q – го элемента AP, $q \in \{b, r, u\}$
$\tilde{\mathbf{p}}_q$	локальные координаты q – го элемента AP, $q \in \{b, r, u\}$
N_Q	число элементов AP BS, RIS, UE, $Q \in \{B, R, U\}$
N'_Q	число элементов составного SA массива AOSA, $Q \in \{B, R, U\}$
L_N	число МЛК в условиях NLOS
$\mathbf{R}_Q$	матрица поворота BS, RIS, UE, $Q \in \{B, R, U\}$
$\mathbf{o}_Q$	вектор ориентации по углам Эйлера BS, RIS, UE, $Q \in \{B, R, U\}$
$\mathbf{t}$	вектор направления в глобальной СК
$\tilde{\mathbf{t}}$	вектор направления в локальной СК
Φ	углы направления в глобальной СК
$\tilde{\Phi}$	углы направления в локальной СК
Φ'	углы направления луча (ДО)
$\mathbf{H}$	канальная матрица традиционной MIMO системы
$\mathcal{H}$	канальная матрица MIMO системы составных массивов AOSA
$\rho e^{-j\kappa}$	комплексный коэффициент канала каждой МЛК
G_B	коэффициент усиления (КУ) BS
G_U	КУ UE
K	число поднесущих (для пилот сигналов)
W	ширина полосы частот
$\mathcal{G}$	число передач
B	сдвиг синхронизации
$\mathcal{A}$	коэффициент составного массива AP
K_a	коэффициент ослабления МЛК
K_N	коэффициент отражения МЛК в условиях NLOS
$\mathbf{s}$	вектор состояния
γ	вектор измерений

3.1.2. Соотношения в глобальной системе координат

Обозначим центр AP базовой станции B , реконфигурируемой интеллектуальной поверхности R или пользовательского устройства U как:

$$\mathbf{p}_Q = [x_Q, y_Q, z_Q]^T, \quad (1)$$

где Q может принимать значения $Q \in \{B, R, U\}$, а число элементов AP равно N_Q .

Под элементом далее понимается минимальный субъект радиосвязи, будь то отдельный элемент единой традиционной AP, составной массив SA в составе решетки из составных массивов AOSA, или же элемент RIS. Обозначим через $\mathbf{p}_b, \mathbf{p}_r, \mathbf{p}_u$ местоположение в глобальной СК отдельного элемента BS, RIS и UE. Сделаем далее допущение о том, что в сценарии передачи-приема (см. рисунок 2) наблюдается N_L МЛК в условиях NLOS, где l -й с однократным отражением – отражатель или рассеиватель с неизвестными координатами $\mathbf{p}_N^{(l)}$ в глобальной СК.

3.1.2. Соотношения в локальной системе координат

Для планарной (UPA) или прямоугольной (URA, аббр. от англ. Uniform Rectangular Array) эквидистантной AP определим ее центр как начало координат в локальной СК, а нормаль к поверхности решетки в ее центре, т. е. прямая, перпендикулярная к касательной плоскости в указанной точке поверхности, определяет ось X . Таким образом, прямоугольная AP лежит в плоскости YZ .

Для $Q \in \{B, R, U\}$ вектор углов Эйлера, описывающих поворот BS, RIS и UE в трехмерном 3D евклидовом пространстве, обозначим выражением:

$$\mathbf{o}_Q = [\alpha_Q, \beta_Q, \gamma_Q]^T, \quad \alpha_Q \in (-\pi, \pi], \quad \beta_Q \in (-\pi/2, \pi/2], \quad \gamma_Q \in (-\pi, \pi]. \quad (2)$$

где сам поворот вокруг осей Z, Y, X используется для описания ориентации AP в глобальной СК.

Обозначим через $\tilde{\mathbf{p}}_q$ координаты q -го элемента решетки BS, RIS и UE в локальной СК, где $q \in \{b, r, u\}$, тогда соотношение между глобальными $\mathbf{p}_q$ и локальными $\tilde{\mathbf{p}}_q$ координатами q -го элемента AP в трехмерном 3D-пространстве определяется как:

$$\mathbf{p}_q = \mathbf{R}_Q \tilde{\mathbf{p}}_q + \mathbf{p}_Q; \quad (3)$$

где $\mathbf{p}_Q$ – центр решетки BS, RIS и UE, в составе которой находится q -й элемент, $\mathbf{R}_Q$ – матрица поворота, получаемая из вектора поворота $\mathbf{o}_Q$ (2) [55]:

$$\mathbf{R}_Q = \begin{bmatrix} c_\alpha c_\beta & c_\alpha s_\beta s_\gamma - c_\gamma s_\alpha & s_\alpha s_\gamma + c_\alpha c_\gamma s_\beta \\ c_\beta s_\alpha & c_\alpha c_\gamma + s_\alpha s_\beta s_\gamma & c_\gamma s_\alpha s_\beta - c_\alpha s_\gamma \\ -s_\beta & c_\beta s_\gamma & c_\beta c_\gamma \end{bmatrix}; \quad (4)$$

где использованы следующие обозначения:

$$c_\alpha = \cos(\alpha_Q), \quad s_\alpha = \sin(\alpha_Q), \quad (5)$$

$$c_\beta = \cos(\beta_Q), s_\beta = \sin(\beta_Q), \quad (6)$$

$$c_\gamma = \cos(\gamma_Q), s_\gamma = \sin(\gamma_Q). \quad (7)$$

Используя обратную матрицу поворота $\mathbf{R}_Q^{-1}$ от (4), можно записать выражение для локальных координат q -го элемента через глобальные координаты $\mathbf{p}_Q$ решетки BS, RIS и UE, в состав которой входит q -й элемент:

$$\tilde{\mathbf{p}}_q = \mathbf{R}_Q^{-1}(\mathbf{p}_q - \mathbf{p}_Q). \quad (8)$$

3.1.3. Вектор направления прихода и ухода сигналов

Рассмотрим вектор направления углов прихода / ухода AOA / AOD сигналов.

Допустим, UE с глобальными координатами $\mathbf{p}_U = [x_U, y_U, z_U]^T$ ведет передачу на BS с глобальными координатами $\mathbf{p}_B = [x_B, y_B, z_B]^T$; тогда расстояние между центрами решеток BS и UE можно представить выражением:

$$d_{BU} = \|\mathbf{p}_U - \mathbf{p}_B\|, \quad (9)$$

где $\|\mathbf{p}_U - \mathbf{p}_B\|$ – евклидова норма вектора направления:

$$\mathbf{p}_U - \mathbf{p}_B = [x_U - x_B, y_U - y_B, z_U - z_B]^T, \quad (10)$$

которая является геометрическим расстоянием между точками $\mathbf{p}_U$ и $\mathbf{p}_B$ в евклидовом пространстве и вычисляется как:

$$\begin{aligned} \|\mathbf{p}_U - \mathbf{p}_B\| &= \\ &= \sqrt{(x_U - x_B)^2 + (y_U - y_B)^2 + (z_U - z_B)^2}. \end{aligned} \quad (11)$$

Вектор направления $\mathbf{t}_{BU}$ в глобальной СК от BS к UE нормируется расстоянием d_{BU} и определяется по выражению:

$$\mathbf{t}_{BU} = -\mathbf{t}_{UB} = \begin{bmatrix} t_{BU,x} \\ t_{BU,y} \\ t_{BU,z} \end{bmatrix} = \frac{\mathbf{p}_U - \mathbf{p}_B}{d_{BU}}, \quad (12)$$

где $\mathbf{t}_{UB}$ – вектор направления в глобальной СК от UE к BS.

Представим вектор направления в локальной СК $\tilde{\mathbf{t}}_{BU}$ выражением:

$$\tilde{\mathbf{t}}_{BU} = \begin{bmatrix} \tilde{t}_{BU,x} \\ \tilde{t}_{BU,y} \\ \tilde{t}_{BU,z} \end{bmatrix} = \frac{\tilde{\mathbf{p}}_u - \tilde{\mathbf{p}}_b}{d_{bu}}, \quad (13)$$

тогда с учетом описанного порядка поворота вектор направления в локальной СК $\tilde{\mathbf{t}}_{BU}$ можно представить через вектор направления в глобальной СК $\mathbf{t}_{BU}$ как:

$$\tilde{\mathbf{t}}_{BU} = \mathbf{R}_B^{-1} \mathbf{t}_{BU} = \mathbf{R}_B^T \mathbf{t}_{BU}. \quad (14)$$

Направление прихода / ухода AOA / AOD сигналов определяется парой углов (см. рисунок 2):

– азимут $\phi \in (\pi, \pi]$ – угол между проекцией вектора $\mathbf{t}$ на плоскость XY и осью Y ;

– угол места $\theta \in [-\pi/2, \pi/2]$ – между вектором $\mathbf{t}$ и плоскостью XY .

Данные пары углов, задающих направления AOA / AOD, могут определяться как в глобальной, так и в локальной СК, но измерения самих углов могут осуществляться исключительно в локальной СК AP с использованием известных алгоритмов оценки направления прихода сигнала [1–3].

Обозначим направление прихода / ухода AOA / AOD в глобальной СК как:

$$\boldsymbol{\varphi} = [\phi, \theta]^T, \quad (15)$$

а направление прихода / ухода AOA / AOD в локальной СК выражением:

$$\tilde{\boldsymbol{\varphi}} = [\tilde{\phi}, \tilde{\theta}]^T. \quad (16)$$

Из тригонометрических определений азимута $\tilde{\phi}_{BU}$ и угла места $\tilde{\theta}_{BU}$, вектор $\tilde{\mathbf{t}}_{BU}$ (14) направления $\tilde{\boldsymbol{\varphi}}_{BU}$ в локальной СК можно представить выражением:

$$\tilde{\mathbf{t}}_{BU} = \mathbf{t}(\tilde{\boldsymbol{\varphi}}_{BU}) = \begin{bmatrix} \cos(\tilde{\phi}_{BU}) \cos(\tilde{\theta}_{BU}) \\ \sin(\tilde{\phi}_{BU}) \cos(\tilde{\theta}_{BU}) \\ \sin(\tilde{\theta}_{BU}) \end{bmatrix}, \quad (17)$$

где $\mathbf{t}(\boldsymbol{\varphi})$ – функция отображения углов $\boldsymbol{\varphi}$ прихода / ухода AOA / AOD на вектор $\mathbf{t}$ направления прихода / ухода AOA / AOD.

Справедливо и обратное отображение вектора $\tilde{\mathbf{t}}_{BU}$ (13) направления прихода / ухода AOA / AOD (от UE на BS) на углы $\tilde{\boldsymbol{\varphi}}_{BU}$ прихода / ухода AOA / AOD:

$$\tilde{\boldsymbol{\varphi}}_{BU} = \begin{bmatrix} \tilde{\phi}_{BU} \\ \tilde{\theta}_{BU} \end{bmatrix} = \begin{bmatrix} \tan^{-1}(\tilde{t}_{BU,y} / \tilde{t}_{BU,x}) \\ \sin^{-1}(\tilde{t}_{BU,z}) \end{bmatrix}. \quad (18)$$

Аналогично преобразованиям (17) и (18) в локальной СК можно выполнить соответствующие преобразования и в глобальной СК для углов $\boldsymbol{\varphi}_{BU}$ прихода / ухода AOA / AOD и векторов $\mathbf{t}_{BU}$ направления прихода / ухода AOA / AOD.

Выше были формализованы геометрические соотношения для учета PPB в радиолинии BS↔UE в условиях LOS между BS и UE: расстояние d_{BU} (9), вектор направления прихода / ухода AOA / AOD в глобальной $\mathbf{t}_{BU}$ (12) и локальной $\tilde{\mathbf{t}}_{BU}$ (13) СК, углы прихода / ухода AOA / AOD в глобальной $\boldsymbol{\varphi}$ (15) и локальной $\tilde{\boldsymbol{\varphi}}$ (16) СК. Похожее соотношение можно записать и для радиолиний BS↔RIS и RIS↔UE как для условий LOS, так и для $l = 1, \dots, L_N$ однократно отраженных МЛК в условиях NLOS; например, d_{BR} , d_{ru} , $d_{BN}^{(l)}$, $d_{NU}^{(l)}$. Далее формализуем модель MIMO канала в дальней зоне для радиолиний BS↔UE, BS↔RIS и RIS↔UE.

3.2. Модель канала MIMO-системы в дальней зоне

3.2.1. Общее выражение модели канала MIMO

Для исследования и разработки методов и алгоритмов позиционирования в сетях шестого поколения миллиметрового и субмиллиметрового или терагерцового диапазона необходима надлежащая модель радиоканала. Известны детерминированные, статистические и гибридные подходы к моделированию физических явлений при РРВ [18]. Для решения задач позиционирования рассмотрим детерминированную модель РРВ в дальней зоне, когда фронт волны предполагается плоским РВМ (аббр. от англ. Planar Wave Model); в модели NF, напротив, учитывают SWM.

Будем полагать, что используются OFDM-сигналы и анализируется восходящий канал UE→BS, тогда матрицу комплексных коэффициентов канала размерности $\mathbf{H} \in \mathbb{C}^{N_B \times N_U}$ можно представить выражением [53]:

$$\mathbf{H} = \mathbf{H}_L + \mathbf{H}_R + \mathbf{H}_N, \quad (19)$$

где $\mathbf{H}_L$ – матрица канала LOS, $\mathbf{H}_R$ – матрица канала RIS, $\mathbf{H}_N$ – матрица канала МЛК в условиях NLOS.

3.2.2. Модель канала в условиях наличия прямой видимости

Матрицу канала LOS для радиолинии BS–UE на центральной несущей частоте f можно представить следующим выражением [53]:

$$\mathbf{H}_L(t, f) = \rho_L(f) e^{-j2\pi(f\tau_{BU} - v_{BU}t)} \times G_B(\tilde{\boldsymbol{\varphi}}_{BU}) G_U(\tilde{\boldsymbol{\varphi}}_{UB}) \mathbf{a}_B(f, \tilde{\boldsymbol{\varphi}}_{BU}) \mathbf{a}_U^T(f, \tilde{\boldsymbol{\varphi}}_{UB}), \quad (20)$$

где $\rho_L(f)$ – коэффициент ослабления LOS на несущей частоте f ; v_{BU} – доплеровский сдвиг частоты; τ_{BU} – задержка при РРВ в радиолинии BS–UE, включающая время распространения сигнала d_{BU}/c на расстояние d_{BU} прямой видимости, а также сдвиг синхронизации B в радиолинии BS–UE; G_B и G_U – КУ антенной решетки BS и UE, соответственно; $\mathbf{a}_B$ и $\mathbf{a}_U$ – направляющие вектора антенной решетки BS

и UE, соответственно; КУ и направляющие вектора зависят от направления прихода AOA на BS $\tilde{\boldsymbol{\varphi}}_{BU}$ и направления ухода AOD от UE $\tilde{\boldsymbol{\varphi}}_{UB}$ в локальной СК.

Амплитудный коэффициент ослабления луча LOS на несущей частоте f определяется выражением:

$$\rho_L(f) = \frac{c}{4\pi f d_{BU}} \mathcal{K}_a(f, d_{BU}), \quad (21)$$

где $\mathcal{K}_a(f, d_{BU})$ – коэффициент затухания, определяемый несущей частотой f и расстоянием d_{BU} ; в диапазоне ММВ коэффициент затухания учитываются коэффициентом $\mathcal{K}_a = \mathcal{K}_{atm}(f, d)$; в терагерцовом диапазоне молекулярное поглощение, вызванное водяными парами и другими газами, увеличивает потери РРВ, и учитывается коэффициентом $\mathcal{K}_a = e^{-\mathcal{K}_{abs}(f)d/2}$ [53].

Используя идеальную секторную модель (ISM, аббр. от англ. Ideal Sector Model), КУ решеток в условиях LOS на BS и UE можно аппроксимировать выражением:

$$G_Q(\tilde{\boldsymbol{\varphi}}) = \begin{cases} \sqrt{G_Q^0}, & \tilde{\boldsymbol{\varphi}} \in \left[-\frac{\Phi_h}{2}; \frac{\Phi_h}{2}\right], \tilde{\boldsymbol{\theta}} \in \left[-\frac{\Theta_h}{2}; \frac{\Theta_h}{2}\right], \\ 0, & \text{в остальных случаях.} \end{cases} \quad (22)$$

где G_Q^0 – КУ в направлении максимального излучения; $\tilde{\boldsymbol{\varphi}}$ – направление прихода / ухода AOA / AOD сигнала, задаваемое парой углов (18) в локальной СК; Φ_h и Θ_h – ширина луча или диаграммы направленности антенны по уровню половинной мощности (HPBW, аббр. от англ. Half Power Beam Width) в горизонтальной и вертикальной плоскостях, соответственно. Для ненаправленных антенн КУ равен единице и справедливо выражение $G_Q(\tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\theta}}) = 1$.

Рисунок 3 иллюстрирует пример использования узконаправленных антенн, для которых КУ можно аппроксимировать выражением [53]:

$$G^0 \approx 4\pi/(\Theta_h \Phi_h). \quad (23)$$

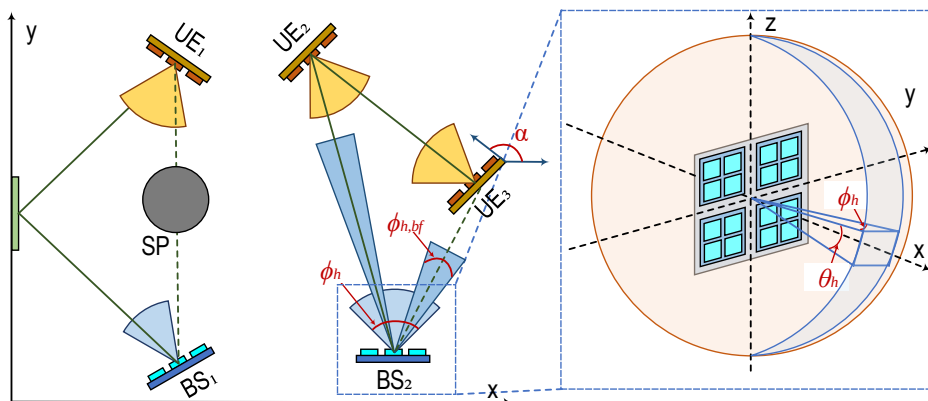


Рис. 3. Учет ширины луча при позиционировании в сетях 6G

Fig. 3. Taking into Account the Beam Width When Positioning in 6G Networks

Следует различать ширину луча по уровню половинной мощности в горизонтальной $\phi_{h,bf}$ и вертикальной $\theta_{h,bf}$ плоскостях при ДО, от НРВW в горизонтальной/вертикальной плоскостях ϕ_h/θ_h , определяемой размерностью АР.

Частотно-зависимые направляющие вектора $\mathbf{a}_B(f, \tilde{\Phi}_{BU})$ и $\mathbf{a}_U(f, \tilde{\Phi}_{UB})$ можно представить следующим выражением [53]:

$$\mathbf{a}_Q(\tilde{\Phi}) = [a_Q(1), \dots, a_Q(q), \dots, a_Q(N_Q)]^T, \quad (24)$$

где q -й элемент вектора $\mathbf{a}_Q(\tilde{\Phi})$ определяется выражениями:

$$\begin{aligned} a_Q(q) &= e^{j\frac{2\pi f}{c}(\mathbf{R}_Q^T(\mathbf{p}_q - \mathbf{p}_Q))^T(\mathbf{R}^T \mathbf{t}(\Phi))} = \\ &= e^{j\frac{2\pi f}{c}(\mathbf{p}_q - \mathbf{p}_Q)^T \mathbf{t}(\Phi)}, \end{aligned} \quad (25)$$

$$a_Q(q) = e^{j\frac{2\pi f}{c}\Psi_q(\tilde{\Phi})} = e^{j\frac{2\pi f}{c}\tilde{\mathbf{p}}_q^T \mathbf{t}(\tilde{\Phi})}, \quad (26)$$

где элементы $a_Q(q)$ направляющего вектора описываются с использованием пары углов в глобальной Φ (15) и локальной $\tilde{\Phi}$ (16) СК, соответственно; отображение вектора углов $\tilde{\Phi}_{BU}$ на вектор направления $\tilde{\mathbf{t}}_{BU}$ определяется в (17).

Допустим далее, что влиянием доплеровского сдвига частоты ν_{BU} можно пренебречь, а коэффициенты затухания одинаковы на всех поднесущих OFDM-сигнала, т. е. $\mathcal{K}_a(f, d) = \mathcal{K}_a(f_c, d)$, где f_c – центральная несущая частота. Тогда матрицу канала LOS на k -й поднесущей можно представить как:

$$\mathbf{H}_L[k] = c_k \rho_L e^{-j\xi_L} e^{-j2\pi\Delta f_k \tau_L} \mathbf{a}_B(\tilde{\Phi}_{BU}) \mathbf{a}_U^T(\tilde{\Phi}_{UB}), \quad (27)$$

где отношение между центральной несущей f_c и поднесущей f_k частотами равно:

$$c_k = f_c / f_k, \quad (28)$$

а частота поднесущих с разносом Δf_k вычисляется по формуле:

$$f_k = f_c + \Delta f_k = f_c + (2k - 1 - K)W / (2K), \quad (29)$$

направляющие вектора $\mathbf{a}_B(\tilde{\Phi}_{BU})$ и $\mathbf{a}_U(\tilde{\Phi}_{UB})$ на k -й поднесущей вычисляются по формуле (24) с подстановкой в (25) и (26) значений f_k (29) вместо f .

Комплексный коэффициент канала $\rho_L e^{-j\xi_L}$ и задержка τ_L времени прихода сигнала TOA из (27) определяются выражениями [53]:

$$\rho_L = \frac{\lambda_c}{4\pi d_{BU}} \mathcal{K}_a(f_c, d_{BU}) G_B(\tilde{\Phi}_{BU}) G_U(\tilde{\Phi}_{UB}), \quad (30)$$

$$\xi_L = 2\pi f_c \tau_L, \quad (31)$$

$$\tau_L = d_{BU} / c + B, \quad (32)$$

где c – скорость света; B – сдвиг синхронизации.

Допущение 1. При $B = 0$ имеем радиолинию BS-UE с идеальной синхронизацией. Сдвиг синхронизации B предполагается одинаковым для всех передач в радиолинии BS-UE. Для многопозиционной

системы ОМП с несколькими BS, каждая из которых может характеризоваться своим сдвигом синхронизации.

Допущение 2. Несмотря на то, что параметры комплексного коэффициента канала ρ_L / ξ_L могут быть установлены из геометрических соотношений взаимного расположения BS и UE в условиях LOS, их обычно трактуют как неизвестные параметры вследствие динамических условий РРВ и нелинейности аппаратного обеспечения приемопередающего тракта миллиметрового и терагерцового диапазона. Такой сценарий позиционирования далее называется сценарием с неизвестным каналом. В сценарии позиционирования с частично известным каналом параметр ρ_L предполагают известным, а ξ_L – неизвестным; дальность d_{BU} можно оценить из коэффициента ослабления луча LOS ρ_L .

3.2.2. Модель канала реконфигурируемой интеллектуальной поверхности

В радиолинии с RIS переданный UE сигнал сначала принимается RIS и, таким образом, образует радиолинию RIS-UE. Затем, после изменения амплитуд и фаз элементами RIS, сигнал передается на BS и, таким образом, образует радиолинию BS-RIS.

Канальную матрицу RIS $\mathbf{H}_R$ можно представить следующим выражением [53]:

$$\mathbf{H}_R = \mathbf{H}_{BR} \mathbf{\Omega} \mathbf{H}_{RU}, \quad (33)$$

где $\mathbf{H}_{BR} \in \mathbb{C}^{N_B \times N_R}$, $\mathbf{H}_{RU} \in \mathbb{C}^{N_R \times N_U}$ – канальные матрицы BS-RIS и RIS-UE, соответственно; комплексные коэффициенты составных канальных матриц $\mathbf{H}_{BR}$ и $\mathbf{H}_{RU}$.

Диагональная матрица $\mathbf{\Omega} \in \mathbb{C}^{N_R \times N_R}$ образована комплексными коэффициентами:

$$\mathbf{\Omega} \triangleq \text{diag}(\beta_1 e^{jw_1}, \dots, \beta_{N_R} e^{jw_{N_R}}), \quad (34)$$

где $w_n \in [0, 2\pi)$ – фазовый сдвиг; $\beta_n \in [0, 1]$ – коэффициент отражения n -го элемента RIS, $n = 1, \dots, N_R$.

Из допущения о ненаправленном излучении и единичном КУ отдельного элемента RIS, а также принимая площадь одного элемента RIS равной $\lambda^2 / (4\pi)$, канальную матрицу $\mathbf{H}_R$ на k -й поднесущей можно представить выражением [53]:

$$\mathbf{H}_R[k] = c_k^2 \rho_R e^{-j\xi_R} e^{-j2\pi\Delta f_k \tau_R} \times \mathbf{a}_B(\tilde{\Phi}_{BR}) \mathbf{a}_R^T(\tilde{\Phi}_{RB}) \mathbf{\Omega} \mathbf{a}_R(\tilde{\Phi}_{RU}) \mathbf{a}_U^T(\tilde{\Phi}_{UR}), \quad (35)$$

где

$$\begin{aligned} \rho_R &= \frac{\lambda_c^2}{16\pi^2 d_{BR} d_{UR}} \mathcal{K}_a(f_c, d_{BR}) \mathcal{K}_a(f_c, d_{UR}) \times \\ &\times G_B(\tilde{\Phi}_{BR}) G_R(\tilde{\Phi}_{RB}) G_R(\tilde{\Phi}_{RU}) G_U(\tilde{\Phi}_{UR}), \end{aligned} \quad (36)$$

$$\xi_R = 2\pi f_c \tau_R, \quad (37)$$

$$\tau_R = \frac{d_{BR} + d_{UR}}{c} + B. \quad (38)$$

3.2.3. Модель канала в условиях отсутствия прямой видимости

В условиях NLOS переданный сигнал может отражаться другими объектами в радиолинии и приниматься после однократных и/или многократных переотражений. Помимо комплексных коэффициентов канала следует дополнительно учитывать и коэффициенты отражения. В диапазоне ММВ коэффициенты отражения можно оценить из соотношений геометрической оптики и статистики ослаблений сигналов после отражений в радиолинии. В диапазоне терагерцовых волн условия NLOS имеют следующие отличия. Во-первых, МЛК терагерцового диапазона испытывают существенно более высокие потери при РРВ, поэтому, значимых с точки зрения энергетики радиоприема МЛК становится существенно меньше; в итоге канал характеризуется еще большей, по сравнению с диапазоном ММВ, разреженностью. Во-вторых, отражающие поверхности, которые для диапазона ММВ могут рассматриваться как гладкие, в терагерцовом диапазоне уже характеризуются неровностями и приводят к рассеянию и дифракции субмиллиметровых волн.

Для учета МЛК терагерцового радиоканала в условиях NLOS могут использоваться статистические подходы, а также модели и методы трассировки лучей. Для задач позиционирования будем далее учитывать в радиолинии между BS и UE наличие L_N разрешаемых в домене времени МЛК и соответствующих им отражателей с координатами $\mathbf{p}_N$ и NLOS коэффициентами $\mathcal{K}_N^{(l)}$; при этом рассеянными компонентами пренебрегают, а каждому отражателю соответствует единственная однократно отраженная МЛК.

Канальную матрицу NLOS $\mathbf{H}_N$ можно представить суммой канальных матриц L_N МЛК:

$$\mathbf{H}_N = \sum_{l=1}^{L_N} \mathbf{H}_N^{(l)}, \quad (39)$$

где канальная матрица k -й поднесущей для l -й МЛК определяется как [53]:

$$\mathbf{H}_N^{(l)}[k] = c_k \rho_N^{(l)} e^{-j\xi_N^{(l)}} e^{-j2\pi\Delta f_k \tau_N^{(l)}} \times \mathbf{a}_R(\tilde{\boldsymbol{\varphi}}_{RU}) \mathbf{a}_U^T(\tilde{\boldsymbol{\varphi}}_{UR}), \quad (40)$$

где

$$\rho_N^{(l)} = \frac{\lambda_c}{4\pi d_N^{(l)}} \times \mathcal{K}_N^{(l)} \mathcal{K}_a^{(l)}(f_c, d_N^{(l)}) G_B^{(l)}(\tilde{\boldsymbol{\varphi}}_{NB}) G_U^{(l)}(\tilde{\boldsymbol{\varphi}}_{UN}), \quad (41)$$

$$\xi_N^{(l)} = 2\pi f_c \tau_N^{(l)}, \quad (42)$$

$$\tau_N^{(l)} = \frac{d_{BN}^{(l)} + d_{NU}^{(l)}}{c} + B. \quad (43)$$

Формализовав модели канала MIMO в дальней зоне при LOS (27) и NLOS (40), а также для RIS (35), далее рассмотрим модели приема в традиционной

MIMO-системе с едиными массивом AP, а также в UM-MIMO-системе, образованной составными массивами решеток AOSA.

3.3. Модель приема MIMO-системы в дальней зоне

3.3.1. Архитектура MIMO-систем

В MIMO-системах диапазона ДМВ обработка сигналов при передаче и приеме для всех элементов AP производится в цифровом виде в информационной полосе частот. Каждый элемент AP непосредственно связан с радиочастотным модулем RFC (*аббр. от англ.* Radio Frequency Chain), т. е. оборудован своим аналоговым трактом: при приеме сигнал поступает с RFC на АЦП, а при передаче сигнал поступает с цифроаналогового преобразователя (ЦАП) на RFC. ДО на передачу и прием реализуется подбором комплексных весовых коэффициентов AP в цифровом виде. Понижающее (при приеме) и повышающее (при передаче) преобразование частоты в модуле RFC обычно реализуется в аналоговой форме. Такой подход называют цифровым ДО.

С увеличением несущей частоты при переходе в диапазон ММВ для компенсации потерь при РРВ необходимо повышать направленность AP, измеряемую КУ, путем увеличения ее размерности, измеряемой числом элементов AP. В таком случае описанный выше подход цифрового ДО, когда каждому элементу AP соответствует свой аналоговый тракт RFC, становится сложно реализуемым на практике ввиду пропорционального роста энергопотребления. Для использования преимуществ ДО многоантенных MIMO-систем с AP высокой размерности на практике реализуется подход гибридного (аналого-цифрового) ДО. Символы данных в информационной полосе частот сначала прекодируются цифровым прекодером, а затем поступают на аналоговую диаграммообразующую схему, реализованную на фазовращателях или линзовых антеннах. AP при этом может быть представлена единым массивом элементов, или же набором из составных массивов элементов AOSA.

Для полносвязной архитектуры с фазовращателями каждый аналоговый тракт RFC связан со всеми элементами *единого массива* AP через индивидуальный набор фазовращателей. В архитектуре *составного массива* AOSA каждый аналоговый тракт RFC связан только с частью элементов всей AP, образующих составной массив SA. Для UM-MIMO-диапазона терагерцовых волн по критериям спектральной эффективности и энергопотребления предпочтительнее оказывается архитектура составного массива AP AOSA. Далее формализуем модели приема сигналов по схеме цифрового и гибридного или аналого-цифрового ДО единым и составным массивами AP.

3.3.2. Модель приема в цифровой ММО-системе с единым массивом АР

В цифровой ММО-системе каждому элементу АР соответствует свой аналоговый тракт RFC и АЦП / ЦАП. Пусть P – средняя мощность передачи. Тогда для модели канала ММО в дальней зоне, определяемой формулами (19)–(43), вектор принятого сигнала на k -й поднесущей при g -й передаче OFDM-символа можно представить выражением:

$$\begin{aligned} \mathbf{y}^{(g)}[k] &= \sqrt{P}\mathbf{H}[k]\mathbf{x}^{(g)}[k] + \mathbf{n}^{(g)}[k] = \\ &= \boldsymbol{\mu}^{(g)}[k] + \mathbf{n}^{(g)}[k], \end{aligned} \quad (44)$$

где $\boldsymbol{\mu}^{(g)}[k] \in \mathbb{C}^{N_B \times 1}$ – вектор принятого сигнала на k -й поднесущей при g -й передаче без учета вектора $\mathbf{n}^{(g)}[k] \in \mathbb{C}^{N_B \times 1}$ комплексных выборок аддитивного белого гауссовского шума (АБГШ):

$$\boldsymbol{\mu}^{(g)}[k] = \sqrt{P}\mathbf{H}[k]\mathbf{x}^{(g)}[k]. \quad (45)$$

Комплексные выборки вектора АБГШ $\mathbf{n}^{(g)}[k]$ имеют нормальное распределение $\mathcal{CN}(0, \sigma_n^2)$ с нулевым средним и дисперсией σ_n^2 , а нормированный вектор переданного сигнала $\mathbf{x}^{(g)}[k] \in \mathbb{C}^{N_U \times 1}$ может быть выбран произвольно, или согласно порядку работы диаграммообразующей схемы, при ограничении:

$$\|\mathbf{x}^{(g)}[k]\|^2 = 1. \quad (46)$$

Если совокупная энергия распределяется по всем $k = 1, \dots, K$ поднесущим и всем $g = 1, \dots, G$ OFDM-символам, нормированный вектор переданного сигнала $\mathbf{x}^{(g)}[k] \in \mathbb{C}^{N_U \times 1}$ удовлетворяет условию:

$$\|\mathbf{x}^{(g)}[k]\|^2 = \frac{1}{KG}. \quad (47)$$

Для многоантенных систем сверхвысокой размерности в терагерцовом диапазоне использование подхода обработки сигналов в полностью цифровой ММО-системе сложно реализуемо на практике, однако позволяет установить пределы точности работы алгоритмов и методов позиционирования. Будем далее полагать, что модель (44) используется для установления пределов точности позиционирования систем, работающих в диапазоне ММВ. Далее рассмотрим модель приема в гибридной ММО-системе с обработкой в едином и составном АOSA массиве АР.

3.3.3. Модель приема в гибридной ММО-системе с единым массивом АР

При высокой размерности АР соединять каждый ее элемент с аналоговым трактом RFC и АЦП / ЦАП становится непрактично, поэтому используется гибридная схема ММО-системы, когда для обработки

сигналов с N_B или N_U антенн используется значительно меньшее число RFC M_B или M_U на BS или UE, при этом:

$$M_B \leq N_B, \quad M_U \leq N_U. \quad (48)$$

Модель принятого сигнала в гибридной ММО-системе можно представить выражением [53]:

$$\mathbf{y}^{(g)}[k] = \sqrt{P}\mathbf{W}_B^T \mathbf{H}[k] \mathbf{W}_U \mathbf{x}_0^{(g)}[k] + \mathbf{W}_B^T \mathbf{n}^{(g)}[k], \quad (49)$$

где $\mathbf{W}_B \in \mathbb{C}^{N_B \times M_B}$, $\mathbf{W}_U \in \mathbb{C}^{N_U \times M_U}$ – аналоговые диаграммообразующие матрицы комбинирования BS и прекодирования UE из комплексных коэффициентов на несущей радиочастоте, соответственно; $\mathbf{x}_0^{(g)}[k] \in \mathbb{C}^{M_U \times 1}$ – вектор переданных символов до обработки в прекоде устройства.

При сохранении ограничения на переданную мощность, согласно (46), введем вектор переданных сигналов после прекодирования в UE:

$$\mathbf{x}^{(g)}[k] = \mathbf{W}_U \mathbf{x}_0^{(g)}[k]. \quad (50)$$

Цифровые диаграммообразующие матрицы комбинирования BS и прекодирования UE в информационной полосе частот в настоящей работе не рассматриваются, поскольку для задач позиционирования обработка первичных измерений осуществляется сразу на выходе аналоговых трактов RFC на несущей радиочастоте.

Коэффициенты матриц $\mathbf{W}_B$ и $\mathbf{W}_U$ представляют собой фазовые сдвиги аналоговых фазовращателей с амплитудами [53]:

$$|[\mathbf{W}_B]_{i,j}| = 1/\sqrt{N_B}, \quad |[\mathbf{W}_U]_{i,j}| = 1/\sqrt{N_U}. \quad (51)$$

Полносвязная архитектура гибридных ММО-систем используется для диапазона ММВ. Однако для диапазона терагерцовых волн она оказывается сложно реализуемой на практике и малоэффективной ввиду ограничения на переданную мощность; дополнительным ее недостатком полносвязной архитектуры гибридных ММО-систем является повышенное энергопотребление вследствие роста числа антенных элементов (АЭ). Далее рассмотрим архитектуру UM-MIMO-систем, использующих составные массивы АOSA решеток.

3.3.4. Модель приема в ММО-системе из составных массивов антенной решетки

В ММО-системе из составных массивов АOSA решетка разделяется на несколько составных подмассивов SA. Каждый составной подмассив эффективно представляет отдельный элемент ММО-системы, обслуживаемый своим аналоговым трактом RFC. В отличие от традиционной цифровой ММО-системы с единым массивом, где каждому элементу АР соответствует свой отдельный RFC, в ММО-системе из составных массивов отдельный

RFC обслуживает составной подмассив SA, что позволяет существенно снизить число RFC и, соответственно, энергопотребление. В отличие от полносвязной архитектуры гибридной MIMO-системы с единым массивом, где каждый RFC связан со всеми элементами AP через набор фазовращателей, в MIMO-системе из составных массивов отдельный RFC обслуживает элементы AP только внутри составного подмассива SA, т.е. аналоговое ДО в заданном направлении реализуется фазовращателями для каждого элемента внутри отдельного составного подмассива SA. Теоретически MIMO-система из составных массивов реализует выигрыш в КУ при ДО за счет индивидуальных фазовращателей каждого элемента подмассива; при этом формирование луча в заданном направлении реализуется составным подмассивом, и каждый SA может осуществлять ДО в своем направлении углов прихода / ухода AOA / AOD. Практически реализация управления фазовращателями каждого антенного элемента составного подмассива затруднена по причине высокой вычислительной сложности, поэтому целесообразнее трактовать каждый SA как отдельный элемент MIMO-системы с фиксированным углом ДО в заданном направлении прихода / ухода AOA / AOD.

Обозначим через $\mathcal{A}_Q(\tilde{\Phi}, \Phi')$ коэффициент AP составного подмассива SA, состоящего из N'_Q антенных элементов (АЭ), $Q \in \{B, U\}$. Данный коэффициент характеризует усиление в направлении прихода / ухода AOA / AOD $\tilde{\Phi}$ в локальной СК при ДО в отдельном составном подмассиве SA для луча, ориентированного в направлении максимального излучения Φ' [53]:

$$\begin{aligned} \mathcal{A}_Q(\tilde{\Phi}, \Phi') &= 1/\sqrt{N'_Q} \mathbf{a}_{st}^T(\tilde{\Phi}) \mathbf{a}_{bf}(\Phi') = \\ &= 1/\sqrt{N'_Q} \sum_{q=1}^{N'_Q} e^{j\frac{2\pi f}{c}(\Psi_{q'}(\tilde{\Phi}) - \Psi_{q'}(\Phi'))}, \end{aligned} \quad (44)$$

где задержка сигнала на q' -м элементе подмассива SA по отношению к его центру определяется выражением:

$$\Psi_{q'}(\tilde{\Phi}) = \tilde{\mathbf{p}}_{q'}^T \mathbf{t}(\tilde{\Phi}). \quad (53)$$

Направляющий $\mathbf{a}_{st}(\tilde{\Phi})$ и диаграммообразующий $\mathbf{a}_{bf}(\Phi')$ вектора подмассива SA в (52) вычисляются по формулам [53]:

$$\mathbf{a}_{st}(\tilde{\Phi}) = [a_{st}(1), \dots, a_{st}(q'), \dots, a_{st}(N'_Q)]^T, \quad (54)$$

$$a_{st}(q') = e^{j\frac{2\pi f}{c}\Psi_{q'}(\tilde{\Phi})} = e^{j\frac{2\pi f}{c}\tilde{\mathbf{p}}_{q'}^T \mathbf{t}(\tilde{\Phi})}, \quad (55)$$

$$\mathbf{a}_{bf}(\Phi') = [a_{bf}(1), \dots, a_{bf}(q'), \dots, a_{bf}(N'_Q)]^T, \quad (56)$$

$$a_{bf}(q') = e^{j\frac{2\pi f}{c}\Psi_{q'}(\Phi')} = e^{j\frac{2\pi f}{c}\tilde{\mathbf{p}}_{q'}^T \mathbf{t}(\Phi')}. \quad (57)$$

Следует отметить, что направление Φ'_q диаграммообразования луча в отдельном составном q -м подмассиве не зависит от геометрического сценария расположения других устройств и коэффициент

$\mathcal{A}_Q(\tilde{\Phi}, \Phi')$ AP составного подмассива SA и достигает максимального значения при $\Phi' = \tilde{\Phi}$. Оба вектора – направляющий $\mathbf{a}_{st}(\tilde{\Phi})$ и диаграммообразующий $\mathbf{a}_{bf}(\Phi')$ – могут быть выражены в глобальной и локальной СК, как, например, в (25) и (26).

Модель принятого сигнала в MIMO-системе из составных массивов AP можно представить выражением:

$$\mathbf{y}^{(g)}[k] = \mathcal{H}[k] \mathbf{x}^{(g)}[k] + \mathbf{n}^{(g)}[k], \quad (58)$$

где $\mathcal{H}[k] \in \mathbb{C}^{N_B \times N_U}$ – канальная матрица комплексных коэффициентов составного массива AOSA; N_B и N_U – число подмассивов SA в составном массиве AOSA на BS и UE; $\mathcal{H}[k]$ определяется как [53]:

$$\begin{aligned} \mathcal{H} &= \mathcal{A}_L \odot \mathcal{H}_L[k] + \\ &+ (\mathcal{A}_{BR} \odot \mathcal{H}_{BR}[k]) \Omega (\mathcal{A}_{RU} \odot \mathcal{H}_{RU}[k]) + \\ &+ \sum_{l=1}^{L_N} \mathcal{A}_N^{(l)} \odot \mathcal{H}_N^{(l)}[k]. \end{aligned} \quad (59)$$

Из (59) следует, что в каждой компоненте (LOS, RIS, NLOS) эффективная канальная матрица составного массива AOSA получается в результате произведения Адамара матрицы коэффициентов $\mathcal{A}_L$, $\mathcal{A}_{BR}$, $\mathcal{A}_{RU}$, $\mathcal{A}_N^{(l)}$ составного массива AP на матрицы $\mathcal{H}_L$, $\mathcal{H}_{BR}$, $\mathcal{H}_{RU}$, $\mathcal{H}_N^{(l)}$ соответствующих им подмассивов SA. Канальные матрицы могут быть получены из модели канала MIMO в дальней зоне по формулам (27), (35), (40) с учетом принципа соответствия отдельного антенного элемента в цифровой MIMO-системе с единым массивом AP отдельному подмассиву SA в составном массиве AOSA решетки: 1) число подмассивов SA соответствует числу антенных элементов в цифровой MIMO-системе с единым массивом; 2) центр подмассива SA соответствует центру отдельного антенного элемента в цифровой MIMO-системе с единым массивом; 3) разнос подмассивов SA соответствует разносу отдельных антенных элементов в цифровой MIMO-системе с единым массивом.

Матрицы коэффициентов $\mathcal{A}_L$, $\mathcal{A}_{BR}$, $\mathcal{A}_{RU}$, $\mathcal{A}_N^{(l)}$ составного AOSA массива AP включают коэффициенты отдельных подмассивов SA, которые вычисляются по формулам [53]:

$$\mathcal{A}_{L,bu} = \mathcal{A}_B(\tilde{\Phi}_{BU}, \Phi'_b) \mathcal{A}_U(\tilde{\Phi}_{UB}, \Phi'_u), \quad (60)$$

$$\mathcal{A}_{BR,br} = \mathcal{A}_B(\tilde{\Phi}_{BR}, \Phi'_b), \quad (61)$$

$$\mathcal{A}_{RU,ru} = \mathcal{A}_U(\tilde{\Phi}_{UR}, \Phi'_u), \quad (62)$$

$$\mathcal{A}_{N,bu}^{(l)} = \mathcal{A}_B^{(l)}(\tilde{\Phi}_{BN}^{(l)}, \Phi'_b^{(l)}) \mathcal{A}_U^{(l)}(\tilde{\Phi}_{UN}^{(l)}, \Phi'_u^{(l)}). \quad (63)$$

В настоящем разделе для сценария дальней зоны были формализованы две модели приема сигналов в MIMO-системе – единым массивом AP и составными подмассивами SA решетки AOSA. Далее обобщим рассмотренные модели канала и приема в MIMO-системе для сценария NF с допущением о SWM.

3.4. Модель канала и приема MIMO-системы в ближней зоне

3.4.1. Модель канала MIMO-системы

Модель дальней зоны справедлива для сценария, когда расстояние между приемопередатчиками значительно превышает геометрические размеры АР. Модель NR определяется в пределах между границей зоны Френеля, определяемой выражением $0,62\sqrt{D^3/\lambda}$, и расстоянием Фраунгофера, определяемого соотношением $2D^2/\lambda$, где λ – длина волны, а D – диаметр АР [53]. С увеличением центральной несущей частоты в терагерцовом диапазоне и соответствующим уменьшением длины волны λ использование АР сверхвысокой размерности может привести к такому ее диаметру D , что модель дальней зоны перестает отвечать условиям сценария и становится необходимо учитывать модель NF. Например, на несущей частоте $f = 0,3$ ТГц с линейной эквидистантной АР ULA из 100 антенных элементов с разномом в половину длины волны получим NF от 0,2 до 5,0 м (скрипт 1).

Скрипт 1. Оценка условий ближней зоны

```
% допущение о ближней/дальней зоне
f=0.3e12; % несущая частота, Гц
c=3e8; % скорость света
lambda=c/f; % длины волны, м
nel=100; % размерность ULA: для URA nel×nel
delta=lambda/2; % расстояние между элементами АР
D=delta*nel; % диаметр АР, м
frb=0.62*sqrt((D^3)/lambda) % граница зоны Френеля
frd=(2*D^2)/lambda % расстояние Фраунгофера
% допущение о широкополосной/узкополосной модели
W=0.1e12; % ширина полосы частот
nw=c/(nel*delta)
```

В модели NR дополнительно вводится допущение о том, что принятые на различные антенные элементы сигналы имеют одинаковую амплитуду ρ . Отличие же модели NR от модели дальней зоны заключается в том, что принятые на различные антенные элементы сигналы имеют различные фазы. Рассмотрим канальную матрицу MIMO-системы в условиях LOS $\mathbf{H}_L[k]$ без учета сдвига синхронизации B . Каждый элемент $h_{L,bu}[k]$ b -й строки и u -го столбца матрицы $\mathbf{H}_L[k]$ в условиях PWM и SWM может быть представлен выражениями [53]:

$$h_{L,bu}^{SWM}[k] = c_k \rho_L e^{-j\frac{2\pi f_k d_{bu}}{c}} = c_k \rho_L e^{-j\frac{2\pi f_k}{c} \|\mathbf{p}_b - \mathbf{p}_u\|}, \quad (64)$$

$$h_{L,bu}^{PWM}[k] = c_k \rho_L e^{-j\frac{2\pi f_k}{c} (\|\mathbf{p}_b - \mathbf{p}_u\| - \tilde{\mathbf{p}}_b^T \mathbf{t}(\tilde{\Phi}_{BU}) - \tilde{\mathbf{p}}_u^T \mathbf{t}(\tilde{\Phi}_{UB}))}. \quad (65)$$

Выражение (65) является элементом канальной матрицы (27). Извлекая элементы из канальных матриц MIMO-системы в дальней зоне (27), (35), (40) на различных поднесущих и учитывая комплексные коэффициенты канала $\rho e^{-j\xi}$ с фиксированной амплитудой ρ и переменной фазой ξ для SWM, можем представить элементы этих канальных матриц для модели NR следующими выражениями [53]:

$$h_{L,bu}^{SWM}[k] = c_k \rho_L e^{-j\xi_L} e^{-j2\pi(\Delta f_k \tau_L + f_k \Delta \tau_{bu})}, \quad (66)$$

$$h_{R,bu}^{SWM}[k] = c_k \rho_R e^{-j\xi_R} \sum_{r=1}^{N_R} e^{-j2\pi(\Delta f_k \tau_R + f_k \Delta \tau_{bru})}, \quad (67)$$

$$h_{N,bu}^{(l)SWM}[k] = c_k \rho_N^{(l)} e^{-j\xi_N^{(l)}} e^{-j2\pi(\Delta f_k \tau_N^{(l)} + f_k \Delta \tau_{bnu}^{(l)})}, \quad (68)$$

где

$$\Delta \tau_{bu} = \tau_{bu} - \tau_L = \frac{d_{bu} - d_{BU}}{c}, \quad (69)$$

$$\Delta \tau_{bru} = \tau_{bru} - \tau_R = \frac{d_{br} + d_{rm} - d_{BR} - d_{RU}}{c}, \quad (70)$$

$$\Delta \tau_{bnu}^{(l)} = \tau_{bnu}^{(l)} - \tau_N^{(l)} = \frac{d_{bn}^{(l)} + d_{nu}^{(l)} - d_{BN}^{(l)} - d_{NU}^{(l)}}{c}. \quad (71)$$

Для канальной матрицы комплексных коэффициентов составного массива AOSA в (59) в условиях модели ближней зоны будем полагать, что на каждом составном подмассиве SA справедливо допущение о PWM, а SWM проявляется в различных фазах принятого сигнала на различных составных подмассивах SA. Это допущение оправдано, поскольку размер отдельного подмассива SA можно считать достаточно малым по сравнению с размером всего составного массива AOSA.

Таким образом, коэффициенты отдельных подмассивов SA в (60)–(63) для модели NR можно уточнить следующими выражениями [53]:

$$\mathcal{A}_{L,bu} = \mathcal{A}_B(\tilde{\Phi}_{bu}, \Phi'_{bu}) \mathcal{A}_U(\tilde{\Phi}_{ub}, \Phi'_{ub}), \quad (72)$$

$$\mathcal{A}_{BR,br} = \mathcal{A}_B(\tilde{\Phi}_{br}, \Phi'_{br}), \quad (73)$$

$$\mathcal{A}_{RU,ru} = \mathcal{A}_U(\tilde{\Phi}_{ur}, \Phi'_{ur}), \quad (74)$$

$$\mathcal{A}_{N,bu}^{(l)} = \mathcal{A}_B^{(l)}(\tilde{\Phi}_{bn}^{(l)}, \Phi'_{bn}^{(l)}) \mathcal{A}_U^{(l)}(\tilde{\Phi}_{nu}^{(l)}, \Phi'_{nu}^{(l)}). \quad (75)$$

Коэффициенты (72)–(75) отдельных подмассивов SA для модели NF отличаются от коэффициентов (60)–(63) отдельных подмассивов SA для модели дальней зоны тем, что в последнем случае пары углов ухода / прихода AOA / AOD рассчитываются для центра всего составного массива AOSA, а в модели NF – для центров каждого составного подмассива SA.

3.4.2. Модель эффекта расщепления луча

Эффект расщепления луча BSE, также известный как эффект «прищуривания» луча, происходит вследствие частотно-независимых постоянных фазовых сдвигов в аналоговой диаграммообразующей схеме. Если в узкополосной системе направляющие вектора комплексных коэффициентов (24) и (54) не зависят от частоты, т. е. справедливо тождество $f = f_c$, то в широкополосной системе направляющие вектора комплексных коэффициентов (24) и (54) являются частотно-зависимыми. При использовании в аналоговой диаграммообразующей схеме идеальных фазовращателей на различных поднесущих f_k широкополосного OFDM-сигнала будет наблюдаться постоянный фазовый сдвиг.

Для учета фазового сдвига коэффициент составного массива AP (52) следует уточнить путем умножения задержки сигнала на q' -м элементе подмассива SA по отношению к его центру (53) на коэффициент:

$$1/c_k = f_k/f_c, \quad (76)$$

после чего получим следующее выражение для $\mathcal{A}_Q(\tilde{\Phi}, \Phi')$ [53]:

$$\mathcal{A}_Q(\tilde{\Phi}, \Phi') = 1/\sqrt{N'_Q} \sum_{q'=1}^{N'_Q} e^{j\frac{2\pi f_c}{c} \left(\frac{f_k}{f_c} \Psi_{q'}(\tilde{\Phi}) - \Psi_{q'}(\Phi') \right)}. \quad (77)$$

Можно выделить следующие три фактора, влияющие на проявление BSE:

- 1) размер AP D ;
- 2) ширина полосы сигнала W ;
- 3) угол в направлении которого выполняется ДО.

Из уравнения (77) следует, что максимальный коэффициент ДО $\mathcal{A}_Q(\tilde{\Phi}, \Phi')$ достигает значения $\sqrt{N'_Q}$ и наблюдается только на центральной несущей частоте f_c , где вектор угла ДО Φ' равен направляющему вектору $\tilde{\Phi}$; на остальных поднесущих коэффициент $\mathcal{A}_Q(\tilde{\Phi}, \Phi')$ оказывается ниже. BSE можно также наблюдать и в канале с RIS; в этом случае матрица коэффициентов Ω в (33) становится частотно-зависимой. Влияние BSE на точность ОК и ориентации планируется оценить в третьей части цикла исследований.

3.4.3. Модель приема с учетом несовершенства аппаратного обеспечения

Несовершенство аппаратного обеспечения проявляется в ухудшении характеристик компонентов аналогового тракта RFC, фазовращателей и элементов RIS. Формализуем далее модели приема, учитывающие ухудшения, непосредственно влияющие на точность позиционирования.

Ухудшение характеристик аналогового тракта RFC проявляется в его дополнительном шуме и приводит к искажению сигнала на передаче и приеме. Принятый сигнал с учетом шума RFC можно представить выражением:

$$\mathbf{y} = \mathbf{H}(\mathbf{x} + \mathbf{n}_t) + \mathbf{n}_r + \mathbf{n}, \quad (78)$$

где $\mathbf{n}_t \sim \mathcal{CN}(0, \kappa_t^2 \bar{P})$, $\mathbf{n}_r \sim \mathcal{CN}(0, \kappa_r^2 \bar{P} |h|^2)$ – шумы искажений от ухудшения характеристик RFC на передаче и приеме, соответственно; κ_t , κ_r – коэффициенты ухудшения характеристик RFC на передаче и приеме, соответственно; $\bar{P}$ – средняя мощность передачи; $|h|^2$ – мгновенный коэффициент мощности канала.

Фазовый шум может оказывать существенное влияние на точность позиционирования, особенно при высокоточном УМ и ДМ разрешении объекта позиционирования. Принятый сигнал с учетом фазового шума можно представить выражением:

$$\mathbf{y} = \Omega_P \mathbf{H} \mathbf{x} + \mathbf{n}, \quad (79)$$

где диагональная матрица Ω_P содержит комплексные компоненты фазового шума по числу N_r приемных антенных элементов [53]:

$$\Omega_P = \text{diag}([e^{jw_{P,1}}, \dots, e^{jw_{P,N_r}}]). \quad (80)$$

Вектор $\mathbf{w}_P$ вещественных значений фазового шума:

$$\mathbf{w}_P = [w_{P,1}, \dots, w_{P,N_r}] \quad (81)$$

обычно описывается нормальным распределением $\mathbf{w}_P \sim \mathcal{N}(0, \sigma_\theta^2 \mathbf{I}_{N_r})$ с нулевым средним и дисперсией σ_θ^2 в единицах измерения рад².

Ошибка квантования фаз элементов RIS также приводит к искажению сигнала на передаче и приеме с участием реконфигурируемой интеллектуальной поверхности. Фазы w , образующие комплексные коэффициенты матрицы Ω , квантуются таким образом, что набор их весьма ограничен. Так, при двухбитном квантовании фаз набор включает четыре значения $Q = \{0, \pi/2, \pi, 3\pi/2\}$. Квантование на небольшое число уровней позволяет снизить энергопотребление и требования к вычислительной сложности и быстродействию, однако приводит к снижению точности сбора первичных ДМ и УМ измерений.

Другими источниками искажению сигнала на передаче и приеме является дисбаланс синфазной / квадратурной I / Q составляющих комплексного сигнала, нелинейности аналоговых трактов передачи и приема и др. Данные искажения приводят к снижению как пределов точности сбора первичных измерений, так и их вторичной обработки. Если ранее для оценки пределов точности позиционирования использовалась известная метрика нижней границы Крамера – Рао, которая не учитывала несовершенство аппаратного обеспечения, то теперь для учета неизбежного снижения точности позиционирования вследствие описанных выше искажений вводится новая метрика – MCRB (аббр. от англ. Misspecified Cramer-Rao Bound) [54].

4. Заключение

В первой части цикла, посвященного комплексу моделей позиционирования устройств 6G, выполнено обоснование новых факторов построения и функционирования сетей шестого поколения, существенных с точки зрения радиотехнического ОМП в диапазоне терагерцовых волн:

1) переход из диапазона миллиметровых в диапазон терагерцовых волн и использование сверхширокополосных радиоканалов с шириной полосы до единиц гигагерц;

2) переход от систем из десятков элементов в составе единой AP к UM-MIMO-системам из сотен и

тысяч элементов с организацией их в решетке из составных подмассивов;

3) использование RIS в радиолиниях между BS и UE для управления отраженными компонентами;

4) учет SWM и эффекта ближней зоны NF в радиолиниях BS–UE, BS–RIS, RIS–UE.

Данные факторы, отсутствовавшие в сетях 5G и оказывающие принципиальное влияние на разработку моделей позиционирования в сетях 6G, легли в основу построения математической модели сбора первичных измерений для позиционирования устройств в сетях 6G. Во второй части цикла выполняется обзор алгоритмов и приводится оценка точности позиционирования в сетях 6G. Практическая

значимость формализованной математической модели сбора первичных измерений заключается в ее последующей программной реализации в имитационной модели для численного обоснования пределов точности ОК и ориентации устройств в сетях 6G во второй части. Теоретическая значимость сопоставительного анализа, выполненного в первой части цикла, заключается в установлении как технологических препятствий, так и новых возможностей по увеличению точности ОК и ориентации UE. Если сети пятого поколения были призваны достичь точности до одного метра, то сети шестого поколения предлагают повысить ее более, чем на порядок, и приблизиться к погрешности ОК менее одного дециметра.

Список источников

1. Фокин Г.А. Комплекс моделей и методов позиционирования устройств в сетях пятого поколения. Дис. ... докт. техн. наук. СПб.: СПбГУТ, 2021. 499 с. EDN:PQMSQX
2. Фокин Г.А. Технологии сетевого позиционирования. СПб.: СПбГУТ, 2020. 558 с. EDN:PQSMAG
3. Фокин Г.А. Технологии сетевого позиционирования 5G. М.: Горячая Линия – Телеком, 2021. 456 с. EDN:BHFAP1
4. Фокин Г. Эволюция технологий позиционирования в сетях 2G–4G. Часть 1 // Первая миля. 2020. № 2(87). С. 32–39. DOI:10.22184/2070-8963.2020.87.2.32.38. EDN:MYRTVE
5. Фокин Г. Эволюция технологий позиционирования в сетях 2G–4G. Часть 2 // Первая миля. 2020. № 3(88). С. 30–35. DOI:10.22184/2070-8963.2020.88.3.30.35. EDN:WWXGQI
6. Фокин Г.А. Модель технологии сетевого позиционирования метровой точности 5G NR. Часть 1. Конфигурация сигналов PRS // Труды учебных заведений связи. 2022. Т. 8. № 2. С. 48–63. DOI:10.31854/1813-324X-2022-8-2-48-63. EDN:OEXILA
7. Фокин Г.А. Модель технологии сетевого позиционирования метровой точности 5G NR. Часть 2. Обработка сигналов PRS // Труды учебных заведений связи. 2022. Т. 8. № 3. С. 80–99. DOI:10.31854/1813-324X-2022-8-3-80-99. EDN:BRJHYG
8. Дворников С.В., Фокин Г.А., Аль-Одхари А.Х., Федоренко И.В. Оценка влияния свойств сигнала PRS LTE на точность позиционирования // Вопросы радиоэлектроники. Серия: Техника телевидения. 2017. № 4. С. 94–103. EDN:YQWNLJ
9. Дворников С.В., Фокин Г.А., Аль-Одхари А.Х., Федоренко И.В. Исследование зависимости значения геометрического фактора снижения точности от топологии пунктов приема // Вопросы радиоэлектроники. Серия: Техника телевидения. 2018. № 2. С. 99–104. EDN:XRZIXB
10. Фокин Г.А. Сетевое позиционирование 5G и вероятностные модели оценки его точности // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 12. С. 4–17. DOI:10.36724/2072-8735-2020-14-12-4-17. EDN:DQRXIK
11. Фокин Г.А., Кучерявый А.Е. Сетевое позиционирование в экосистеме 5G // Электросвязь. 2020. № 9. С. 51–58. DOI:10.34832/ELSV.2020.10.9.006. EDN:FNNQSH
12. Фокин Г.А. Использование методов сетевого позиционирования в экосистеме 5G // Электросвязь. 2020. № 11. С. 29–37. DOI:10.34832/ELSV.2020.12.11.002. EDN:LKBGPU
13. Лазарев В.О., Фокин Г.А. Оценка точности позиционирования источника радиоизлучения разностно-дальным и угломерным методами. Часть 1 // Труды учебных заведений связи. 2019. Т. 5. № 2. С. 88–100. DOI:10.31854/1813-324X-2019-5-2-88-100. EDN:FFMJWI
14. Фокин Г.А., Лазарев В.О. Оценка точности позиционирования источника радиоизлучения разностно-дальным и угломерным методами. Часть 2. 2D-моделирование // Труды учебных заведений связи. 2019. Т. 5. № 4. С. 65–78. DOI:10.31854/1813-324X-2019-5-4-65-78. EDN:RJHISC
15. Фокин Г.А., Лазарев В.О. Оценка точности позиционирования источника радиоизлучения разностно-дальным и угломерным методами. Часть 3. 3D-моделирование // Труды учебных заведений связи. 2020. Т. 6. № 2. С. 87–102. DOI:10.31854/1813-324X-2020-6-2-87-102. EDN:FKSYIZ
16. Фокин Г.А. Процедуры позиционирования в сетях 5G // Вестник связи. 2021. № 11. С. 2–8. EDN:DEFMNY
17. Фокин Г.А. Методика идентификации прямой видимости в радиолиниях сетей мобильной связи 4-го поколения с пространственной обработкой сигналов // Труды Научно-исследовательского института радио. 2013. № 3. С. 78–82. EDN:RVFDCV
18. Фокин Г.А. Имитационное моделирование процесса распространения радиоволн в радиолиниях сетей мобильной связи 4-го поколения с пространственной обработкой сигналов // Труды Научно-исследовательского института радио. 2013. № 3. С. 83–89. EDN:RVFDDF
19. Shahmansoori A., Garcia G.E., Seco-Granados G., Wymeersch H. 5G Position and Orientation Estimation through Millimeter Wave MIMO // Proceedings of the 2015 IEEE Globecom Workshops (GC Wkshps, San Diego, USA, 06–10 December 2015). IEEE, 2015. DOI:10.1109/GLOCOMW.2015.7413967

20. Shahmansoori A., Garcia G.E., Destino G., Seco-Granados G., Wymeersch H. Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems // IEEE Transactions on Wireless Communications. 2018. Vol. 17. Iss. 3. PP. 1822–1835. DOI:10.1109/TWC.2017.2785788
21. Talvitie J., Valkama M., Destino G., Wymeersch H. Novel Algorithms for High-Accuracy Joint Position and Orientation Estimation in 5G mmWave Systems // Proceedings of the 2017 IEEE Globecom Workshops (GC Wkshps, Singapore, 04–08 December 2017). IEEE, 2017. DOI:10.1109/GLOCOMW.2017.8269069
22. Abu-Shaban Z., Zhou X., Abhayapala T., Seco-Granados G., Wymeersch H. Error Bounds for Uplink and Downlink 3D Localization in 5G Millimeter Wave Systems // IEEE Transactions on Wireless Communications. 2018. Vol. 17. Iss. 8. PP. 4939–4954. DOI:10.1109/TWC.2018.2832134
23. Abu-Shaban Z., Wymeersch H., Abhayapala T., Seco-Granados G. Single-Anchored Two-Way Localization Bounds for 5G mmWave Systems // IEEE Transactions on Vehicular Technology. 2020. Vol. 69. Iss. 6. PP. 6388–6400. DOI:10.1109/TVT.2020.2987039
24. Guidi F., Guerra A., Dardari D. Personal Mobile Radars with Millimeter-Wave Massive Arrays for Indoor Mapping // IEEE Transactions on Mobile Computing. 2016. Vol. 15. Iss. 6. PP. 1471–1484. DOI:10.1109/TMC.2015.2467373
25. Guerra A., Guidi F., Dardari D. Position and orientation error bound for wideband massive antenna arrays // Proceedings of the International Conference on Communication Workshop (ICCW, London, UK, 08–12 June 2015). IEEE, 2015. PP. 853–858. DOI:10.1109/ICCW.2015.7247282
26. Guerra A., Guidi F., Dardari D. Single-Anchored Localization and Orientation Performance Limits Using Massive Arrays: MIMO vs. Beamforming // IEEE Transactions on Wireless Communications. 2018. Vol. 17. Iss. 8. PP. 5241–5255. DOI:10.1109/TWC.2018.2840136
27. Alsabah M., Naser M.A., Mahmmoud B.M., Abdulhussain S.H., Eissaet M.R., Al-Baidhanial A., et al. 6G Wireless Communications Networks: A Comprehensive Survey // IEEE Access. 2021. Vol. 9. PP. 148191–148243. DOI:10.1109/ACCESS.2021.3124812
28. Tataria H., Shafi M., Molisch A.F., Dohler M., Sjöland H., Tufvesson F. 6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities // Proceedings of the IEEE. 2021. Vol. 109. Iss. 7. PP. 1166–1199. DOI:10.1109/JPROC.2021.3061701
29. Jiang W., Han B., Habibi M.A., Schotten H.D. The Road Towards 6G: A Comprehensive Survey // IEEE Open Journal of the Communications Society. 2021. Vol. 2. PP. 334–366. DOI:10.1109/OJCOMS.2021.3057679
30. De Lima C., Belot D., Berkvens R., Bourdoux A., Dardari D., Guillaud M., Isomursu M., et al. Convergent Communication, Sensing and Localization in 6G Systems: An Overview of Technologies, Opportunities and Challenges // IEEE Access. 2021. Vol. 9. PP. 26902–26925. DOI:10.1109/ACCESS.2021.3053486
31. Liu F., Cui Y., Masouros C., Xu J., Han T.X., Eldar Y.C., Buzzi S., et al. Integrated Sensing and Communications: Toward Dual-Functional Wireless Networks for 6G and Beyond // IEEE Journal on Selected Areas in Communications. 2022. Vol. 40. Iss. 6. PP. 1728–1767. DOI:10.1109/JSAC.2022.3156632
32. Wymeersch H., Pärssinen A., Abrudan T.E., Wolfgang A., Haneda K., Sarajlic M. et al. 6G Radio Requirements to Support Integrated Communication, Localization, and Sensing // Proceedings of the Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit, Grenoble, France, 07–10 June 2022). IEEE, 2022. PP. 463–469. DOI:10.1109/EuCNC/6GSummit54941.2022.9815783
33. Wymeersch H., Shrestha D., de Lima C.M., Jaynarayana V., Richerzhagen B., Keskin M.F., et al. Integration of Communication and Sensing in 6G: a Joint Industrial and Academic Perspective // Proceedings of the 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC, Helsinki, Finland, 13–16 September 2021). IEEE, 2021. PP. 1–7. DOI:10.1109/PIMRC50174.2021.9569364
34. Wymeersch H., Seco-Granados G. Radio Localization and Sensing–Part I: Fundamentals // IEEE Communications Letters. 2022. Vol. 26. Iss. 12. PP. 2816–2820. DOI:10.1109/LCOMM.2022.3206821
35. Wymeersch H., Seco-Granados G. Radio Localization and Sensing–Part II: State-of-the-Art and Challenges // IEEE Communications Letters. 2022. Vol. 26. Iss. 12. PP. 2821–2825. DOI:10.1109/LCOMM.2022.3206846
36. González-Prelcic N., Keskin M.F., Kalliokallio O., Valkama M., Dardari D., Shen X., et al. The Integrated Sensing and Communication Revolution for 6G: Vision, Techniques, and Applications // Proceedings of the IEEE. 2024. DOI:10.1109/JPROC.2024.3397609
37. Behravan A., Jaynarayana V., Keskin M.F., Chen H., Shrestha D., Abrudan T.E., et al. Positioning and Sensing in 6G: Gaps, Challenges, and Opportunities // IEEE Vehicular Technology Magazine. 2023. Vol. 18. Iss. 1. PP. 40–48. DOI:10.1109/MVT.2022.3219999
38. Zheng P., Ballal T., Chen H., Wymeersch H., Al-Naffouri T.Y. Localization Coverage Analysis of THz Communication Systems with a 3D Array // Proceedings of the Global Communications Conference (GLOBECOM, Rio de Janeiro, Brazil, 04–08 December 2022). IEEE, 2022. PP. 5378–5383. DOI:10.1109/GLOBECOM48099.2022.10000653
39. Zheng P., Ballal T., Chen H., Wymeersch H., Al-Naffouri T.Y. Coverage Analysis of Joint Localization and Communication in THz Systems With 3D Arrays // IEEE Transactions on Wireless Communications. 2024. Vol. 23. Iss. 5. PP. 5232–5247. DOI:10.1109/TWC.2023.3325192
40. Jaynarayana V., Wymeersch H. Multistatic Sensing of Passive Targets Using 6G Cellular Infrastructure // Proceedings of the Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit, Gothenburg, Sweden, 06–09 June 2023). 2023. PP. 132–137. DOI:10.1109/EuCNC/6GSummit58263.2023.10188243
41. Mateos-Ramos J.M., Song J., Wu Y., Häger C., Keskin M.F., Jaynarayana V., et al. End-to-End Learning for Integrated Sensing and Communication // Proceedings of the International Conference on Communications (Seoul, Republic of Korea, 16–20 May 2022). IEEE, 2022. PP. 1942–1947. DOI:10.1109/ICC45855.2022.9838308

42. Rivetti S., Mateos-Ramos J.M., Wu Y., Song J., Keskin M.F., Yajnanarayana V., et al. Spatial Signal Design for Positioning via End-to-End Learning // IEEE Wireless Communications Letters. 2023. Vol. 12. Iss. 3. PP. 525–529. DOI:10.1109/LWC.2022.3233475
43. Huang C., Hu S., Alexandropoulos G.C., Zappone A., Zappone A., Yuen C., et al. Holographic MIMO Surfaces for 6G Wireless Networks: Opportunities, Challenges, and Trends // IEEE Wireless Communications. 2020. Vol. 27. Iss. 5. PP. 118–125. DOI:10.1109/MWC.001.1900534
44. Elzanaty A., Guerra A., Guidi F., Dardari D., Alouini M.-S. Toward 6G Holographic Localization: Enabling Technologies and Perspectives // IEEE Internet of Things Magazine. 2023. Vol. 6. Iss. 3. PP. 138–143. DOI:10.1109/IOTM.001.2200218
45. Basar E., Yildirim I., Kilinc F. Indoor and Outdoor Physical Channel Modeling and Efficient Positioning for Reconfigurable Intelligent Surfaces in mmWave Bands // IEEE Transactions on Communications. 2021. Vol. 69. Iss. 12. PP. 8600–8611. DOI:10.1109/TCOMM.2021.3113954
46. He J., Jiang F., Keykhosravi K., Kokkonen J., Wymeersch H., Juntti M. Beyond 5G RIS mmWave Systems: Where Communication and Localization Meet // IEEE Access. 2022. Vol. 10. PP. 68075–68084. DOI:10.1109/ACCESS.2022.3186510
47. Киреев А.В., Фокин Г.А. Оценка точности локального позиционирования мобильных устройств с помощью радиокарт и инерциальной навигационной системы // Труды учебных заведений связи. 2017. Т. 3. № 4. С. 54–62. EDN:YMIHOI
48. Фокин Г.А., Владыко А.Г. Позиционирование транспортных средств в сверхплотных сетях радиодоступа V2X/5G с использованием расширенного фильтра Калмана // Труды учебных заведений связи. 2020. Т. 6. № 4. С. 45–59. DOI:10.31854/1813-324X-2020-6-4-45-59. EDN:PYHUMZ
49. Фокин Г.А., Владыко А.Г. Позиционирование транспортных средств с комплексированием дальномерных, угломерных и инерциальных измерений в расширенном фильтре Калмана // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 51–67. DOI:10.31854/1813-324X-2021-7-2-51-67. EDN:AIEESO
50. Фокин Г.А. Процедуры выравнивания лучей устройств 5G NR // Электросвязь. 2022. № 2. С. 26–31. DOI:10.34832/ELSV.2022.27.2.003. EDN:GWPZQH
51. Фокин Г.А. Модели управления лучом в сетях 5G NR. Часть 1. Выравнивание лучей при установлении соединения // Первая миля. 2022. № 1(101). С. 42–49. DOI:10.22184/2070-8963.2022.101.1.42.49. EDN:UVALJF
52. Фокин Г. Модели управления лучом в сетях 5G NR. Часть 2. Выравнивание лучей при ведении радиосвязи // Первая миля. 2022. № 3(103). С. 62–69. DOI:10.22184/2070-8963.2022.103.3.62.68. EDN:PTALDP
53. Chen H., Srieddeen H., Ballal T., Wymeersch H., Alouini M.-S., Al-Naffouri T.Y. A Tutorial on Terahertz-Band Localization for 6G Communication Systems // IEEE Communications Surveys & Tutorials. 2022. Vol. 24. Iss. 3. PP. 1780–1815. DOI:10.1109/COMST.2022.3178209
54. Chen H., Aghdam S.R., Keskin M.F., Wu Y., Lindberg S., Wolfgang A., et al. MCRB-based Performance Analysis of 6G Localization under Hardware Impairments // Proceedings of the International Conference on Communications Workshops (ICC Workshops, Seoul, Republic of Korea, 16–20 May 2022). IEEE, 2022. PP. 115–120. DOI:10.1109/ICCWorkshops53468.2022.9814598

References

1. Fokin G.A. *A set of Models and Methods for Positioning Devices in Fifth-Generation Networks*. D.Sc Thesis. St. Petersburg: The Bonch-Bruевич Saint-Petersburg State University of Telecommunications Publ.; 2021. 499 p. (in Russ.) EDN:PQMSQX
2. Fokin G.A. *Technologies of Network Positioning*. St. Petersburg: The Bonch-Bruевич State University of Telecommunications Publ.; 2020. 558 p. (in Russ.) EDN:PQSMAG
3. Fokin G.A. *5G Network Positioning Technologies*. Moscow: Hot Line – Telecom Publ.; 2021. 456 p. (in Russ.) EDN:BHFAP1
4. Fokin G. Evolution of positioning technologies in 2G-4G networks. Part 1. *Last mile*. 2020;2(87):32–39. (in Russ.) DOI:10.22184/2070-8963.2020.87.2.32.38. EDN:MYRTVE
5. Fokin G. Evolution of positioning technologies in 2G-4G networks. Part 2. *Last mile*. 2020;3(88):30–35. (in Russ.) DOI:10.22184/2070-8963.2020.88.3.30.35. EDN:WWXGQI
6. Fokin G. Model of 5G NR Precision Metro Network Positioning Technology. Part 1. Configuration of PRS Signals. *Proceedings of Telecommunication Universities*. 2022;8(2):48–63. DOI:10.31854/1813-324X-2022-8-2-48-63 (in Russ.) EDN:OEXILA
7. Fokin G. Model of 5G NR Precision Metro Network Positioning Technology. Part 2. PRS Signal Processing. *Proceedings of Telecommunication Universities*. 2022;8(3):80–99. DOI:10.31854/1813-324X-2022-8-3-80-99 (in Russ.) EDN:BRJHYG
8. Dvornikov S.V., Fokin G.A., Al-Odhari A.Kh., Fedorenko I.V. Assessing the influence of PRS LTE signal properties on positioning accuracy. *Voprosy radioelektroniki. Seriya: Tekhnika teledeniya*. 2017;4:94–103. (in Russ.) EDN:YQWNLJ
9. Dvornikov S.V., Fokin G.A., Al-Odhari A.Kh., Fedorenko I.V. Study of the dependence of the value of the geometric factor of reducing accuracy on the topology of receiving points. *Voprosy radioelektroniki. Seriya: Tekhnika teledeniya*. 2018;2:99–104. (in Russ.) EDN:XRZIXB
10. Fokin G.A. 5G network positioning and probabilistic models for assessing its accuracy. *T-Comm*. 2020;14(12):4–17. (in Russ.) DOI:10.36724/2072-8735-2020-14-12-4-17. EDN:DQRXIK
11. Fokin G.A., Kucheryavy A.E. Network positioning in the 5G ecosystem. *Elektrosvyaz*. 2020;9:51–58. (in Russ.) DOI:10.34832/ELSV.2020.10.9.006. EDN:FNHQSH
12. Fokin G.A. Using network positioning methods in the 5G ecosystem. *Elektrosvyaz*. 2020;11:29–37. (in Russ.) DOI:10.34832/ELSV.2020.12.11.002. EDN:LKBGPU
13. Lazarev V., Fokin G. Positioning Accuracy Evaluation of Radio Emission Sources Using Time Difference of Arrival and Angle of Arrival Methods. Part 1. *Proceedings of Telecommunication Universities*. 2019;5(2):88–100. (in Russ.) DOI:10.31854/1813-324X-2019-5-2-88-100. EDN:FFMJWI

14. Fokin G., Lazarev V. Positioning Accuracy Evaluation of Radio Emission Sources Using Time Difference of Arrival and Angle of Arrival Methods. Part 2. 2D-Simulation. *Proceedings of Telecommunication Universities*. 2019;5(4):65–78. (in Russ.) DOI:10.31854/1813-324X-2019-5-4-65-78. EDN:RJHISC
15. Fokin G., Lazarev V. Positioning Accuracy Evaluation of Radio Emission Sources Using Time Difference of Arrival and Angle of Arrival Methods. Part 3. 3D-Simulation. *Proceedings of Telecommunication Universities*. 2020;6(2):87–102. (in Russ.) DOI:10.31854/1813-324X-2020-6-2-87-102. EDN:FKSYIZ
16. Fokin G.A. Positioning procedures in 5G networks. *Vestnik sviazy*. 2021;11:2–8. (in Russ.) EDN:DEFMNY
17. Fokin G.A. Methodology for identifying line of sight in radio links of 4th generation mobile communication networks with spatial signal processing. *Proceedings of the Radio Research Institute*. 2013;3:78–82. (in Russ.) EDN:RVFDCV
18. Fokin G.A. Simulation modeling of the process of radio wave propagation in radio links of 4th generation mobile communication networks with spatial signal processing. *Proceedings of the Radio Research Institute*. 2013;3:83–89. (in Russ.) EDN:RVFDDF
19. Shahmansoori A., Garcia G.E., Destino G., Seco-Granados G., Wymeersch H. 5G Position and Orientation Estimation through Millimeter Wave MIMO. *Proceedings of the 2015 IEEE Globecom Workshops*, GC Wkshps, 06–10 December 2015, San Diego, USA. IEEE; 2015. DOI:10.1109/GLOCOMW.2015.7413967
20. Shahmansoori A., Garcia G.E., Destino G., Seco-Granados G., Wymeersch H. Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems. *IEEE Transactions on Wireless Communications*. 2018;17(3):1822–1835. DOI:10.1109/TWC.2017.2785788
21. Talvitie J., Valkama M., Destino G., Wymeersch H. Novel Algorithms for High-Accuracy Joint Position and Orientation Estimation in 5G mmWave Systems. *Proceedings of the 2017 IEEE Globecom Workshops*, GC Wkshps, 04–08 December 2017, Singapore. IEEE; 2017. DOI:10.1109/GLOCOMW.2017.8269069
22. Abu-Shaban Z., Zhou X., Abhayapala T., Seco-Granados G., Wymeersch H. Error Bounds for Uplink and Downlink 3D Localization in 5G Millimeter Wave Systems. *IEEE Transactions on Wireless Communications*. 2018;17(8):4939–4954. DOI:10.1109/TWC.2018.2832134
23. Abu-Shaban Z., Wymeersch H., Abhayapala T., Seco-Granados G. Single-Anchor Two-Way Localization Bounds for 5G mmWave Systems. *IEEE Transactions on Vehicular Technology*. 2020;69(6):6388–6400. DOI:10.1109/TVT.2020.2987039
24. Guidi F., Guerra A., Dardari D. Personal Mobile Radars with Millimeter-Wave Massive Arrays for Indoor Mapping. *IEEE Transactions on Mobile Computing*. 2016;15(6):1471–1484. DOI:10.1109/TMC.2015.2467373
25. Guerra A., Guidi F., Dardari D. Position and orientation error bound for wideband massive antenna arrays. *Proceedings of the International Conference on Communication Workshop*, ICCW, 08–12 June 2015. IEEE; 2015. p.853–858. DOI:10.1109/ICCW.2015.7247282
26. Guerra A., Guidi F., Dardari D. Single-Anchor Localization and Orientation Performance Limits Using Massive Arrays: MIMO vs. Beamforming. *IEEE Transactions on Wireless Communications*. 2018;17(8):5241–5255. DOI:10.1109/TWC.2018.2840136
27. Alsabah M., Naser M.A., Mahmmoud B.M., Abdulhussain S.H., Eissaet M.R., Al-Baidhanial A., et al. 6G Wireless Communications Networks: A Comprehensive Survey. *IEEE Access*. 2021;9:148191–148243. DOI:10.1109/ACCESS.2021.3124812
28. Tataria H., Shafi M., Molisch A.F., Dohler M., Sjöland H., Tufvesson F. 6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities. *Proceedings of the IEEE*. 2021;109(7):1166–1199. DOI:10.1109/JPROC.2021.3061701
29. Jiang W., Han B., Habibi M.A., Schotten H.D. The Road Towards 6G: A Comprehensive Survey. *IEEE Open Journal of the Communications Society*. 2021;2:334–366. DOI:10.1109/OJCOMS.2021.3057679
30. De Lima C., Belot D., Berkvens R., Bourdoux A., Dardari D., Guillaud M., Isomursu M., et al. Convergent Communication, Sensing and Localization in 6G Systems: An Overview of Technologies, Opportunities and Challenges. *IEEE Access*. 2021;9:26902–26925. DOI:10.1109/ACCESS.2021.3053486
31. Liu F., Cui Y., Masouros C., Xu J., Han T.X., Eldar Y.C., Buzzi S., et al. Integrated Sensing and Communications: Toward Dual-Functional Wireless Networks for 6G and Beyond. *IEEE Journal on Selected Areas in Communications*. 2022;40(6):1728–1767. DOI:10.1109/JSAC.2022.3156632
32. Wymeersch H., Pärssinen A., Abrudan T.E., Wolfgang A., Haneda K, Sarajlic M. et al. 6G Radio Requirements to Support Integrated Communication, Localization, and Sensing. *Proceedings of the Joint European Conference on Networks and Communications & 6G Summit, EuCNC/6G Summit, 07–10 June 2022, Grenoble, France*. IEEE; 2022. p.463–469. DOI:10.1109/EuCNC/6GSummit54941.2022.9815783
33. Wymeersch H., Shrestha D., de Lima C.M., Yajnanarayana V., Richerzhagen B., Keskin M.F., et al. Integration of Communication and Sensing in 6G: a Joint Industrial and Academic Perspective. *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, Helsinki, Finland, 13–16 September 2021*. IEEE; 2021. p.1–7. DOI:10.1109/PIMRC50174.2021.9569364
34. Wymeersch H., Seco-Granados G. Radio Localization and Sensing–Part I: Fundamentals. *IEEE Communications Letters*. 2022;26(12):2816–2820. DOI:10.1109/LCOMM.2022.3206821
35. Wymeersch H., Seco-Granados G. Radio Localization and Sensing–Part II: State-of-the-Art and Challenges. *IEEE Communications Letters*. 2022;26(12):2821–2825. DOI:10.1109/LCOMM.2022.3206846
36. González-Prelcic N., Keskin M.F., Kaltiokallio O., Valkama M., Dardari D., Shen X., et al. The Integrated Sensing and Communication Revolution for 6G: Vision, Techniques, and Applications. *Proceedings of the IEEE*. 2024. DOI:10.1109/JPROC.2024.3397609
37. Behravan A., Yajnanarayana V., Keskin M.F., Chen H., Shrestha D., Abrudan T.E., et al. Positioning and Sensing in 6G: Gaps, Challenges, and Opportunities. *IEEE Vehicular Technology Magazine*. 2023;18(1):40–48. DOI:10.1109/MVT.2022.3219999

38. Zheng P., Ballal T., Chen H., Wymeersch H., Al-Naffouri T.Y. Localization Coverage Analysis of THz Communication Systems with a 3D Array. *Proceedings of the Global Communications Conference, GLOBECOM, 04–08 December 2022, Rio de Janeiro, Brazil*. IEEE; 2022. p.5378–5383. DOI:10.1109/GLOBECOM48099.2022.10000653
39. Zheng P., Ballal T., Chen H., Wymeersch H., Al-Naffouri T.Y. Coverage Analysis of Joint Localization and Communication in THz Systems With 3D Arrays. *IEEE Transactions on Wireless Communications*. 2024;23(5):5232–5247. DOI:10.1109/TWC.2023.3325192
40. Yajnanarayana V., Wymeersch H. Multistatic Sensing of Passive Targets Using 6G Cellular Infrastructure. *2023 Joint European Conference on Networks and Communications & 6G Summit, EuCNC/6G Summit, Gothenburg, Sweden, 06–09 June 2023*. p.132–137. DOI:10.1109/EuCNC/6GSummit58263.2023.10188243
41. Mateos-Ramos J.M., Song J., Wu Y., Häger C., Keskin M.F., Yajnanarayana V., et al. End-to-End Learning for Integrated Sensing and Communication. *Proceedings of the International Conference on Communications, 16–20 May 2022, Seoul, Republic of Korea*. IEEE; 2022. p.1942–1947. DOI:10.1109/ICC45855.2022.9838308
42. Rivetti S., Mateos-Ramos J.M., Wu Y., Song J., Keskin M.F., Yajnanarayana V., et al. Spatial Signal Design for Positioning via End-to-End Learning. *IEEE Wireless Communications Letters*. 2023;12(3):525–529. DOI:10.1109/LWC.2022.3233475
43. Huang C., Hu S., Alexandropoulos G.C., Zappone A., Zappone A., Yuen C., et al. Holographic MIMO Surfaces for 6G Wireless Networks: Opportunities, Challenges, and Trends. *IEEE Wireless Communications*. 2020;27(5):118–125. DOI:10.1109/MWC.001.1900534
44. Elzanaty A., Guerra A., Guidi F., Dardari D., Alouini M.-S. Toward 6G Holographic Localization: Enabling Technologies and Perspectives. *IEEE Internet of Things Magazine*. 2023;6(3):138–143. DOI:10.1109/IOTM.001.2200218
45. Basar E., Yildirim I., Kilinc F. Indoor and Outdoor Physical Channel Modeling and Efficient Positioning for Reconfigurable Intelligent Surfaces in mmWave Bands. *IEEE Transactions on Communications*. 2021;69(12):8600–8611. DOI:10.1109/TCOMM.2021.3113954
46. He J., Jiang F., Keykhosravi K., Kokkonen J., Wymeersch H., Juntti M. Beyond 5G RIS mmWave Systems: Where Communication and Localization Meet. *IEEE Access*. 2022;10:68075–68084. DOI:10.1109/ACCESS.2022.3186510
47. Kireev A., Fokin G. Accuracy Evaluation of Local Positioning by Radiomap Building and Inertial Navigation System. *Proceedings of Telecommunication Universities*. 2017;3(4):54–62. (in Russ.) EDN:YMIHOI
48. Fokin G., Vladiko A. The Vehicles Positioning in Ultra-Dense 5G/V2X Radio Access Networks Using the Extended Kalman Filter. *Proceedings of Telecommunication Universities*. 2020;6(4):45–59. (in Russ.) DOI:10.31854/1813-324X-2020-6-4-45-59. EDN:PYHUMZ
49. Fokin G., Vladiko A. Positioning of Vehicles with the Fusion of Time of Arrival, Angle of Arrival and Inertial Measurements in the Extended Kalman Filter. *Proceedings of Telecommunication Universities*. 2021;7(2):51–67. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-51-67. EDN:AIEESO
50. Fokin G.A. Beam alignment procedures for 5G NR devices. *Elektrosvyaz*. 2022;2:26–31. (in Russ.) DOI:10.34832/ELSV.2022.27.2.003. EDN:UVALJF
51. Fokin G.A. Beam steering models in 5G NR networks. Part 1. Beam alignment when establishing a connection. *Last mile*. 2022;1(101):42–49. (in Russ.) DOI:10.22184/2070-8963.2022.101.1.42.49. EDN:PTALDP
52. Fokin G. Beam control models in 5G NR networks. Part 2. Alignment of beams during radio communication. *Last mile*. 2022;3(103):62–69. DOI:10.22184/2070-8963.2022.103.3.62.68
53. Chen H., Sardeddeen H., Ballal T., Wymeersch H., Alouini M.-S., Al-Naffouri T.Y. A Tutorial on Terahertz-Band Localization for 6G Communication Systems. *IEEE Communications Surveys & Tutorials*. 2022;24(3):1780–1815. DOI:10.1109/COMST.2022.3178209
54. Chen H., Aghdam S.R., Keskin M.F., Wu Y., Lindberg S., Wolfgang A., et al. MCRB-based Performance Analysis of 6G Localization under Hardware Impairments // Proceedings of the International Conference on Communications Workshops, ICC Workshops, 16–20 May 2022, Seoul, Republic of Korea. IEEE; 2022. p.115–120. DOI:10.1109/ICCWorkshops53468.2022.9814598

Статья поступила в редакцию 09.07.2024; одобрена после рецензирования 23.08.2024; принята к публикации 28.08.2024.

The article was submitted 01.07.2024; approved after reviewing 23.08.2024; accepted for publication 28.08.2024.

Информация об авторе:

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>

Автор сообщает об отсутствии конфликтов интересов.

The author declares 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**

Обзорная статья

УДК 621.391

<https://doi.org/10.31854/1813-324X-2024-10-4-100-109>

Энергоэффективный алгоритм выбора маршрута передачи данных в беспроводных сенсорных сетях высокой плотности

Татьяна Николаевна Астахова¹, ctn_af@mail.ru
Михаил Олегович Колбанёв², mokolbanev@mail.ru
Александр Сергеевич Лямин¹, a.s.lyamin@gmail.com
Никита Сергеевич Маслов¹, j-knaginino@yandex.ru
Дарья Александровна Маслова¹✉, dasha.kirilova.96@bk.ru

¹Нижегородский государственный инженерно-экономический университет,
Княгинино, 606340, Российская Федерация

²Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»,
Санкт-Петербург, 197376, Российская Федерация

Аннотация

Актуальность. Повышение числа сенсорных устройств на единицу площади, как следствие, сокращает физическое расстояние между устройствами сенсорной сети. Такие сети, как правило, разворачиваются на большой площади и сенсорное устройство, которое хочет передать пакет данных, располагается далеко от базовой станции. В таком случае перед устройством-источником ставится задача выбрать такой путь передачи, при котором затрачивается наименьшее количество энергетических ресурсов и удовлетворяет требованиям времени доставки.

Целью настоящего исследования является разработка и обоснование эффективности применения эмпирического алгоритма выбора маршрута передачи данных, снижающего энергопотребление беспроводных сенсорных сетей высокой плотности. Используются **методы** системного анализа, аналитического моделирования, геометрии и теорий вероятностей.

Решение. Предполагается, что сенсорная сеть развернута на ограниченной территории и представляет собой совокупность устройств, которые связаны друг с другом информационно и энергетически. При построении маршрутов передачи данных допускается использование любых сенсорных устройств в качестве ретрансляторов.

Новизна. Разработана новая математическая модель, основанная на теореме косинусов и формуле Фрииса, которая позволяет выбрать рациональный маршрут передачи данных от оконечного устройства на системный координатор по критерию совокупного энергопотребления всеми устройствами, составляющими маршрут.

Значимость (теоретическая). Получены зависимости уровня энергопотребления от различных системных параметров, влияющих на процессы функционирования беспроводных сенсорных сетей высокой плотности.

Значимость (практическая). Предложенный эмпирический алгоритм выбора рационального маршрута передачи данных в беспроводной сенсорной сети позволяет определить среди всех альтернативных тот маршрут к координатору, который требует наименьшей мощности. Эффективность предложенного эмпирического энергосберегающего алгоритма подтверждена имитационным моделированием.

Ключевые слова: энергосбережение, беспроводная сенсорная сеть, алгоритм выбора маршрута, сенсорное устройство, мощность сигнала

Ссылка для цитирования: Астахова Т.Н., Колбанёв М.О., Лямин А.С., Маслов Н.С., Маслова Д.А. Энергоэффективный алгоритм выбора маршрута передачи данных в беспроводных сенсорных сетях высокой плотности // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 100–109. DOI:10.31854/1813-324X-2024-10-4-100-109. EDN:VPHYUT

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-100-109>

Energy-Efficient Algorithm for Data Path Selection in High-Density Wireless Sensor Networks

 **Tatiana N. Astakhova**¹, ctn_af@mail.ru
 **Mikhail O. Kolbanev**², mokolbanev@mail.ru
 **Alexandr S. Lyamin**¹, a.s.lyamin@gmail.com
 **Nikita S. Maslov**¹, j-knaginino@yandex.ru
 **Daria A. Maslova**¹ , dasha.kirilova.96@bk.ru

¹Nizhny Novgorod State Engineering and Economic University,
 Knyaginino, 606340, Russian Federation

²St. Petersburg State Electrotechnical University "LETI",
 St. Petersburg, 197376, Russian Federation

Annotation

Relevance. Increasing the number of sensor devices per unit area consequently reduces the physical distance between the devices in the sensor network. Such networks are usually deployed over a large area and the sensor device that wants to transmit a data packet is located far away from the base station. In such a case, the source device is challenged to choose a transmission path that consumes the least amount of energy resources and satisfies the delivery time requirements.

The objective of this study is to develop and validate the effectiveness of an empirical algorithm for selecting a transmission path that reduces the energy consumption of high-density wireless sensor networks. **Methods** of system analysis, analytical modeling, geometry and probability theories are used.

Solution. It is assumed that the sensor network is deployed in a limited area and is a set of devices that are connected to each other informationally and energetically. When building data transmission routes, any sensor devices can be used as repeaters. At the same time, the increase in the number of repeaters leads to an increase in the time of data delivery.

Novelty. It is assumed that the sensor network is deployed in a limited area and is a set of devices that are connected to each other informationally and energetically. Any sensor devices may be used as repeaters when constructing data transmission routes.

Significance (theoretical). Dependences of power consumption level on various system parameters affecting the processes of functioning of high-density wireless sensor networks have been obtained.

Significance (practical). The proposed empirical algorithm for selecting a rational data transmission route in a wireless sensor network allows us to determine, among all alternatives, the route to the coordinator that requires the least power. The effectiveness of the proposed empirical power-saving algorithm is confirmed by simulation modeling.

Keywords: energy saving, wireless sensor networks, route selection algorithm, sensor device, signal strength

For citation: Astakhova T.N., Kolbanev M.O., Lyamin A.S., Maslov N.S., Maslova D.A. Energy-Efficient Algorithm for Data Path Selection in High-Density Wireless Sensor Networks. *Proceedings of Telecommunication Universities*. 2024;10(4):100–109. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-100-109. EDN:VPHYUT

Введение

Четвертая промышленная революция и цифровая трансформация процессов деятельности невозможны без создания, внедрения и эксплуатации сенсорных сетей, которые собирают данные о состоянии физических или социальных сущностей. Сенсорные сети охватывают большие территории

и комплексно обрабатывают информацию в широком спектре предметных областей. Под сенсорной сетью будем понимать совокупность информационно и энергетически связанных сенсорных устройств (далее – устройств, сенсоров), каждое из которых в реальном масштабе времени проводит измерения тех или иных параметров в относи-

тельно малой окружающей их области естественного, общественного, кибернетического или другого вида пространств и передачу данных о результатах измерений к некоторому централизованному системному координатору (точке доступа, базовой станции, центру обработки данных и т. п.). Координатор тем или иным образом способен сам или при помощи других информационных систем проводить комплексный анализ собранных данных и формировать варианты управленческих решений в интересах субъектов деятельности [1].

В качестве примера сенсорной сети приведем архитектуру широко распространенной технологии энергоэффективных систем дальнего радиуса действия LPWAN (аббр. от англ. Low-Power Wide-Area Network). Это класс технологий с радиальной топологией, разработанных для интернета вещей и нацеленных на передачу небольших по объему данных на дальние расстояния. К числу практических реализаций соответствующей концепции относятся, например, системы LoRaWAN [2], СТРИЖ (<https://strij.tech>), NB-FI [3], SNB [4] и др.

В сетях LPWAN координатором при передаче данных является базовая станция. Она принимает данные от сенсоров через радиальные соединения, отвечает за накопление данных, поступающих от оконечных устройств, и отправку этих данных на серверы облачной инфраструктуры для обработки и хранения. Схематично архитектура систем LPWAN очень проста и может быть представлена так, как показано на рисунке 1.

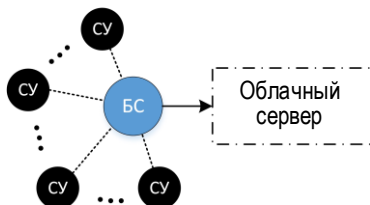


Рис. 1. Архитектура технологии LPWAN

Fig. 1. LPWAN Technology Architecture

Реализация систем LPWAN отличается большим разнообразием:

- данные между сенсорными устройствами и базовой станцией могут передаваться с помощью многих беспроводных технологий;
- базовая станция может располагаться на более или менее высоком здании или вышке, что существенно влияет на дальность связи;
- радиоканалы могут быть организованы как на лицензированных, так и на нелицензируемых диапазонах частот и т. д.

Тенденции развития энергоэффективных технологий дальнего радиуса действия нацелены на преодоление целого ряда ограничений, к числу которых относятся:

- увеличение пропускной способности системы за счет создания более эффективных алгоритмов множественного доступа сенсорных устройств к частотному ресурсу каналов связи;
- привлечение технологий сотовых сетей или сетей другого типа для расширения зоны покрытия;
- внедрение все более надежных технологий безопасности и многое другое.

Однако главным трендом развития систем LPWAN остается стремление к экономии той энергии, которая требуется сенсорным устройствам для передачи данных на базовую станцию. Очевидно, что потребляемая энергия при прочих равных условиях увеличивается из-за роста интенсивности информационного взаимодействия устройств, расстояния от сенсорных устройств до базовой станции и неблагоприятных условий распространения сигналов между удаленными элементами системы.

В данной работе для экономии энергетических ресурсов отдельных устройств и всей системы в целом допускается возможность создания маршрутов передачи данных от устройства-источника на координатор через последовательность нескольких сенсорных устройств системы, которые выступают в качестве ретрансляторов. Это означает, что система в целом для построения маршрутов должна реализовать некоторые алгоритмы, использующие в качестве исходных данных не только расстояние до базовой станции, но и расстояния между любой парой сенсорных устройств и методы оценки энергетических характеристик процесса передачи информации между ними.

Материалы и методы

Для рассматриваемой сенсорной сети первостепенное значение приобретают характеристики связей между всей совокупностью сенсорных устройств. Они зависят от множества внешних и внутренних факторов, таких как:

- масштаб зоны охвата сенсорной сети и количество сенсорных устройств в этой зоне;
- законы распределения и движения сенсоров в пространстве;
- требования ко времени доставки результатов измерений к координатору и энергопотребление этого процесса, а также многое другое.

Будем далее предполагать, что система работает в определенных условиях.

Во-первых, сенсорная сеть объединяет устройства разной конструкции, типа и назначения (рисунок 2).

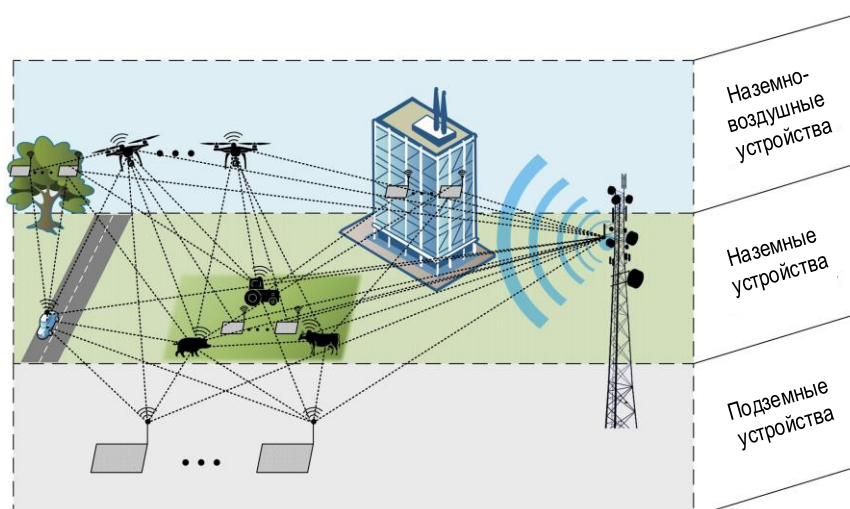


Рис. 2. Типы сенсорных устройств

Fig. 2. Types of Sensor Devices

Во-вторых, сенсорные устройства имеют автономные источники электропитания и по этой причине энергозатраты устройств относятся к важнейшим параметрам системы. Это, в частности, приводит к тому, что время жизни сенсорной сети в целом зависит от емкости электрических батарей отдельных устройств. В работе [5] выявлены условия связности системы, которая может нарушаться из-за потери энергопитания частью устройств.

В-третьих, критерием энергоэффективности системы является общая мощность, потребляемая всеми сенсорами построенного маршрута при передаче сообщения от оконечных сенсорных устройств (источников данных) до координатора. Чем меньше эта мощность, тем эффективнее система в целом.

В-четвертых, энергозатраты связаны, в основном, с передачей сигналов электросвязи по беспроводной сети между сенсорными устройствами. Они зависят от мощности, потребляемой радиопередатчиками и приемниками сигналов электросвязи, и времени занятия радиоканала [6].

В-пятых, устройства передают данные внутри системы с помощью беспроводных технологий связи. Мощность сигнала, необходимая для надежной доставки данных от излучающей на принимающую антенны взаимодействующих устройств, определяется эмпирическими законами, рассмотренными в работе [4], в частности, формулой Фрииса [7].

В-шестых, допустимое время доставки данных к координатору от сенсорных устройств ограничено. При превышении допустимого времени данные «стареют» и их привлечение для дальнейшего анализа становится нецелесообразным. Информация разных устройств может иметь разное время старения. В общем случае время доставки складывается из времени предварительной обработки данных

устройствами, времени доступа к общей для устройств среде передачи, времени распространения сигналов и других временных задержек. Далее будем оценивать время доставки числом ретрансляций, которые использовались в процессе передачи данных координатору, пренебрегая другими параметрами.

В-седьмых, сенсорная сеть состоит из достаточно большого числа устройств, размещенных в ограниченном пространстве так, что систему в целом можно считать плотной. Количественная оценка плотности – это число устройств, расположенных в среднем в единице площади или объема контролируемого пространства.

В-восьмых, в общем случае каждое устройство имеет возможность установить соединение по радиоканалу с любым другим устройством для обмена информацией напрямую без участия транзитных устройств, однако внутри сенсорной сети могут существовать запреты на некоторые из связей, вызванные ограничениями протоколов взаимодействия, пропускной способности каналов связи, остаточной мощности батарей устройств, условиями распространения радиосигналов и др., что ведет к запрету на передачу данных между некоторыми парами устройств.

Сделанные предположения позволяют описать топологические, геометрические, информационные и энергетические связи внутри сенсорной сети в виде нагруженного графа, общий вид которого представлен на рисунке 3. Здесь вершины соответствуют отдельным устройствам системы, а существование ребра указывает на возможность передачи данных между соответствующими устройствами. Нагрузка на ребра представляет собой весь набор физических ресурсов, необходимых цифровым системам для выполнения своих функций: рас-

стояние (дистанция) между устройствами, пропускная способность соответствующего канала связи, информационный объем передаваемых сообщений, энергетические характеристики, необходимые для реализации процесса передачи [8].

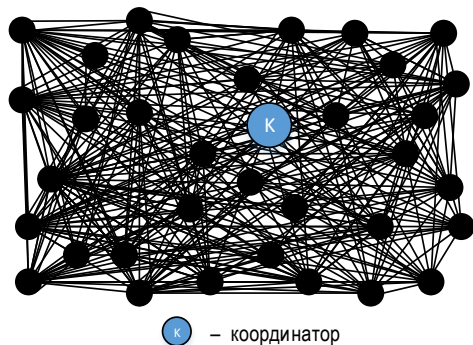


Рис. 3. Фрагмент графа, отображающего сенсорную сеть

Fig. 3. A Fragment of a Graph Showing the Sensory Network

При реализации такой архитектуры сенсорной сети возникает задача создания эффективных алгоритмов поиска и установления маршрутов передачи данных. С учетом сделанных выше предположений о свойствах сенсорной сети задача поиска энергоэффективного маршрута может быть сведена к задаче поиска кратчайшего в энергетическом смысле пути в графе, который отображает топологические и геометрические свойства сенсорной сети. Предлагаемый ниже подход к решению этой задачи может быть реализован при двух условиях:

1) существует механизм вычисления мощности, необходимой для формирования радиосигнала, как функции расстояния между взаимодействующими устройствами, и условий распространения радиосигнала;

2) время доставки данных от устройства источника данных до координатора может быть выражено через число транзитных устройств в маршруте передачи.

Если оба условия выполнены, то ребрам графа можно сопоставить не только расстояния между вершинами, но и мощность, которая необходима для надежной передачи данных, а допустимое время доставки обеспечить ограничением на число ретрансляций.

Особенностью предлагаемого энергоэффективного алгоритма поиска маршрутов передачи данных является использование свойств тех функций, которые связывают энергетические характеристики радиосигнала с дистанцией передачи данных. К числу важных свойств алгоритма следует отнести три обстоятельства:

– алгоритм не гарантирует нахождение самого энергоэффективного маршрута, но предлагаемые им решения всегда являются рациональными;

– по мере повышения плотности устройств в сенсорной сети точность алгоритма растет;

– с вычислительной точки зрения алгоритм существенно проще известных алгоритмов маршрутизации.

В основе предлагаемого алгоритма лежат следующие предпосылки.

В-первых, затухание радиосигнала в среде распространения между передающей и приемной антеннами определяется степенной функцией от расстояния. Например, в соответствии с формулой Фрииса затрачиваемая мощность при передаче радиосигналов зависит от расстояния между взаимодействующими устройствами следующим образом:

$$P_{\text{пер}} = \frac{16P_{\text{пр}}\pi^2 R^2 f^2}{C_{\text{пер}}C_{\text{пр}}v^2}, \quad (1)$$

где $C_{\text{пер}}$ и $C_{\text{пр}}$ – коэффициенты усиления передающей и приемной антенн, соответственно; $P_{\text{пр}}$ – мощность радиосигнала на приемной антенне, Вт (без учета потерь); f – частота радиосигнала, Гц; v_c – скорость света ($\sim 3 \cdot 10^8$ м/с).

Во-вторых, следствием из формулы Фрииса, устанавливающей квадратичную зависимость мощности сигнала от расстояния между антеннами, является утверждение 1 (рисунок 4).

УТВЕРЖДЕНИЕ 1. При передаче данных от устройства-источника данных A к устройству-приемнику B через транзитное устройство $B2$ общая затрачиваемая мощность будет меньше, чем при передаче данных непосредственно на устройство B без ретрансляции, если только угол $B2$ тупой. Наибольший выигрыш в мощности достигается в том случае, когда ретранслюющее устройство расположено на середине отрезка AB , таким образом, как устройство $B1$. Следствие данного утверждения можно отобразить следующим образом:

$$\begin{aligned} (a1 + a2)^2 &> a1^2 + a2^2, \\ (a1 + a2)^2 &> c^2 + b^2, \\ a1^2 + a2^2 &< c^2 + b^2. \end{aligned} \quad (2)$$

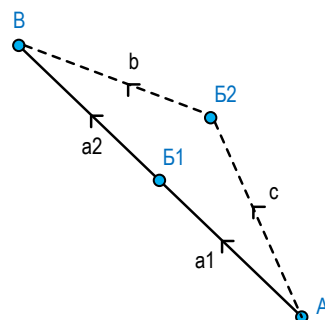


Рис. 4. Следствие квадратичной зависимости мощности от расстояния для сенсорных сетей

Fig. 4. Corollary of the Power Quadratic Distance Dependence for Sensor Network

В [9] для оценки мощности сигнала в специфических условиях предлагается кубическая, четвертичная и даже бóльшая степень зависимости мощности от расстояния. Очевидно, что в этих случаях ретрансляция через правильно выбранные устройства приведет к еще большей экономии заряда батарей электропитания устройств.

В-третьих, в случае, когда на прямой нет сенсорных устройств – кандидатов для ретрансляции, в целях оценки экономии мощности следует использовать теорему косинусов [10], следствием которой является утверждение 2.

УТВЕРЖДЕНИЕ 2. Чем ближе сенсорное устройство, которое можно было бы использовать в качестве ретранслятора, расположено к середине прямой, соединяющей устройство-источник с устройством-приемником, тем бóльшая экономия энергии будет достигнута за счет ретрансляции.

В-четвертых, если принять предположение, что сенсорные устройства распределены в пространстве по закону Пуассона, то вероятность нахождения хотя бы одного устройства в секторе площадью D , такого, что его использование как транзитного уменьшает энергопотребление, определяется формулой Рэлея [11]:

$$P = 1 - e^{-\lambda D}, \quad (3)$$

где λ – плотность распределения сенсорных устройств в пространстве (среднее число устройств на единицу площади).

По мере увеличения плотности сенсорной сети вероятность P достаточно быстро приближается к единице, а это увеличивает возможности экономии энергии за счет транзитов. Заметим, что повышение плотности – это характерная тенденция разви-

тия сенсорных сетей. Так, плотность устройств в сетях 5-го поколения в среднем может составлять 1 устройство на 1 м² (10⁶ устройств на 1 км²), а предполагаемая плотность сетей 6-го поколения – уже 100 устройств на 1 м² (10⁹ устройств на 1 км²).

Алгоритм поиска энергоэффективного маршрута состоит из N этапов. На первом этапе прямое направление от устройства-источника данных к координатору длиной L разобьем на секторы круга так, как показано на рисунке 5. Такое деление на секторы отражает свойства антенн сенсорных устройств, ориентированных в направлении координатора. Каждый выделенный базовый сектор характеризуется центральным углом θ_k и радиусом R_k ($k = \overline{1, K}$, K – число базовых секторов) и содержит две опорные точки: нижнюю (начальную, передающую данные) и верхнюю (конечную, принимающую данные). Сектора сопрягаются таким образом, чтобы верхняя опорная точка предыдущего базового сектора одновременно являлась нижней опорной точкой последующего базового сектора. Нижняя точка первого базового сектора ($k = 1$) соответствует устройству-источнику данных, а верхняя точка последнего базового сектора ($k = K$) – координатору. Все промежуточные опорные точки ($k = 2, \dots, K-1$) соответствуют устройствам-ретрансляторам. Зависящие от базовых углов и радиусов геометрические свойства секторов выбираются так, чтобы именно в них умещались все устройства системы, ретрансляция через которые ведет к сокращению энергии, потребляемой при передаче данных. Для этого центральный угол k -го сектора θ_k должен удовлетворять условию $\theta_k < 90^\circ$, так как экономия энергии достигается только при ретрансляции через устройства, образующие тупые углы с прямой, соединяющей узел-источник данных с координатором.

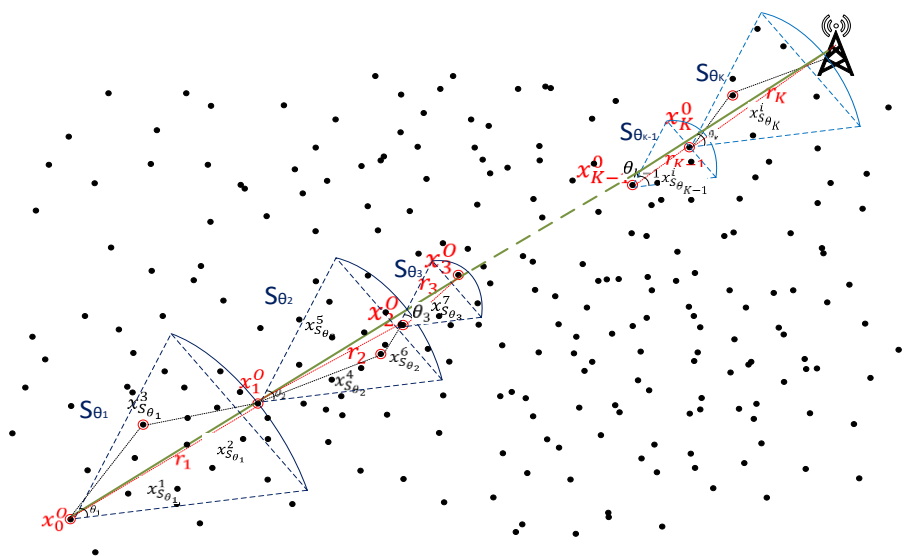


Рис. 5. Алгоритм выбора маршрута

Fig. 5. Route Selection Algorithm

При выборе радиуса k -го сектора R_k ($k = \overline{1, K}$) надо учитывать расстояние L от устройства-источника данных до координатора и количество отрезков K , на которое делится это расстояние. С учетом длины отрезка, которая равна $\frac{L}{K}$, для поиска R_k сначала отметим такую точку на прямой, которая удалена от узла-источника на расстояние $k \cdot \frac{L}{K}$. Затем назовем верхней опорной точкой такое устройство сенсорной сети, которое расположено в k -м секторе и наиболее близко к отмеченной точке на прямой. Радиус R_k равен расстоянию от нижней опорной точки сектора до назначенной верхней опорной точки. В результате реализации такой процедуры радиусы R_k секторов могут отличаться друг от друга.

На этапах 2, ..., i , ..., N работы алгоритма некоторые из секторов, выделенных на предыдущем этапе, рассматриваются в качестве подсистемы сенсорной сети, внутри которой между опорными точками при помощи аналогичной процедуры осуществляется поиск устройств для дополнительных ретрансляций.

Алгоритм может быть реализован и в один этап. Разделение его на несколько этапов становится целесообразным при неравномерном распределении сенсорных устройств в зоне охвата, при котором количество устройств в некоторых выделяемых секторах существенно отличается от количества устройств в других секторах.

Сформулированные выше утверждения 1 и 2 указывают на то, что рациональный по критерию энергозатрат маршрут передачи данных состоит только из устройств сенсорной сети, расположенных в выделяемых секторах, а привлечение для ретрансляций других устройств ведет к увеличению совокупного энергопотребления. Если в некоторых секторах отсутствуют устройства, доступные для ретрансляции, то следует увеличить площадь сектора за счет увеличения центрального угла и (или) радиуса, но в плотных сенсорных сетях вероятность такого события достаточно мала даже для небольших по площади секторов, в соответствии с формулой (3). Энергоэффективность выбранного маршрута будет повышаться при уменьшении центральных углов секторов θ_k и увеличении числа секторов (уменьшении радиусов R_k) и, следовательно, числа ретрансляций. Возможность уменьшения R_k ограничена допустимым временем доставки данных.

Эффект от использования предложенного подхода проявляется в уменьшении вычислительной сложности поиска маршрута, так как вместо вычислений в полном графе с большим количеством вершин строятся маршруты внутри секторов, охватывающих значительно меньшее число устройств системы.

Численный пример

Для проведения расчетов использовалась модель беспроводной сенсорной сети, включающая в себя координатор с антенной LoRa 868 МГц. Количество секторов $K = 3$, $v = 3 \cdot 10^8$ м/с, $C_{\text{пер}} = 1$, $C_{\text{пр}} = 1$, $P_{\text{пр}} = 2,2 \cdot 10^{-11}$ Вт, $f = 868 \cdot 10^6$ Гц. Допустимое число ретрансляций – 6. Предполагалось, что устройства распределены в пространстве по закону Пуассона. Использовался двухэтапный алгоритм, причем на втором этапе в каждом из секторов допускалась одна или несколько ретрансляций.

В основе методики расчета лежит следующая процедура. Случайным образом генерировались плотные сенсорные поля, образующие сенсорные сети с пуассоновским распределением точек на плоскости, для которых применялся эмпирический алгоритм поиска маршрута передачи данных. Для найденных маршрутов фиксировались показатели энергопотребления.

Зависимости потребляемой мощности при разных значениях плотности сенсорной сети и допустимого числа ретрансляций представлены на рисунке 6.

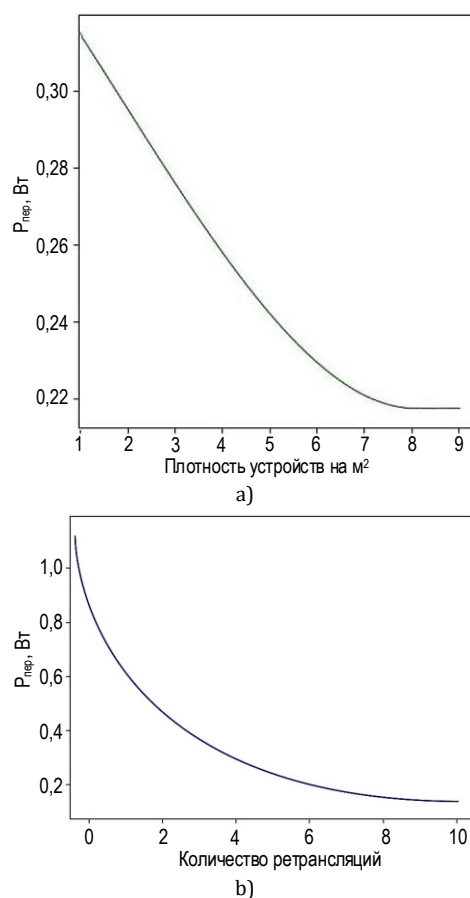


Рис. 6. Зависимость затрачиваемой мощности на передачу пакета данных к базовой станции от плотности распределения устройств (а) и количества ретрансляций (б)

Fig. 6. Dependence of Power Consumption for Data Packet Transmission to the Base Station on Device Distribution Density (a) and Number of Retransmissions (b)

Полученные графики соответствуют следствию из теоремы косинусов [10], а также показывают эффективность предлагаемого алгоритма. При увеличении плотности устройств на квадратный метр мощность, затрачиваемая на передачу сигнала, уменьшается, это свидетельствует о том, что эффективность предлагаемого алгоритма повышается в сверхплотных системах. Такая же тенденция наблюдается и при увеличении количества ретрансляций. Однако увеличение числа ретрансляций ведет к дополнительным энергетическим затратам вычислительных ресурсов, что связано с необходимостью обработки большого количества информации на каждом этапе ретрансляции.

Было проведено сравнение результатов работы предложенного эмпирического алгоритма с точными алгоритмами поиска рациональных путей, которые не учитывают свойства функций, устанавливающих зависимость мощности радиосигнала от расстояния. Результаты сравнения приведены в таблице 1.

ТАБЛИЦА 1. Результаты сравнения эмпирического и точного алгоритмов выбора маршрута

TABLE 1. Results of Comparison of Route Selection Algorithms

Допустимое число ретрансляций – 6		
Плотность (сенсорных устройств в среднем на м ²)	Среднее значение мощности на передачу данных, Вт	
	Предлагаемый алгоритм	Точный алгоритм
1	~ 0,33	~ 0,30
2	~ 0,29	~ 0,27
3	~ 0,27	~ 0,26

Можно заметить, что точность эмпирического алгоритма повышается при увеличении плотности сети, поскольку в этом случае увеличивается вероятность нахождения опорного узла P в секторе с меньшим углом θ_k в соответствии с выражением (2). Одновременно по мере увеличения плотности уменьшается и минимальная мощность, требуемая для передачи данных, так как повышается вероятность нахождения транзитных устройств, в большей степени удовлетворяющих условиям теоремы косинусов.

Результаты и их обсуждение

Для оптимизации энергозатрат сенсорных сетей при передаче данных известен целый ряд алгоритмов. Среди них – EADVR (аббр. от англ. Energy-Aware Distance Vector Routing) – протокол векторной маршрутизации расстояния с учетом затрачиваемой энергии, основанный на алгоритме Белмана – Форда; EALSR (аббр. от англ. Energy-Aware Link State Routing) [12, 13]. Известны и другие про-

токолы маршрутизации, позволяющие минимизировать энергопотребление: AODV (аббр. от англ. Ad hoc On-Demand Distance Vector), DSR (аббр. от англ. Dynamic Source Routing), OLSR (аббр. от англ. Optimized Link State Routing), LOAD (аббр. от англ. Link and Energy Aware Multipath Routing) и пр.

Алгоритмы маршрутизации различаются способом выбора маршрута и разным принципом управления процессом передачи данных. Они нацелены на оптимизацию либо длины маршрута на основе данных о расстоянии между устройствами, либо информации об остаточной емкости батарей питания в разных устройствах.

В отличие от популярного протокола EALSR, в котором используется алгоритм Дейкстры (поиска кратчайшего пути), предложенный алгоритм при выборе маршрута использует данные о необходимой мощности для передачи данных между конкретной парой устройств и свойствах функции, устанавливающей зависимость мощности от расстояния. В то время как алгоритмы EADVR и EALSR нацелены на минимизацию задержек при ограниченной пропускной способности каналов связи, предлагаемый алгоритм обеспечивает экономию мощности, затрачиваемой при передаче, при ограниченном времени доставки.

Очевидным преимуществом предлагаемого алгоритма является нацеленность на минимизацию мощности по всему выбранному маршруту в конкретном маршруте, что приводит к точным оценкам затрачиваемых энергетических ресурсов системы в целом.

Точность предлагаемого алгоритма повышается по мере увеличения плотности сенсорной сети. Он имеет преимущество и для сенсорных сетей, развернутых на большой площади, так как использует информацию о расположении в пространстве относительно небольшого количества сенсорных устройств, расположенных в относительной близости к прямой, соединяющей источник данных с координатором.

Использование теоремы косинусов при выборе маршрута позволяет исключить из рассмотрения заведомо неэффективные в энергетическом смысле маршруты, так как при выборе пути учитывается не только расстояние, но и углы между взаимодействующими сенсорами по всему маршруту относительно прямого пути.

Выводы

В статье рассмотрена сенсорная сеть, процессы функционирования которой характеризуются не только взаимодействиями сенсорных устройств друг с другом и координатором во времени и пространстве, но и энергетическими характеристиками, влияющие на связность и время жизни сети.

В ходе исследования были сформулированы утверждения, являющиеся следствиями теоремы косинусов и формулы Фрииса и, на их основе предложен эмпирический алгоритм выбора маршрута передачи данных от устройства-источника данных к координатору, снижающий совокупное потребление энергии устройствами, составляющими маршрут. Алгоритм гарантирует директивное время доставки данных потребителям.

Построена модель, позволяющая оценить зависимости потребляемой мощности от допустимого числа ретрансляций и плотности сети и представлены соответствующие расчеты.

Проведено сравнение с точными алгоритмами поиска рациональных путей. Выявлена и объяс-

нена закономерность, согласно которой при увеличении плотности сенсорной сети повышается точность предложенного эмпирического алгоритма, практическое использование которого становится целесообразным уже при плотности системы $\sim 1,5$ устройства в среднем на 1 м^2 .

Доказано, что предложенный эмпирический алгоритм выбора рационального маршрута передачи данных в беспроводной сенсорной сети позволяет определить среди всех альтернативных тот маршрут к координатору, который требует наименьшей мощности, снижая тем самым количество энергии, потребляемой в совокупности всеми сенсорами сети.

Список источников

1. Романова А.А. Модели и методы оценки вероятностно-энергетических характеристик информационного взаимодействия в интернете вещей. Автореферат дис. ... канд. техн. наук. СПб.: Санкт-Петербургский государственный электротехнический университет "ЛЭТИ" им. В.И. Ульянова (Ленина), 2023. 21 с.
2. Almuhaaya M.A.M., Jabbar W.A., Sulaiman N., Abdulmalek S. A survey on LoraWAN Technology: Recent Trends, Opportunities, Simulation Tools and Future Directions // *Electronics*. 2022. Vol. 11. Iss. 1. P. 164. DOI:10.3390/electronics11010164
3. Levchenko P., Bankov D., Khorov E., Lyakhov A. Performance Comparison of NB-Fi, Sigfox, and LoRaWAN // *Sensors*. 2022. Vol. 22. Iss. 24. P. 9633. DOI:10.3390/s22249633
4. Зверев Б., Сартаков А. SNB новая LPWAN-технология «Интернета вещей» с высокой пропускной способностью // *Control Engineering Россия*. 2019. № S. С. 38–41. EDN:JRNENT
5. Астахова Т.Н., Верзун Н.А., Касаткин В.В., Колбанев М.О., Шамин А.А. Исследование моделей связности сенсорных сетей // *Информационно-управляющие системы*. 2019. № 5(102). С. 38–50. DOI:10.31799/1684-8853-2019-5-38-50. EDN:GLMGRC
6. Татарникова Т.М., Бимбетов Ф., Горина Е.В. Алгоритм энергоэффективного взаимодействия узлов беспроводной сенсорной сети // *Научно-технический вестник информационных технологий, механики и оптики*. 2022. Т. 22. №. 2. С. 294–301. DOI:10.17586/2226-1494-2022-22-2-294-301. EDN:FCHFEL
7. Уоллес Р. Максимальная дальность связи по радиоканалу в системе: как этого добиться? // *Новости электроники*. 2015. № 11. С. 3.
8. Колбанёв М.О., Татарникова Т.М. Физические ресурсы информационных процессов и технологий // *Научно-технический вестник информационных технологий, механики и оптики*. 2014. № 6(94). С. 113–123. EDN:TBDGTL
9. Аунг Т.А. Разработка методов и средств защищенного сбора данных окружений Интернета Вещей (ИВ). Дис. ... канд. техн. наук. СПб.: Национальный исследовательский университет ИТМО, 2020. 297 с. EDN:QTBXRI
10. Астахова Т.Н., Кирилова Д.А., Колбанёв М.О., Маслов Н.С., Шамин А.А. Критерий выбора оптимального маршрута передачи сообщения в беспроводных сенсорных сетях // *Телекоммуникации*. 2020. № 7. С. 6–12. EDN:VIEYJZ
11. Астахова Т.Н., Верзун Н.А., Колбанев М.О., Полянская Н.А., Шамин А.А. Вероятностно-энергетические характеристики взаимодействия умных вещей // *Вестник НГИЭИ*. 2019. № 4(95). С. 66–77. EDN:IEAGYT
12. Kocakulak M., Butun I. An overview of Wireless Sensor Networks towards internet of things // *Proceedings of the 7th Annual Computing and Communication Workshop and Conference (CCWC, Las Vegas, USA, 09–11 January 2017)*. IEEE, 2017. DOI:10.1109/CCWC.2017.7868374
13. Mahlke S., Madani S.A., Roetzer M. Energy aware distance vector routing Scheme for Data Centric Low Power Wireless Sensor Networks // *Proceedings of the 4th IEEE International Conference on Industrial Informatics (Singapore, 16–18 August 2006)*. IEEE, 2006. PP. 1030–1035. DOI:10.1109/INDIN.2006.275739
14. Levchenko P., Bankov D., Khorov E., Lyakhov A. Performance Comparison of NB-Fi, Sigfox, and LoRaWAN // *Sensors*. 2022. Vol. 22. Iss. 24. P. 9633. DOI:10.3390/s22249633

References

1. Romanova A.A. *Models and methods for evaluating probabilistic-energetic characteristics of information interaction in the Internet of Things*. Phd Theses. St. Petersburg: St. Petersburg State Electrotechnical University "LETI" Publ.; 2023. 21 p. (in Russ.)
2. Almuhaaya M.A.M., Jabbar W.A., Sulaiman N., Abdulmalek S. A survey on LoraWAN Technology: Recent Trends, Opportunities, Simulation Tools and Future Directions. *Electronics*. 2022;11(1):164. DOI:10.3390/electronics11010164
3. Levchenko P., Bankov D., Khorov E., Lyakhov A. Performance Comparison of NB-Fi, Sigfox, and LoRaWAN. *Sensors*. 2022;22(24):9633. DOI:10.3390/s22249633
4. Zverev B., Sartakov A. SNB new LPWAN technology of "Internet of Things" with high throughput. *Control Engineering Russia*. 2019;S:38–41. (in Russ.) EDN:JRNENT

5. Astakhova T.N., Verzun N.A., Kasatkin V.V., Kolbanov M.O., Shamin A.A. Sensor network connectivity models. *Information and Control Systems*. 2019;5(102):38–50. (in Russ.) DOI:10.31799/1684-8853-2019-5-38-50. EDN:GLMGRC
6. Tatarnikova T.M., Bimbetov F., Gorina E.V. Algorithm for energy-efficient interaction of wireless sensor network nodes. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2022;22(2):294–301. (in Russ.) DOI:10.17586/2226-1494-2022-22-2-294-301. EDN:FCHFEL
7. Uolles R. Maximum radio communication range in the system: how to achieve this? *News of Electronics*. 2015;11:3–13. (in Russ.)
8. Kolbanov M.O., Tatarnikova T.M. Physical resources of information processes and technologies. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2014;6(94):113–123. (in Russ.) EDN:TBDGTL
9. Aung T.A. *Development of methods and means of protected data collection of Internet of Things (IoT) districts*. Phd Theses. St. Petersburg: ITMO University Publ.; 2020. 297 p. (in Russ.)
10. Astakhova T.N., Kirilova D.A., Kolbanov M.O., Maslov N.S., Shamin A.A. Criterion for selecting the optimal route of message transmission in wireless sensor networks. *Telecommunications*. 2020;7:6–12. (in Russ.) EDN:VIEYJZ
11. Astakhova T.N., Verzun N.A., Kasatkin V.V., Kolbanov M.O., Shamin A.A. Sensor network connectivity models. *Information and Control Systems*. 2019;5(102):38–50. (in Russ.) DOI:10.31799/1684-8853-2019-5-38-50. EDN:GLMGRC
12. Kocakulak M., Butun I. An overview of Wireless Sensor Networks towards internet of things. *Proceedings of the 7th Annual Computing and Communication Workshop and Conference, CCWC, 09–11 January 2017, Las Vegas, USA*. IEEE; 2017. DOI:10.1109/CCWC.2017.7868374
13. Mahlknecht S., Madani S.A., Roetzer M. Energy aware distance vector routing Scheme for Data Centric Low Power Wireless Sensor Networks. *Proceedings of the 4th IEEE International Conference on Industrial Informatics, 16–18 August 2006, Singapore*. IEEE; 2006. p.1030–1035. DOI:10.1109/INDIN.2006.275739

Статья поступила в редакцию 12.07.2024; одобрена после рецензирования 05.08.2024; принята к публикации 07.08.2024.

The article was submitted 12.07.2024; approved after reviewing 05.08.2024; accepted for publication 07.08.2024.

Информация об авторах:

АСТАХОВА
Татьяна Николаевна

кандидат физико-математических наук, доцент, доцент кафедры информационных систем и технологии Нижегородского государственного инженерно-экономического университета

 <https://orcid.org/0000-0002-7032-0697>

КОЛБАНЁВ
Михаил Олегович

доктор технических наук, профессор, профессор кафедры информационных систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ»

 <https://orcid.org/0000-0003-4825-6972>

ЛЯМИН
Александр Сергеевич

кандидат экономических наук, доцент, доцент кафедры информационных систем и технологии Нижегородского государственного инженерно-экономического университета

 <https://orcid.org/0000-0003-0027-2404>

МАСЛОВ
Никита Сергеевич

старший преподаватель кафедры информационных систем и технологии Нижегородского государственного инженерно-экономического университета

 <https://orcid.org/0000-0003-0209-7267>

МАСЛОВА
Дарья Александровна

старший преподаватель кафедры информационных систем и технологии Нижегородского государственного инженерно-экономического университета

 <https://orcid.org/0000-0002-3329-4821>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

Научная статья

УДК 004.27+004.056

<https://doi.org/10.31854/1813-324X-2024-10-4-110-125>

Обеспечение информационной безопасности интегрируемых информационных систем на базе доверия

Виталий Владимирович Грызунов¹✉, viv1313r@mail.ru
Александр Сергеевич Крюков², steelrat76@mail.ru
Александр Викторович Шестаков¹, alexander.shestakov@yandex.ru
Игорь Алексеевич Зикратов³, zikratov.ia@sut.ru

¹Санкт-Петербургский университет ГПС МЧС России,

Санкт-Петербург, 196105, Российская Федерация

²Российский государственный университет правосудия,

Москва, 117418, Российская Федерация

³Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича,

Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. Информационные системы интегрируются между собой, что приводит к необходимости обеспечить доверие к интегрированной системе. Обеспечение доверия требует formalизовать понятие доверия, изучить его природу и структуру.

Цель работы: снять противоречие между потребностями, предоставить доступ к ресурсам результирующей интегрированной системы и обеспечить выполнение требований информационной безопасности каждой из интегрируемых систем за счет формулировки понятия доверия с позиции информационной безопасности. Для чего были использованы **методы** системного анализа, теории управления рисками, резолюций, синтеза операторного уравнения iSOFT.

Результаты. Выявлены основные недостатки существующих подходов к формализации понятия «доверие». На базе модели FIST информационной системы разработана функциональная структура доверия и оформлена в нотации IDEF0 для всех уровней интегрируемых информационных систем: обеспечивающего уровня, уровня персонала, уровней аппаратного и программного обеспечений. Приведены примеры нарушения доверия и примеры инструментов создания доверия для каждого уровня информационной системы. Адекватность модели проиллюстрирована на примере реальной интеграции информационных систем. Применение предложенной модели доверия позволило выявить особенности, которые увеличивают риски информационной безопасности для интегрированной информационной системы из примера.

Новизна. Предложена трактовка доверия как меры информационной безопасности в отличие от «риска» как меры опасности, а также – инструмент количественного оценивания доверия. Сформулировано и доказано методом резолюций необходимое и достаточное условие создания максимального доверия в информационной системе.

Практическая значимость. Предлагаемая модель доверия может использоваться при разработке руководящих документов, регламентирующих процесс интеграции информационных систем, при выдвигании требований к обслуживающему персоналу и созданию программ его обучения, для разработки средств защиты информации и методик их применения.

Ключевые слова: риск информационной безопасности, доверие, аргумент доверия, интеграция информационных систем

Ссылка для цитирования: Грызунов В.В., Крюков А.С., Шестаков А.В., Зикратов И.А. Обеспечение информационной безопасности интегрируемых информационных систем на базе доверия // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 110–125. DOI:10.31854/1813-324X-2024-10-4-110-125. EDN:MZMYXF

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-110-125>

Ensuring Information Security of Information Systems to be Integrated Based on Trust

 Vitaliy V. Gryzunov¹✉, viv1313r@mail.ru

 Alexandr S. Krjukov², steelrat76@mail.ru

 Alexander V. Shestakov¹, alexander.shestakov@yandex.ru

 Igor A. Zikratov³, zikratov.ia@sut.ru

¹Saint Petersburg University of State Fire Service of Emercom of Russia,
St. Petersburg, 196105, Russian Federation

²Russian State University of Justice,
Moscow, 117418, Russian Federation

³The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance. Information systems are integrated with each other, which leads to the need to ensure the protection of the integrated system. The level of trust requires formalizing the concept of trust and studying its nature and structure.

The purpose of the article is to remove the contradiction between the needs to provide access to the resources of the resulting integrated system and ensure compliance with the information security requirements of each of the integrated systems by formulating the concept of trust from the information security perspective.

Methods used: systems analysis, risk management theory, resolutions, iSoft operator equation synthesis method.

Results. Main shortcomings of existing approaches to the formalization of the concept of “trust” are identified. Based on the FIST information system model, a functional trust structure has been developed and formalized in IDEF0 notation for all levels of integrated information systems: supporting level, personnel level, hardware and software levels. Examples of violation of trust and examples of tools for creating trust for each level of the information system are given. The adequacy of the model is illustrated by the example of real integration of information systems. Application of the proposed trust model made it possible to identify features that increase information security risks for the integrated information system from the example.

Novelty. An interpretation of trust as a measure of information security is proposed, in contrast to “risk” as a measure of danger. A tool for quantitative assessment of trust is proposed. A necessary and sufficient condition for creating maximum trust in an information system is formulated and proven by the resolution method

Practical significance. The proposed trust model can be used in the development of guidance documents regulating the process of integration of information systems, in setting requirements for service personnel and creating training programs for them, for developing information security tools and methods for their application.

Keywords: information security risk, trust, assurance argument, information systems integration

For citation: Gryzunov V.V., Krjukov A.S., Shestakov A.V., Zikratov I.A. Ensuring Information Security of Information Systems to be Integrated Based on Trust. *Proceedings of Telecommunication Universities*. 2024;10(4):110–125. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-110-125. EDN: MZMYXF

1. Введение

Информационные системы с каждым годом все больше интегрируются друг с другом. Это происходит как в рамках локальной бизнес-сферы: Сбер, Яндекс, электронные торговые площадки и т. д., так и на международном уровне, согласно решению Коллегии евразийской экономической комиссии

(ЕЭК). На практике, под интеграцией информационных систем (ИС) понимают процесс установления связей между информационными системами для получения единого информационного пространства, организации поддержки сквозных бизнес-процессов. Под интегрированной системой (от англ. Integrated System) понимается «система, в которой все входящие в нее подсистемы работают по

единому алгоритму, т.е. имеет единую точку управления» (ГОСТ Р 55062–2021).

В России действует «Стратегия развития информационного общества на 2017–2030 годы», утвержденная указом Президента Российской Федерации от 9 мая 2017 г. № 203. Целью данной стратегии является улучшение взаимодействия бизнеса и общества с государственными органами власти за счет решения «задачи интеграции федеральных государственных информационных систем (ФГИС) и инфраструктуры пространственных данных на уровне среды (создающей условия для развития платформ и технологий их эффективного взаимодействия с субъектами органов власти) различных сфер деятельности» [1].

Для кардинального улучшения взаимодействия между государством и обществом необходимо устойчивое функционирование информационной инфраструктуры Российской Федерации. Информационная инфраструктура страны в данный момент насчитывает более 50 государственных информационных систем, которые взаимодействуют между собой тем или иным способом.

Интеграция ИС на международном, государственном и корпоративном уровнях позволяет в полной мере раскрыть преимущества, которые дает цифровизация: снижение издержек, связанных с информационной безопасностью (ИБ), как указано в G20 Ministerial in Tsukuba, Japan: Statement on Trade and Digital (<https://www.economy.gov.ru/material/file/d53673906a29de7bde82260e21ebcd8b/G20.pdf>), повышение эффективности эксплуатации ИС, согласно докладу корпорации PricewaterhouseCoopers «Building Digital Trust The confidence to take risks» (<https://www.pwc.com/cy/en/technology-consulting-services/assets/building-digital-trust-january-2016.pdf>), создание уверенности в предоставляемых цифровых услугах [2–3], упрощение адаптации ИС к действию деструктивных факторов [4].

В ходе интеграции образуются новые связи между ИС, поэтому возникает проблема согласования целей ИБ, моделей угроз, требований к обеспечению безопасности интегрируемых систем. На интуитивном уровне согласование вопросов ИБ обычно описывается понятием «доверие». То есть ИС могут интегрироваться, если они доверяют друг другу.

Для решения проблемы интеграции информационных систем принят ГОСТ Р 55062-2021 «Информационные технологии. Интероперабельность. Основные положения». Данный документ описывает единый подход к созданию интероперабельных ИС, создание которых позволит полноценно интегрировать информационные системы, но документ не

затрагивает вопросы обеспечения ИБ интегрируемых систем. Действующие нормативные документы по ИБ информационных систем определяют требования к ее обеспечению посредством аттестации отдельного экземпляра ИС, а также в ходе эксплуатации этого экземпляра [5, 6]. Обеспечение безопасности интегрированных ИС при этом не рассматривается.

Таким образом, в настоящий момент наблюдается противоречие между необходимостью интегрировать информационные системы и обеспечить их ИБ. Цель настоящего исследования – снять обозначенное противоречие за счет разработки модели обеспечения ИБ интегрируемых ИС, в которой согласуются цели ИБ, модели угроз, требования к обеспечению ИБ интегрируемых систем. Модель является продолжением исследования [7] и базируется на понятии «доверие».

Понятие доверия относится к области ИБ, и понимается научными исследователями по-разному.

2. Обзор современных исследований

Так, например, в работе [8] авторы затрагивают только аспект управления доступом в сложных, территориально распределенных ИС и определяют доверие как меру готовности стороны *A* с некоторой относительной уверенностью предоставить стороне *B* запрашиваемый доступ, несмотря на возможные негативные последствия, т.е., принимая в расчет возможный ущерб от действий стороны *B*, не согласующихся с заявленной ролью. Поскольку под интегрированной системой понимается система, в которой все входящие в нее подсистемы работают по единому алгоритму (см. выше), можно утверждать, что интеграция сегментов сложной, территориально распределенной ИС и отдельных ИС носит идентичный характер.

В качестве развития предыдущей идеи можно рассматривать работу [9], где авторы определяют доверие как меру готовности стороны *A* положиться на кого-то или что-то в данной ситуации с некоторой относительной уверенностью, несмотря на возможные негативные последствия. В работе отмечается, что в это же определение входит учет потерь в случае обмана. Доверие выражается через риск, преддоверие (репутацию) и канал передачи информации:

$$D = \frac{PD \cdot X}{R + \mu},$$

где *D* – доверие; *PD* – преддоверие; *X* – канал передачи данных; μ – роль (ответственность) абонента; *R* – риск.

Достоинство данной модели состоит в том, что она предполагает количественную оценку доверия. Однако вызывают вопросы единицы измерения

как элементов доверия, так и самого доверия. В работе [9] не приведены измерения единицы, не указано, чему соответствует доверие, если R и μ равны нулю.

В исследовании [10] доверие к информационной системе управления (ИСУ) определяется как готовность зависеть от нее и быть уязвимым перед ней, не имея возможности отслеживать или контролировать ее функционирование, т. е. в условиях неопределенности и риска. Авторы рассматривают доверие к ИСУ как пережитое состояние отдельного пользователя, которое включает в себя как когнитивные, так и аффективные аспекты. Доверие возникает и изменяется как функция воспринимаемой надежности ИСУ и склонности человека доверять технологиям в целом. Авторы определяют воспринимаемую надежность ИСУ в качестве ее когнитивной оценки пользователем, имеющей положительные свойства в ситуациях, когда пользователи сталкиваются с потенциальными негативными последствиями. Воспринимаемая надежность включает в себя восприятие пользователями надежности, функциональности, полезности и достоверности ИСУ. Кроме того, авторы предполагают, что испытанное доверие влияет на поведенческие намерения пользователей, которые, в свою очередь, предсказывают фактическое использование ИСУ в рабочих процессах и решениях. В работе рассматривается доверительное использование ИСУ не как «наивное», а, скорее как рефлексивное и преднамеренное использование, не ощущая необходимости в дополнительных обходных путях.

В данной работе доверие устанавливается между пользователем (одним человеком) и информационной системой. Представленная в работе «теоретическая модель доверия» в информационных системах управления не отображает составные части ИСУ, и предполагает доверие сразу ко всей системе. Способы получения количественной оценки доверия опущены.

В работе [11] авторы рассматривают фактор доверия как фактор продолжения работы пользователей в социальной сети Facebook, что только косвенно относится к ИБ. Аналогично предыдущей работе доверие рассматривается между пользователем и информационной системой. Пользователь имеет самую низкую компетенцию в ИБ.

Научные изыскания авторов в [12] касаются доверительных отношений двух технологических ядер фирмы: информационных технологий (ИТ) и исследований и разработок. В статье показано, что отношение доверия между технологическими ядрами не зарегулированы документально и больше строятся на доверии между работниками фирмы. Сам термин «доверие» не формализован и не может быть измерен количественно.

В работе [13] делается вывод, что в широком смысле «доверие к технологии» означает готовность зависеть от конкретной технологии в конкретной ситуации, в которой возможны негативные последствия. Подобно доверию к людям, доверие к конкретной технологии формируется двумя отдельными компонентами – доверительным намерением и доверительными убеждениями. Доверие к технологии, ориентированной на конкретные намерения, означает готовность людей зависеть от конкретной технологии. Когда доверие высоко, люди выражают готовность зависеть от конкретной технологии в неопределенных и рискованных ситуациях. Данный подход описывает доверие качественно, и ограничен к применению на практике, где требуется количественное измерение.

В источнике [14] доверие к ИБ рассматривается как услуга. Основой доверия выступает «готовность стороны (доверителя) быть уязвимой для действий другой стороны (доверительного управляющего) на основе ожидания того, что другой (доверительный управляющий) выполнит определенное действие, важное для доверителя, независимо от возможности мониторить или контролировать эту другую сторону (доверенного управляющего)». Приведены семь предположений по взаимосвязи и на их основе предложена концептуальная модель доверия, в которой сервис-провайдер услуг безопасности интегрирует свои сервисы в корпоративную инфраструктуру заказчика. Однако данная модель не в полной мере позволяет понять, как формируется доверие, и как оно может быть измерено.

В работе [15] авторы приводят «конструкт доверия» к интеллектуальным роботам в сфере туризма. В этой работе предлагается модель многогранного доверия к сервисным роботам, состоящая из трех конструктов – производительность, процесс и цель. Само доверие понимается как «установка, что агент поможет достичь целей человека в ситуации, характеризующейся неопределенностью и уязвимостью». Т. е. можно сказать, что в работе рассматривается только один аспект ИБ – доступность.

Исследование [16–17] посвящено доверию пользователей к мобильному банкингу. Выдвинуто предположение, что доверие, влекущее за собой использование мобильного банкинга, включает в себя: осведомленность, инновационность, конфиденциальность, безопасность, полезность, простоту в использовании. В такой постановке ИБ характеризуется понятиями «конфиденциальность» и «безопасность». Но не приводятся инструменты, позволяющие измерить данные понятия количественно.

В исследовании [18] представлено определение отношения доверия между информационными системами. Отношением доверия между системами A

и B – подмножество $T_{a,b} \subset S(A) \times S(B)$. Если пара (a, b) принадлежит $T_{a,b}$, то субъект a может получить доступ к объектам системы B посредством субъекта b . В этом случае автор считает, что субъект b доверяет субъекту a . Данное определение доверия касается только получения доступа. В приведенной формулировке доверие – бинарная величина, поэтому невозможно выделить уровни доверия для более гибкого применения данного термина. Некоторые исследователи не формализуют понятие доверия, но подразумевают его существование, и описывают только состав доверия и/или порядок применения. Так, например, авторы [2] приводят главные факторы установления доверия: уверенность, риск, честность, законность, безопасность.

В работе [19] утверждается, что доверие должно быть реализовано с помощью корпоративного WEB-портала, обеспечивающего единую точку входа в информационное пространство корпоративной сети со всеми необходимыми атрибутами – межсетевыми экранами, VPN, системами обнаружения вторжений, антивирусной защитой, системой фильтрации спама и т. д., Доверием может управлять корпоративный удостоверяющий центр с разными режимами доступа, желательно с разной криптографией – западной и отечественной, с разной длиной ключа и т. д.

В своем докладе «Building digital trust into better experiences» компания IBM указывает, что цифровое доверие состоит из нескольких элементов: пользователь и его уникальные атрибуты; аутентификация устройства и конечной точки; деятельность, связанная с данными, приложением и пользователем; среда пользователя, а также сеть; поведение пользователя, при использовании информации; надежность; прозрачность; честность; безопасность. В докладе IBM высказывается мнение, что целью построения доверия должно быть создание безопасного функционала для пользователей. Вопросы количественного измерения доверия отсутствуют.

PricewaterhouseCoopers в своем докладе «Building Digital Trust The confidence to take risks» указывает, что для построения доверия нужна уверенность в:

- безопасности (системы безопасны для защиты данных и нет проблем с идентификацией пользователей);
- данных (в целостности данных и способность извлекать выгоду для развития и получения прибыли);
- системах (ИС осуществляется контроль и мониторинг, позволяющие гарантировать, что они выполняют то, что от них требуется, независимо от того, являются ли они собственными системами или предоставляются в виде облачных сервисов);

– рисках (технологические риски понятны и хорошо управляются, а цифровые платформы будут доступны при необходимости круглосуточно);

– программе цифровой трансформации (в информационной системе можно начать и реализовать следующую комплексную программу цифровой трансформации так, чтобы обеспечить ожидаемые выгоды, в срок и в рамках бюджета).

В данном докладе не уточняется, что такое «уверенность», как ее измерить и как быть, если доверие целенаправленно атакуется [20].

В решении Совета ЕЭК от 05.12.2018 г. № 96 «О требованиях к созданию, развитию и функционированию трансграничного пространства доверия» идет речь о пространстве доверия при межгосударственном обмене данными, выдвигаются требования к созданию, развитию и функционированию трансграничного пространства доверия, требования к архитектуре, обеспечивающей пространство доверия. Однако само определение «пространства доверия» отсутствует.

Согласно ГОСТ Р 54581-2011 доверие (*от англ. Assurance*): выполнение соответствующих действий или процедур для обеспечения уверенности в том, что оцениваемый объект соответствует своим целям безопасности. Здесь стоит отметить следующие моменты:

1) налицо неудачный перевод «assurance», более точным является слово «гарантия», «доверию» соответствует английское слово «trust»;

2) доверие является процессом, что немного не согласуется с дальнейшим текстом документа;

3) данный ГОСТ базируется на международном стандарте, который в 2012 г. утратил силу и заменен на ISO/IEC TR 15443-1:2012, где assurance – это основания для обоснованной уверенности в том, что объект оценки соответствует функциональным требованиям безопасности.

ГОСТ Р 54581-2011 касается безошибочной работы ИТ-систем, а не вопросов обеспечения ИБ.

Аналогичная ситуация с определением доверия в ГОСТ Р ИСО/МЭК 15408-1-2012, где под доверием (assurance) понимается «основание для уверенности в том, что объект оценки отвечает конкретным функциональным требованиям безопасности». Суть стандарта в том, чтобы доказать, что для конкретного типа объекта оценки, в роли которого в настоящей статье выступают интегрируемые информационные системы, применяемые средства защиты позволяют гарантировать достижение одного из оценочных уровней доверия [21–23].

На каждый тип объекта оценки составляется профиль защиты – «независимое от реализации изложение потребностей в безопасности для некоторого типа объекта оценки». Применение ГОСТ 15408 для интегрируемых ИС ограничено тем, что:

1) в ходе их интеграции гарантированно возникают особенности, которые сложно типизировать и предусмотреть заранее; 2) интегрированная ИС изменяется во времени, что требует периодического пересмотра профиля защиты и самой системы, а также нового запуска всех оценочных процедур.

Обозначенная ситуация, когда информационная система изменяется во времени, и применение нормативного подхода для обеспечения ее безопасности не вполне достаточное, привело к появлению новой концепции нулевого уровня доверия (Zero Trust Architecture) в NIST Special Publication 800-207 (<https://doi.org/10.6028/NIST.SP.800-207>). В стандарте речь идет именно о доверии (trust), а не о гарантии (assurance). Нулевое доверие (ZT) «представляет собой набор концепций и идей, предназначенных для минимизации неопределенности при обеспечении точных решений о доступе с наименьшими привилегиями для каждого запроса в информационных системах и службах в условиях сети, которая считается скомпрометированной», т. е. фактически говорится о том, как действовать в ИС, если никакому элементу доверять нельзя [24–26]. Сам термин «доверие» оставлен за скобками стандарта.

Таким образом можно утверждать, что термин «доверие» понимается интуитивно, нуждается в формализации и введении соответствующей метрики. Предложенные в статье определение доверия и количественный показатель для его измерения базируются на модели FIST [27]. FIST разработана в ходе исследования геоинформационных систем. Поскольку геоинформационные системы являются самым сложным типом информационных систем, то любые другие ИС могут рассматриваться как упрощенные версии геоинформационных систем. Следовательно, модель FIST, описывающая произвольные геоинформационные системы, применима для описания произвольной ИС.

3. Доверие в информационной системе согласно модели FIST

Обобщая опыт исследователей, продиктованный в первую очередь требованиями практического применения, можно сделать вывод, что доверие является некоторым антиподом риска. Если считать риск мерой опасности, то *доверие является мерой безопасности*. Когда доверие максимально, то риск минимален. Когда риск максимален, то о доверии говорить не приходится.

Риск – влияние неопределенности на достижение поставленных целей (ГОСТ Р 51897-2021. Менеджмент риска. Термины и определения). Риск ИБ (*по англ.* Information Security Risk) – «возможность того, что данная угроза сможет воспользоваться уязвимостью актива или группы активов и тем самым нане-

сет ущерб организации. Измеряется, исходя из комбинации вероятности события и его последствия» (ГОСТ Р ИСО/МЭК 27005-2010. Информационная технология. Методы и средства обеспечения безопасности. Менеджмент риска информационной безопасности). Далее под риском понимается риск ИБ.

Специалисты ИБ выполняют обеспечивающие, а не целевые функции, поэтому обычно не вмешиваются в соответствующие процессы. Поскольку ущерб рассчитывается именно по результатам целевых процессов, то специалисты ИБ не могут его уменьшить, но могут влиять на вероятность возникновения ущерба (далее – вероятность риска P_{risk}) за счет применения средств защиты информации, которые нейтрализуют угрозы.

Взаимосвязь вероятности риска и доверия аналогична связи вероятности отказа и вероятности безотказной работы: если доверия нет, то вероятность риска равна единице, если доверие максимально и равно единице, то вероятность риска равна нулю:

$$D = 1 - R, \quad (1)$$

где D – доверие $\in [0; 1]$; R – вероятность риска $\in [0; 1]$.

Следует отметить, что на практике доверие не может быть равно 1, потому что всегда существует остаточный риск.

Практическое использование формулы (1) будет представлено в дальнейших исследованиях, связанных непосредственно с расчетом количественного значения доверия в рамках модели доверия интегрируемых ИС.

Согласно модели FIST, все ИС имеют «вложенные друг в друга уровни»: уровень программного обеспечения (УПО) «вложен» в уровень аппаратного обеспечения (УАО), УАО «вложен» в уровень персонала (УП), УП «вложен» в обеспечивающий уровень (рисунок 1). Другими словами – обеспечивающим является метауровень для УП, который является метауровнем для УАО, представляющий, в свою очередь, метауровнем для УПО. Метауровень задает требуемые пространственно-временные состояния «вложенного» уровня, т. е. предъявляет требования к его функционированию, накладывает ограничения [27]. На каждом уровне иерархии к доверию предъявляются свои требования, существуют свои особенности его создания; доверие имеет свой состав.

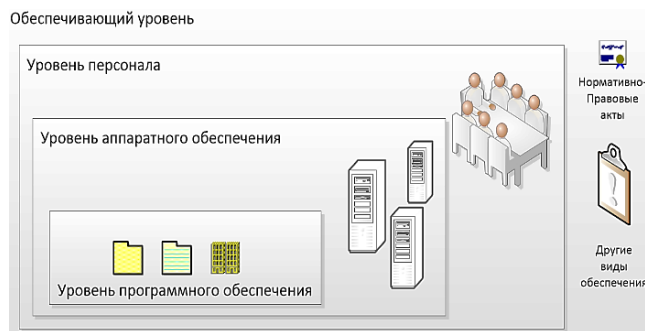


Рис. 1. Модель FIST

Fig. 1. Model FIST

На уровне программного обеспечения доверие связано с программными средствами ИС. Здесь оказывают деструктивное влияние незашифрованный трафик, вредоносное программное обеспечение, эксплойты и т. д. Пример инструмента создания доверия на данном уровне – внедрение протокола https.

На уровне аппаратного обеспечения доверие касается аппаратных средств ИС. Его нарушают аппаратные закладки, подслушивающие устройства, средства радиоэлектронной борьбы и т. д. На данном уровне доверия формируется при помощи модулей доверенной загрузки.

На уровне персонала доверие создается посредством работы с людьми. Действия, совершаемые социальными инженерами, халатными сотрудниками, болтунами и т. д. приводят к снижению доверия к организации [28]. Пример инструмента создания доверия на данном уровне – обучение персонала, проверка его лояльности компании.

На обеспечивающем уровне доверие связано с организационными, правовыми, финансовыми и другими видами обеспечения интеграции ИС [29]. Доверие нарушается за счет сокращения финансирования, ослабления или чрезмерного ужесточения требований к персоналу, программно-аппаратным средствам, необязательностью выполнения регламентов работы ИС и т. д. Пример инструментов создания доверия – создание регламентов, оформление требований к ИБ, разделение сфер ответственности персонала, обслуживающего интегрируемые ИС, и т. п.

Управление доверием непосредственно связано с бизнес-процессами и может быть построено аналогично системе управления адаптацией ИС к деструктивным воздействиям [18–20]. Деструктивные воздействия могут быть детерминированными, стохастическими или агрессивными, целенаправленными [27–30]. Одним наиболее удобных инструментов моделирования в этом случае выступает нотация IDEF0, так как он специально разработан для функциональной декомпозиции и моделирования управленческих процедур [31].

4. Функциональная структура доверия

Функциональная структура модели доверия интегрируемых ИС в нотации IDEF0 представлена на рисунках 2 и 3. Точка зрения – специалист ИБ. Элементы структуры доверия выделены различными цветами для облегчения визуального восприятия. Представленная модель доверия отличается своей функциональной структурой, охватывающей все уровни информационной системы согласно модели FIST.

Цель моделирования – выявление и формулировка условия существования доверия, показателей эффективности создания доверия для интегрируемых информационных систем. Для достижения целей моделирования достаточен первый уровень детализации модели доверия в нотации IDEF0. Более глубокая детализация будет приведена в методике построения пространства доверия.

Цель создания доверия – снизить риск интегрируемых информационных систем до величины остаточного риска. Для этого необходимо знать, какие риски есть, какими активами располагают интегрируемые информационные системы (блок А0). Под активами ИС понимаются информационные ресурсы и средства, а также системы информатизации [32–33], имеющие существенное значение для интегрируемых ИС. Наличие доверия подтверждается аргументом доверия ИС.

Согласно ГОСТ Р ИСО/МЭК 21827-2010:

- аргумент доверия (*от англ. Assurance Argument*): совокупность структурированных заявлений о доверии, подтвержденных доказательствами и аргументацией, четко демонстрирующих, каким образом были удовлетворены потребности в доверии;
- заявление о доверии (*от англ. Assurance Claim*): утверждение или поддержка утверждения того, что система удовлетворяет требованиям безопасности;
- свидетельство обеспечения доверия (*от англ. Assurance Evidence*): результаты анализа обеспечения доверия к объекту (включая итоговые отчеты или другие обоснования), поддерживающие утверждение о доверии.

Формирование доверия обусловлено экономическими, физическими и правовыми ограничениями, а также требованиями бизнеса (целевых процессов). Доверие охватывает все уровни иерархии, согласно модели FIST, и создается организационными мероприятиями, персоналом, аппаратными и программными средствами.

Снижение рисков включает в себя четыре блока функций (см. рисунок 3): снижение рисков обеспечивающего уровня (блок А1), снижение рисков на уровне персонала (блок А2), снижение рисков на аппаратном уровне (блок А3), снижение рисков на программном уровне (блок А4).

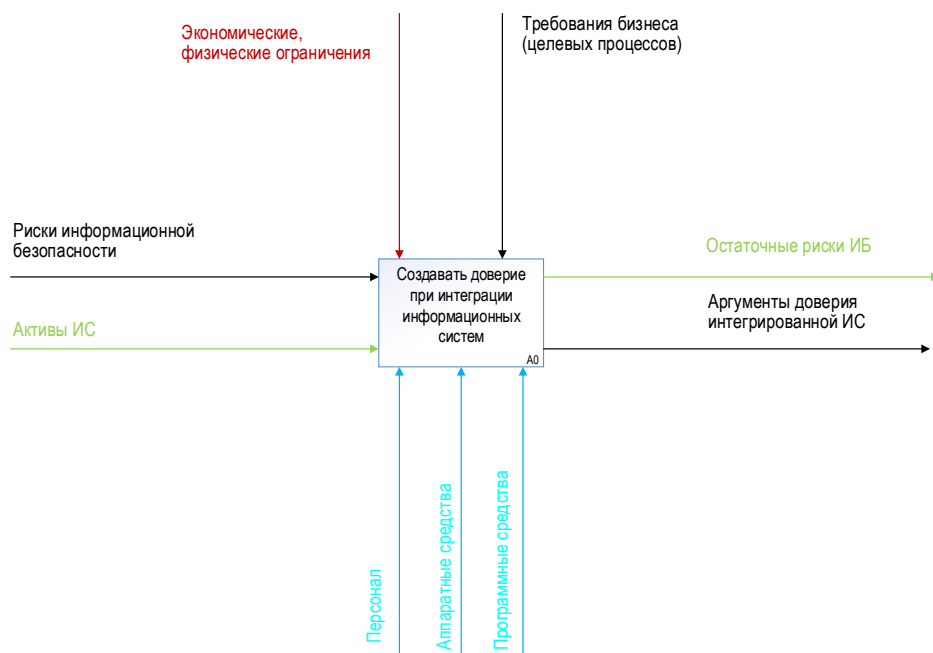


Рис. 2. Доверие в интегрируемых ИС

Fig. 2. Trust in Integrated IS

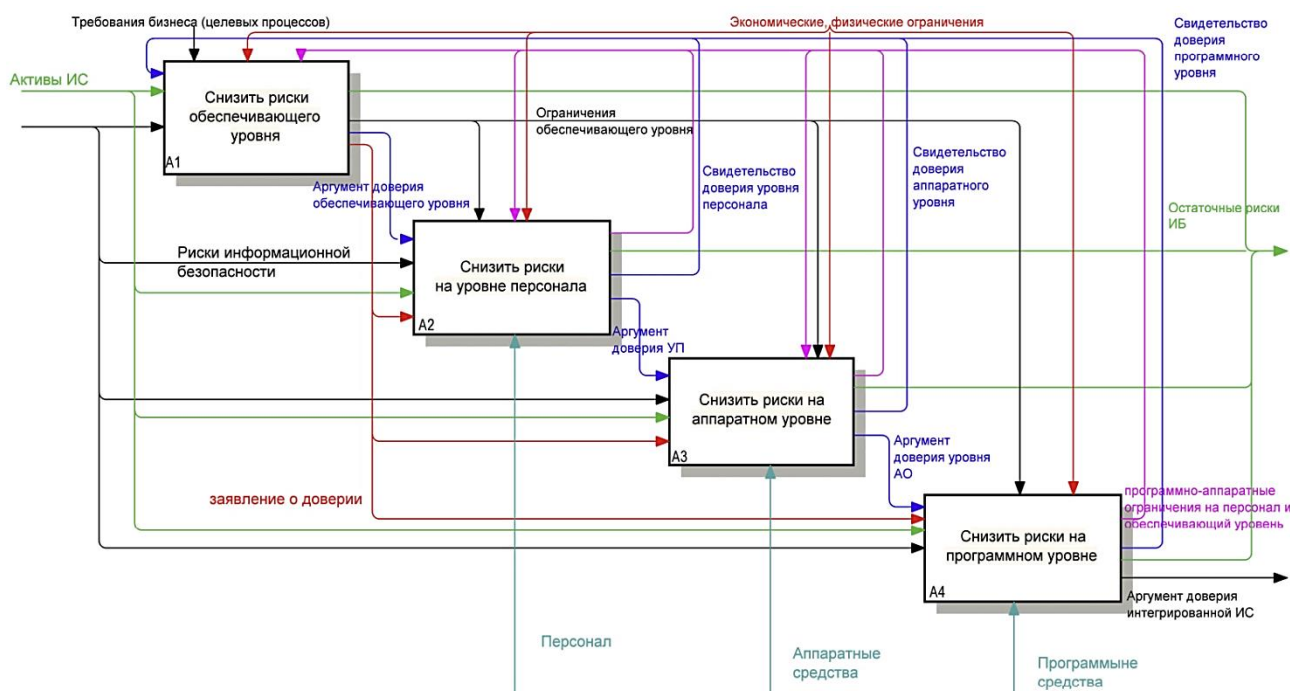


Рис. 3. Функциональная структура доверия в интегрируемых ИС согласно модели FIST

Fig. 3. The Functional Structure of Trust in Integrated IS According to the FIST Model

Снижение риска при создании доверия на обеспечивающем уровне (блок A1) подразумевает подготовку соответствующих организационно-правовых документов, выделение финансирования, предъявление требований ИБ к «вложенным» уровням персонала, аппаратного и программного обеспечения. Разработка документов и финансирование производится для существующих активов интегрируемых информационных систем и рисков

этим активам с учетом действующих требований бизнеса (целевых процессов) и экономических и физических ограничений. При этом принимается во внимание обратная связь от «вложенных» уровней в виде ограничений, накладываемых этими уровнями. Например, в качестве требования обеспечивающего уровня может выступать применение сертифицированного программного, аппаратного или программно-аппаратного обеспечения.

Однако, если такого обеспечения не существует, то это является ограничением «вложенных» уровней, накладываемым на обеспечивающий уровень.

Результатом работы обеспечивающего уровня выступают заявление о доверии и аргумент доверия, подтверждающий, что на обеспечивающем уровне вопросы ИБ решены. Аргумент доверия может содержать доказательство того, что правовая экспертиза документов пройдена, документы составлены и утверждены должным образом, запланированное финансирование поступит в требуемом объеме.

Заявление о доверии относится к разряду формальных требований, являющихся видом обеспечения, поэтому заявления о доверии для всех уровней формируются на обеспечивающем уровне.

Снижение риска на уровне персонала (блок А2) выполняется для существующих активов интегрируемых информационных систем, рисков для этих активов. Снижение выполняется на основе заявлений о доверии и аргумента доверия, сформированных на обеспечивающем уровне, с учетом его ограничений, а также ограничений уровней аппаратного и программного обеспечения. Например, для организации выдвигается требование к подтвержденной квалификации персонала (ограничение обеспечивающего уровня). При этом в интегрируемых информационных системах используется платформа Wintel, что выдвигает дополнительные требования к квалификации персонала со стороны «вложенных» уровней аппаратного и программного обеспечения. В свою очередь уровень персонала выдвигает ограничения, обусловленные персоналом. Например, способность усваивать информацию в процессе обучения.

Аргумент доверия, созданный на уровне персонала, содержит элементы аргумента доверия обеспечивающего уровня, дополненные доказательствами того, что на уровне персонала доверие может быть сформировано: персонал прошел аттестационные испытания, участвовал в киберучениях и т. д.

Снижение риска на уровне аппаратного обеспечения (блок А3) выполняется на основе аргумента доверия, сформулированного на уровне персонала, включающего в себя элементы аргумента доверия обеспечивающего уровня, заявления о доверии защищаемых активов информационных систем и рисков для защищаемых активов.

Снижение риска на уровне аппаратного обеспечения выполняется с учетом ограничений, накладываемых другими уровнями информационных систем, требованиями бизнеса (целевых процессов) и экономических и физических ограничений. Например, если необходимо сформировать доверие к рас-

пределенной мобильной информационной системе, то требования к весу и энергопотреблению являются ограничениями бизнеса (целевых процессов), минимально возможные габариты и стоимость – физическими и экономическими ограничениями, требования к эргономичности – ограничениями уровня персонала, к безопасности эксплуатации – ограничениями обеспечивающего уровня, применение специального программного обеспечения – ограничениями уровня программного обеспечения.

Ограничения, рождаемые на уровне аппаратного обеспечения, касаются возможностей аппаратуры: пропускная способность линии связи, производительность, емкость накопителей и т. д.

Генерируемый аргумент доверия содержит элементы метауровней, плюс подтверждение того, что доверие сформировано на уровне аппаратного обеспечения. Например, аппаратное обеспечение прошло проверки на отсутствие недеklarированных возможностей.

Доверие, создаваемое на уровне программного обеспечения (блок А4), охватывает активы информационной системы, риски для этих активов, использует заявление о доверии и аргумент доверия, поступающий с уровня аппаратного обеспечения.

В процессе формирования доверия учитываются ограничения внешних систем и остальных уровней информационной системы. Например, требование к лицензионной чистоте или наличие антивируса – это ограничения обеспечивающего уровня, способности пользователей запомнить сложные пароли – ограничения уровня персонала, использование процессоров RISC – ограничения уровня аппаратного обеспечения, требование к стоимости разработки программного обеспечения – требование внешней системы.

Уровень программного обеспечения аналогично другим уровням интегрированной информационной системы формирует свои ограничения. Пусть на обеспечивающем уровне выдвинуто требование использовать антивирус марки X. Однако этот антивирус работает только под управлением операционной системы Y. Операционная система Y разработана только для архитектуры CISC. В этом случае требования к архитектуре CISC – ограничение, накладываемое уровнем программного обеспечения на уровень аппаратного обеспечения, требования к квалификации персонала работать с архитектурой CISC и операционной системой Y – ограничение для уровня персонала, необходимость сертифицировать операционную систему Y – ограничение для обеспечивающего уровня.

По итогам создания доверия на уровне программного обеспечения формируется аргумент до-

верия для интегрированной системы, включающей элементы доверия метауровней и элементы уровня программного обеспечения. Например, подтверждение того, что применение сканера уязвимостей не выявило уязвимостей.

Аргумент доверия содержит структурированные заявления о доверии, которые, в свою очередь, утверждают, что система соответствует требованиям безопасности, и учитывают и прямые, и косвенные угрозы (ГОСТ Р ИСО/МЭК 21827-2010). Поэтому можно сказать, что заявление о доверии учитывает риск, или, другими словами, риск входит в заявление о доверии. Риски информационной системы, которые не могут быть нейтрализованы, формируют остаточный риск интегрированной информационной системы.

Вычисление конкретного значения доверия требует разработки соответствующей методики. На основе предложенных определений и модели доверия необходимое и достаточное условие создания максимального доверия в ИС формулируется следующим образом (результатом интеграции информационных систем является также информационная система, поэтому утверждение сформулировано к ИС как к таковой):

Утверждение. Чтобы в информационной системе доверие было максимальным, необходимо и достаточно, чтобы каждый риск ИБ входил в заявление о доверии и аргумент доверия интегрированной ИС включал каждое заявление о доверии.

Ограничение. Каждый риск ИБ можно нейтрализовать посредством защиты до величины остаточного риска.

Доказываемое утверждение в виде резолюций записывается в следующем виде:

$$\{J\} \Leftrightarrow \{R \wedge S\},$$

где в фигурные скобки $\{ \}$ берется система формул или множество дизъюнктов; $\{J\} \rightarrow \{R \wedge S\}$ – необходимость; $\{R \wedge S\} \rightarrow \{J\}$ – достаточность.

Доказательство непротиворечивости выражения $\{J\} \Rightarrow \{R \wedge S\}$.

Докажем необходимость методом резолюций. Согласно методу, чтобы доказать истинность утверждения $F1 \Rightarrow F2$, необходимо и достаточно убедиться, что формула $F1 \rightarrow F2$ является тавтологией, где $F1, F2$ – это посылки. Для этого необходимо и достаточно убедиться, что система формул $\{F1, \overline{F2}\}$ содержит в себе тождественно ложный дизъюнкт $0 \vee 0$.

Система формул утверждения в случае доказательства необходимости примет вид:

$$\{J, \overline{(R \wedge S)}\} = \{J, \overline{R} \vee \overline{S}\}.$$

Добавим в систему формул неявно существующие утверждения (таблица 1).

ТАБЛИЦА 1. Неявно существующие утверждения

TABLE 1. Implicitly Existing Statements

Обозначение	Высказывание	Формула
F	Каждый риск ИБ можно нейтрализовать средством защиты до величины остаточного риска (ограничение 1)	
$F \rightarrow M$	Если каждый риск ИБ можно нейтрализовать, то он нейтрализуется	$\overline{F} \vee M$
R	Каждый риск ИБ входит в заявление о доверии	
J	Доверие максимально	
M	Каждый риск ИБ нейтрализован до величины остаточного риска	
S	Каждое заявление о доверии входит в аргумент доверия	
$R \wedge M \rightarrow S$	Если каждый риск ИБ входит в заявление о доверии, и он нейтрализован до величины остаточного риска, то каждое заявление о доверии входит в аргумент доверия (требование утверждения)	$\overline{R} \vee (\overline{M} \vee S)$
$S \rightarrow M \wedge R$	Если каждое заявление о доверии входит в аргумент доверия, значит, каждый риск ИБ нейтрализован до величины остаточного риска (следует из определения аргумента доверия)	$\overline{S} \vee (R \wedge M) = \overline{S} \vee \overline{R} \vee \overline{M}$

Используем подстановки для представления выражения $\overline{R} \vee (\overline{M} \vee S)$ через совокупность тождественных выражений, являющихся бинарными операциями:

$$(\overline{M} \vee S) = X,$$

$$(\overline{R} \vee \overline{M}) = Y,$$

$$(\overline{R} \vee S) = Z,$$

$$\overline{R} \vee (\overline{M} \vee S) = \{\overline{R} \vee X, Y \vee S, Z \vee \overline{M}\}.$$

Используем подстановки для представления выражения $\overline{S} \vee \overline{R} \vee \overline{M}$ через совокупность тождественных выражений, являющихся бинарными операциями:

$$\overline{S} \vee \overline{R} = X2,$$

$$\overline{R} \vee \overline{M} = Y2,$$

$$\overline{S} \vee \overline{M} = Z2,$$

$$\overline{S} \vee \overline{R} \vee \overline{M} = X2 \vee \overline{M}, Y2, Z2 \vee \overline{R}.$$

В этом случае система формул утверждения для доказательства необходимости примет вид:

$$J, \overline{R} \vee \overline{S}, \overline{R} \vee X, Y \vee S, Z \vee \overline{M}, \overline{F} \vee M, X2 \vee \overline{M}, \overline{S} \vee Y2,$$

$$Z2 \vee \overline{R}, F, R.$$

Если R – контрарный литерал, то для дизъюнктов $\{R, \overline{R} \vee X\}$ резольвентой будет $X = \overline{M} \vee S$.

Добавим резольвенту в систему формул:

$$J, \bar{R} \vee \bar{S}, \bar{R} \vee X, Y \vee S, Z \vee \bar{M}, \bar{F} \vee M, X2 \vee \bar{M}, \bar{S} \vee Y2, \\ Z2 \vee \bar{R}, F, R, X, \bar{M} \vee S.$$

Если S – контрарный литерал, то для дизъюнктов $\{\bar{R} \vee \bar{S}, \bar{M} \vee S\}$ резольвентой будет $\bar{M} \vee \bar{R}$.

Добавим резольвенту в систему формул:

$$J, \bar{R} \vee \bar{S}, \bar{R} \vee X, Y \vee S, Z \vee \bar{M}, \bar{F} \vee M, X2 \vee \bar{M}, \bar{S} \vee Y2, \\ Z2 \vee \bar{R}, F, R, X, \bar{M} \vee S, \bar{M} \vee \bar{R}.$$

Если R – контрарный литерал, то $R, \bar{M} \vee \bar{R} \Rightarrow \bar{M}$. Если M – контрарный литерал, то $\bar{M}, \bar{F} \vee M \Rightarrow \bar{F}$. Если F – контрарный литерал, то $F \vee 0, \bar{F} \vee 0 \Rightarrow \underline{0 \vee 0}$ тождественно ложный дизъюнкт. Следовательно, утверждение истинно.

Доказательство непротиворечивости достаточности:

$$\{R \wedge S\} \Rightarrow \{J\} = \{R \wedge S, \bar{J}\}.$$

Система формул утверждения в случае доказательства достаточности методом резолюций примет следующий вид (согласно методу резолюций, конъюнкции должны быть удалены):

$$\{R \wedge S, \bar{J}\} = \{R, S, \bar{J}\}.$$

С учетом неявных закономерностей:

$$R, S, J, \bar{R} \vee X, Y \vee S, Z \vee \bar{M}, X2 \vee \bar{M}, \bar{S} \vee Y2, Z2 \vee \bar{R}, \\ \bar{F} \vee M, F.$$

Если R – контрарный литерал, то $R, Z2 \vee \bar{R} \Rightarrow Z2 = \bar{S} \vee \bar{M}$. Если F и M – контрарные литералы, то $F, \bar{F} \vee M \Rightarrow M$, $\bar{S} \vee \bar{M} \Rightarrow \bar{S}$, соответственно. Если S – контрарный литерал, то $S \vee 0, \bar{S} \vee 0 \Rightarrow \underline{0 \vee 0}$ – тождественно ложный дизъюнкт. Следовательно, утверждение истинно.

Следствие 1. Если доверие максимально, то вероятность риска минимальна (1).

Следствие 2. Чтобы создать доверие в интегрированной ИС, необходимо описать активы и риски интегрируемых информационных систем.

Рассмотрим работу предложенной модели доверия на примере интеграции информационной системы персональных данных Петербургского государственного университета путей сообщения (далее – ИСПДн ПГУПС) и Государственной информационной системы «Современная цифровая образовательная среда» (далее – ГИС СЦОС), созданной согласно постановлению правительства от 16.11.2020 г. № 1836 «О государственной информационной системе «Современная цифровая образовательная среда».

5. Контрольный пример применения предложенной модели доверия

При интеграции обозначенных ИС требованиями бизнеса (целевых процессов) являются установленные в регламенте требования по подключению к ГИС СЦОС, такие как: приказ о назначении ответственных лиц образовательной организации на подключение к ГИС СЦОС, обеспечение требований к ИБ, заключающиеся в установке определенных средств защиты согласно регламенту. Минимальный класс – КС1 – является ограничением целевых процессов. Сведения об ИС образовательной организации, требования к составу отправляемых данных, минимальные требования к автоматизированному рабочему месту (далее – АРМ) операторов являются ограничениями блока А4.

Требование по наличию лицензии ФСТЭК России и ФСБ России для установки средств криптографической защиты информации (СКЗИ) является правовым ограничением на блок А1. Стоимость установки средств защиты организаций, имеющей лицензии на проведение установочных работ СКЗИ, является экономическим ограничением на блоки А3 и А4. Удовлетворение требований к пропускной способности сети представляется проблематичным, потому что подразумевает прокладку нового сетевого кабеля в культурно-охраняемом здании ПГУПС. Это физическое ограничение от блоков А3 и А4.

Рисками ИБ для двух информационных систем при интеграции являются те риски, которые описаны в актуальной модели угроз информационных систем. Например, в настоящий момент актуальными угрозами для ИСПДн ПГУПС являются: утечка акустической (речевой) информации, НСД с применением стандартных функций операционной системы, утечка информации путем преднамеренного копирования защищаемой информации на неучтенные (в том числе съемные) носители.

Активами информационной системы выступают:

- 1) данные, которые необходимо передавать в ходе интеграции (персональные данные студентов, информация о направлении подготовки и успеваемости);
- 2) АРМ, на которых производится обработка данных;
- 3) программное обеспечение, с помощью которого реализуется взаимодействие.

Заявлением доверия является сообщение о готовности к подключению, а аргументом доверия – совокупность структурированных заявлений о доверии, подтвержденных актами установки средств защиты информации (СЗИ), защите сведений об организации и ответственных сотрудниках.

Инструменты для создания доверия в ходе интеграции:

- персонал информационной системы: операторы и специалисты по ИБ (блок А2);
- аппаратные СЗИ: модули доверенной загрузки на АРМ оператора (блок А3);
- программные СЗИ: Secret Net Studio с дополнительным модулем межсетевого экрана, Kaspersky Endpoint Security для Windows, Континент TLS. Версия 2 (блок А4).

Остаточным риском ИБ являются риски от новых угроз и возможных инсайдеров в ИС. Свидетельством доверия в данном примере выступает закрывающий документ (акт установления интеграции) об успешной интеграции с ГИС СЦОС.

Пример формирования доверия по уровням

Блок А1, обеспечивающий уровень. Данные формируются и движутся между уровнями ИС согласно инструкции по эксплуатации ГИС СЦОС, утвержденной в ПГУПС (требования блока для А1). Активностями ИС на данном уровне являются передаваемые данные и АРМ оператора. Свидетельствами доверия с блоков А2–А4 будет информация о конфигурации, настройке СКЗИ и уровень образования сотрудников. Аргументом доверия обеспечивающего уровня (выходным параметром блока А1) являются листы ознакомления операторов с нормативными документами (инструкциями по взаимодействию с ГИС СЦОС, моделью угроз ПГУПС и т. д.), а заявлением о доверии для блоков А2–А4 будет требование о наличии обозначенных документов.

Блок А2, уровень персонала. Предполагается, что квалифицированный оператор несет меньше рисков, чем неквалифицированный, поэтому оператор должен иметь навыки работы с программными и аппаратными средствами (требования для блока А2). Свидетельством доверия уровня персонала (выходной параметр блока А2) выступает результат сдачи зачета на допуск к самостоятельной работе.

Блок А3, уровень аппаратного обеспечения. Требования бизнеса (целевых процессов) на данном уровне (ограничения для блока А3) – стоимость аппаратного обеспечения. Свидетельством доверия на данном уровне является аттестат АРМ оператора, который выдает испытательная лаборатория, имеющая лицензии ФСТЭК России и ФСБ России (выходной параметр блока А3). Ограничением, накладываемым блоком А3 на блок А2, является требование к квалификации оператора работать с данными аппаратными средствами. Ограничения, накладываемые блоком А4 на блок А3: процессор не ниже *i*-10, оперативная память не меньше 16 ГБ и т. д.

Блок А4, уровень программного обеспечения. Оператор формирует пакет данных для передачи из базы данных ИСПДн ПГУПС. Данные из БД ИСПДн

ПГУПС через защищенное соединение (TLS) передаются в ГИС СЦОС. Свидетельством доверия будет сообщение об успешной доставке информации до ГИС СЦОС. Ограничения для блока А4 от блока А1 состоят в требованиях настроить парольную политику ПГУПС (например, в парольной политике установлено требование для пароля не менее 8 символов, а программное средство устанавливает максимум пароль из 7 символов), а также программные СЗИ (например, включить межсетевой экран в Secret Net Studio). Ограничениями блока А4 на блок А2 является осведомленность оператора об особенностях работы с данными программными средствами. Аргументом доверия блока А4, который является аргументом доверия ко всей интегрированной ИС, выступает акт об успешном выполнении всех требований ИБ, результаты периодического контроля безопасности.

Если по результатам внедрения СЗИ вероятность риска стала $R = 0,002$, то, согласно (1), доверие в интегрированной ИС будет $D = 0,998$.

В ходе практической эксплуатации и, исходя из моделирования, установлено, что необходимое и достаточное условия существования доверия интегрируемых систем (ИСПДн ПГУПС и ГИС СЦОС) не выполнены. Необходимое и достаточное условие, сформулированное в Утверждении, связано с элементами структуры доверия (см. рисунок 3): *J* – цель создания доверия (на рисунке 3 не отражено явно); *R* – говорит о том, что каждой стрелке, обозначающей риск ИБ, ставится в соответствие стрелка, обозначающая заявление о доверии; *S* – требует, чтобы результирующая стрелка «Аргумент доверия интегрированной ИС» включала в себя все стрелки, описывающие аргументы доверия других уровней, которые, в свою очередь, включают в себя все стрелки, описывающие заявления о доверии. Следовательно, необходимое и достаточное условия существования доверия (см. Утверждение) на рисунке 3 выглядит как наличие всех «стрелочек» в структуре доверия, т. е. выполнение и наличие всех ограничений, аргументов, заявления доверия и т. д. По причине отсутствия некоторых ограничений существуют следующие возможности нарушения ИБ:

- удаленный доступ на АРМ оператора при подключении к ГИС СЦОС (нарушает аргумент доверия уровня персонала, что может привести к нарушению ИБ ИС);

- подключения другим оператором / посторонним лицом с помощью действующего сертификата оператора (управление ИС оператором, не ознакомленным с инструкцией по ее эксплуатации, является нарушением аргумента доверия уровня персонала);

- перемещение АРМ операторов в другие кабинеты (нарушает аргумент доверия обеспечивающего уровня, что может привести к утечке информации по визуальным каналам утечки);

- эксплуатация системы даже в случае увольнения работников, ответственных за интеграцию ИС (нарушает аргумент доверия уровня персонала, что приводит к сложностям при расследовании инцидента);

- подключение съемных носителей к АРМ оператора (возможен перенос вирусов и различной вредоносной информации, в том числе и такой, что ее не обнаружит последняя версия антивируса, т. е. нарушены аргументы доверия программного и аппаратного уровней).

Кроме того, выявлено отсутствие в нормативных документах (блок А1):

- ответственности за передачу неверных данных и / или вредоносной информации, что повышает риски нарушения ИБ на уровне персонала (блок А2);

- обоснование необходимости применения средств защиты и разъяснение последствий инцидентов ИБ; в результате персонал не видит необходимости в использовании СЗИ и может их отключить для более комфортной работы в ИС.

6. Заключение

В работе на основе анализа литературы сформулировано определение доверия как меры ИБ, дополняющей риск безопасности до полной группы событий и распространяющееся на каждый уровень иерархии ИС согласно модели FIST. На каждом уровне иерархии ИС доверие имеет свою специфику, но цель доверия, заключающаяся в снижении риска ИБ, возникающего при интеграции ИС, остается неизменной.

Список источников

1. Черных А.М. Основные направления интеграции федеральных государственных информационных систем и иностранных данных // Правовая информатика. 2018. № 2. С. 47–56. EDN: XRPRMT
2. Yan Z., Holtmanns S. Trust Modeling and Management: From Social Trust to Digital Trust // Computer Security, Privacy and Politics: current Issues, Challenges and Solutions. 2008. PP. 290–323. DOI:10.4018/978-1-59904-804-8.ch013
3. Chahal R.K., Kumar N., Batra S. Trust management in social Internet of Things: A taxonomy, open issues, and challenges // Computer Communications. 2020. Vol. 150. PP. 13–46. DOI:10.1016/j.comcom.2019.10.034
4. Burlov V.G., Gryzunov V.V. Evaluation of the effectiveness of geographic information systems adaptation to destabilizing factors // Journal of Physics: Conference Series. 2020. Vol. 1703. P. 012016. DOI:10.1088/1742-6596/1703/1/012016
5. Селифанов В.В., Гордеев А.С., Карманов И.Н. Требования по защите информации при межсетевом взаимодействии государственных информационных систем с иными информационными системами // Интерэкспо Гео-Сибирь. 2018. № 7. С. 277–282. EDN:YORFLV
6. Прокушев Я.Е., Пономаренко С.В., Пономаренко С.А. Моделирование процессов проектирования систем защиты информации в государственных информационных системах // Computational nanotechnology. 2021. Т. 8. № 1. С. 26–37. DOI:10.33693/2313-223X-2021-8-1-26-37. EDN:XJMNND
7. Грызунов В.В., Корниенко А.А., Глухарев М.Л., Крюков А.С. Выбор моделей доверия при интеграции распределенных информационных систем критического применения // Проблемы информационной безопасности. Компьютерные системы. 2021. № 4. С. 79–90. DOI:10.48612/jisp/ev3e-fmtu-x25h. EDN:VMALWC

Разработана модель доверия, отличающаяся охватом всех уровней интегрируемых информационных систем и позволяющая оценить доверие количественно. Согласно предложенной модели доверия, на каждом уровне иерархии ИС существует остаточный риск ИБ, из которого состоит остаточный риск интегрированной ИС. Расчет остаточного риска и доверия требует разработки соответствующих методик.

С продвижением от уровня (см. рисунок 3) с меньшим номером (А1) к уровню с большим номером (А4) появляется все больше ограничений, однако каждый последующий уровень в некоторых случаях может снизить остаточный риск других уровней ИС.

Необходимым и достаточным условием максимизации доверия в ИС при некоторых ограничениях является формирование аргумента доверия, включающего в себя все заявления о доверии интегрируемых ИС.

Адекватность модели продемонстрирована на контрольном примере интеграции ИСПДн ПГУПС и ГИС СЦОС. Данного примера достаточно для демонстрации адекватности модели, т. к. он отражает все связи в структуре модели и варианты использования модели.

Поскольку доверие по определению (см. выше) строится, исходя из удовлетворения целей безопасности, риски и активы формулируются согласно целям безопасности и моделям угроз, то можно утверждать, что предложенная модель доверия дает возможность обеспечить ИБ интегрируемых ИС, и значит, цель исследования достигнута.

Дальнейшим направлением исследований выступит разработка методики практического применения разработанной модели доверия.

8. Кругликов С.В., Дмитриев В.А., Степанян А.Б., Максимович Е.П. Информационная безопасность информационных систем с элементами централизации и децентрализации // Вопросы кибербезопасности. 2020. № 1(35). С. 2–7. DOI:10.21681/2311-3456-2020-01-02-07. EDN:HVFMTK
9. Шиверов П.К., Бондаренко В.В. Понятие доверия в контексте информационной безопасности // Международная конференция и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2016, Самара, Российская Федерация, 17–19 мая 2016). Самара: Самарский государственный аэрокосмический университет, 2016. С. 414–418. EDN:WMPXCP
10. Meeßen S.M., Thielsch M.T., Hertel G. Trust in Management Information Systems (MIS) // Zeitschrift für Arbeits-und Organisationspsychologie A&O. 2019. № 64. Iss. 1. PP. 6–16. DOI:10.1026/0932-4089/a000306
11. Maqableh M., Hmoud H.Y., Jaradat M., Masadeh R. Integrating an information systems success model with perceived privacy, perceived security, and trust: the moderating role of Facebook addiction // Heliyon. 2021. Vol. 7. Iss. 9. PP. 1–15. DOI:10.1016/j.heliyon.2021.e07899
12. Ettlie J.E., Tucci C., Gianiodis P.T. Trust, integrated information technology and new product success // European Journal of Innovation Management. 2017. Vol. 20. Iss. 3. PP. 406–427. DOI:10.1108/EJIM-12-2015-0128
13. McKnight H., Carter M., Clay P. Trust in technology: Development of a set of constructs and measures // DIGIT 2009 Proceedings. 2009. URL: <https://aisel.aisnet.org/digit2009/10> (Accessed 10.06.2024)
14. Ngo-Ye T.L., Nazareth D.L., Choi J.J. Trust in security as a service: a theoretical model // Issues in Information Systems. 2020. Vol. 21. Iss. 2. PP 64–74.
15. Park S. Multifaceted trust in tourism service robots // Annals of Tourism Research. 2020. Vol. 81. P. 102888. DOI:10.1016/j.annals.2020.102888
16. Ramos F.L., Ferreira J.B., Freitas A.S., Rodrigues J.W. The Effect of Trust in the Intention to Use m-banking // BBR. Brazilian Business Review. 2018. Vol. 15. Iss. 2. PP. 175–191. DOI:10.15728/bbr.2018.15.2.5
17. Putra G.C., Astiti N.P.Y., Gunadi G.N.B. The Exploring of Trust that Influences Customer's Intention to Use FinTech M-Banking Application on Regional Banks // International Journal of Economics and Business Administration. 2020. Vol. 8. Iss. 4. PP. 407–421.
18. Иткес А.А. Объединение моделей логического разграничения доступа для сложноорганизованных распределенных информационных систем // Проблемы информатики. 2010. № 1(5). С. 85–94. EDN:NBRZPN
19. Глухова Л.В., Губанова С.Е. Некоторые аспекты менеджмента информационной безопасности промышленных комплексов // Вестник Волжского университета им. В.Н. Татищева. 2015. №3(34). С. 135–144. EDN:VBWJDX
20. Gryzunov V.V. Conceptual Model for Adaptive Control of a Geographic Information System under Conditions of Destabilization // Automatic Control and Computer Sciences. 2021. Vol. 55. Iss. 8. PP. 1222–1227. DOI:10.3103/S0146411621080381
21. Покровский И.А. Разобраться в понятиях // Безопасность информационных технологий. 2023. Т. 30. №. 2. С. 21–22.
22. Калашников А.О., Бугайский К.А., Бирин Д.С., Дерябин Б.О., Цепенда С.О., Табаков К.В. Применение логико-вероятностного метода в информационной безопасности (часть 1) // Вопросы кибербезопасности. 2023. №. 4(56). С. 23–32. DOI:10.21681/2311-3456-2023-4-23-32. EDN:GIHSBN
23. Zefferer T., Prunster B., Kollmann C., Corici A.A. A Security-Evaluation Framework for Mobile Cross-Border e-Government Solutions // Proceedings of the 24th Annual International Conference on Digital Government Research (Gdansk, Poland, 11–14 July 2023). New York: Association for Computing Machinery, 2023. PP. 536–543. DOI:10.1145/3598469.359852
24. Phiayura P., Teerakanok S. A Comprehensive Framework for Migrating to Zero Trust Architecture // IEEE Access. 2023. Vol. 11. PP. 19487–19511. DOI:10.1109/ACCESS.2023.3248622
25. Ahmadi S. Zero Trust Architecture in Cloud Networks: Application, Challenges and Future Opportunities // Journal of Engineering Research and Reports. 2024. Vol. 26. Iss. 2. PP. 215–228. DOI:10.9734/jerr/2024/v26i21083
26. Khan M.J. Zero trust architecture: Redefining network security paradigms in the digital age // World Journal of Advanced Research and Reviews. 2023. Vol. 19. Iss. 3. PP. 105–116. DOI:10.30574/wjarr.2023.19.3.1785
27. Грызунов В.В. Модель геоинформационной системы FIST, использующей туманные вычисления в условиях дестабилизации // Вестник Дагестанского государственного технического университета. Технические науки. 2021. Т. 48. № 1. С. 76–89. DOI:10.21822/2073-6185-2021-48-1-76-89. EDN:IDEYPX
28. Gryzunov V.V. Model of Purpose Aggressive Actions on the Information-Computing System // Proceedings of the 3rd International Conference on Human Factors in Complex Technical Systems and Environments (ERGO, St. Petersburg, Russia, 04–07 July 2018). IEEE, 2018. PP. 119–121. DOI:10.1109/ERGO.2018.8443814
29. Gryzunov V.V., Bondarenko I.Yu. A Social Engineer in Terms of Control Theory // Proceedings of the 3rd International Conference on Human Factors in Complex Technical Systems and Environments (ERGO, St. Petersburg, Russia, 04–07 July 2018). IEEE, 2018. PP. 202–204. DOI:10.1109/ERGO.2018.8443835
30. Gryzunov V., Gryzunova D. Problems of Providing Access to a Geographic Information System Processing Data of Different Degrees of Secrecy // Khanna K., Estrela V.V., Rodrigues J.J.P.C. (eds.) Cyber Security and Digital Forensics. Lecture Notes on Data Engineering and Communications Technologies. Singapore: Springer, 2022. Vol. 73. PP. 191–198. DOI:10.1007/978-981-16-3961-6_17
31. Ананьев И.В., Серова Е.Г. Области эффективного применения нотации IDEF0 для задач описания бизнес-процессов // Вестник Санкт-Петербургского университета. Менеджмент. 2008. № 2. С. 161–172. EDN:JUBTXH
32. Канев С.А. Акцент на эффект. Определение характеристик эффективности использования информационных активов компаний // Креативная экономика. 2010. № 8(44). С. 42–47. EDN:MSVWFV
33. Манжосов А.В., Болодурина И.П., Сабуров В.С., Долгушев Н.А. Разработка специальной классификации информационных активов в сфере информационной безопасности // Вестник Пермского университета. Математика. Механика. Информатика. 2022. № 4(59). С. 54–60. DOI:10.17072/1993-0550-2022-4-54-60. EDN:ZHZNWB

References

1. Chernykh A. The main directions of federal state information systems and classified data. *Legal Informatics*. 2018;2: 47–56. (in Russ.) EDN:XRPRMT
2. Yan Z., Holtmanns S. Trust Modeling and Management: From Social Trust to Digital Trust. *Computer Security, Privacy and Politics: current Issues, Challenges and Solutions*. 2008:290–323. DOI:10.4018/978-1-59904-804-8.ch013
3. Chahal R.K., Kumar N., Batra S. Trust management in social Internet of Things: A taxonomy, open issues, and challenges. *Computer Communications*. 2020;150:13–46. DOI:10.1016/j.comcom.2019.10.034
4. Burlov V.G., Gryzunov V.V. Evaluation of the effectiveness of geographic information systems adaptation to destabilizing factors. *Journal of Physics: Conference Series*. 2020;1703:012016. DOI:10.1088/1742-6596/1703/1/012016
5. Selifanov V.V., Gordeev A.S., Karmanov I.N. Requirements for information security in cross-network interaction of the state information systems with other information systems. *Interexpo GEO-Sibiria*. 2018;7:277–282. (in Russ.) EDN:YORFLV
6. Prokushev Ya.E., Ponomarenko S.V., Ponomarenko S.A. The Modeling of information security system design processes in state information systems. *Computational Nanotechnology*. 2021;1:26–37. (in Russ.) DOI:10.33693/2313-223X-2021-8-1-26-37. EDN:XJMNND
7. Gryzunov V.V. Kornienko A.A., Glukharev M.L., Kryukov A.S. Selection of trust models when integrating distributed information systems of critical application. *Information Security Problems. Computer Systems*. 2021;479–90. (in Russ.) DOI:10.48612/jisp/ev3e-fmtu-x25h. EDN:VMALWC
8. Kruglikov S., Dmitriev V., Stepanian A., Maksimovich E. Information security of information systems with elements of centralization and decentralization. *Voprosy kiberbezopasnosti*. 2022;1(35):2–7. (in Russ.) DOI:10.21681/2311-3456-2020-01-02-07. EDN:HVFMFK
9. Shiverov P.K., Bondarenko V.V. The concept of trust in the context of information security. *Proceedings of the International Conference and Youth School on Information Technologies and Nanotechnologies, ITNT-2016, 17–19 May 2016, Samara, Russian Federation*. Samara: Samara State Aerospace University Publ.; 2016. p.414–418. (in Russ.) EDN:WMPXCP
10. Meeßen S.M., Thielsch M.T., Hertel G. Trust in Management Information Systems (MIS). *Zeitschrift für Arbeits-und Organisationspsychologie A&O*. 2019;64(1):6–16. DOI:10.1026/0932-4089/a000306
11. Maqableh M., Hmoud H.Y., Jaradat M., Masadeh R. Integrating an information systems success model with perceived privacy, perceived security, and trust: the moderating role of Facebook addiction. *Heliyon*. 2021;7(9):1–15. DOI:10.1016/j.heliyon.2021.e07899
12. Ettlie J.E., Tucci C., Gianiodis P.T. Trust, integrated information technology and new product success. *European Journal of Innovation Management*. 2017;20(30):406–427. DOI:10.1108/EJIM-12-2015-0128
13. McKnight H., Carter M., Clay P. Trust in technology: Development of a set of constructs and measures. *DIGIT 2009 Proceedings*. 2009. URL: <https://aisel.aisnet.org/digit2009/10> [Accessed 10.06.2024]
14. Ngo-Ye T.L., Nazareth D.L., Choi J.J. Trust in security as a service: a theoretical model. *Issues in Information Systems*. 2020;21(2):64–74.
15. Park S. Multifaceted trust in tourism service robots. *Annals of Tourism Research*. 2020;81:102888. DOI:10.1016/j.annals.2020.102888
16. Ramos F.L., Ferreira J.B., Freitas A.S., Rodrigues J.W. The Effect of Trust in the Intention to Use m-banking. *BBR. Brazilian Business Review*. 2018;15(2):175–191. DOI:10.15728/bbr.2018.15.2.5
17. Putra G.C., Astiti N.P.Y., Gunadi G.N.B. The Exploring of Trust that Influences Customer's Intention to Use FinnTech M-Banking Application on Regional Banks. *International Journal of Economics and Business Administration*. 2020;8(4):407–421.
18. Itkes A.A. Combining logical access control models for complex distributed information systems. *Problemy informatiki*. 2010;3:85–94. (in Russ.) EDN:NBRZPN
19. Glukhova L.V., Gubanova S.E. Some aspects of information security management of industrial complexes. *Vestnik of Volzhsky University named after V.N. Tatishchev*. 2015;3(34):135–144. (in Russ.) EDN:VBWJDX
20. Gryzunov V.V. Conceptual Model for Adaptive Control of a Geographic Information System under Conditions of Destabilization. *Automatic Control and Computer Sciences*. 2021;55(8):1222–1227. DOI:10.3103/S0146411621080381
21. Pokrovskij I.A. To understand the concepts. *IT Security*. 2023;2:21–22. (in Russ.).
22. Kalashnikov A.O., Bugajskij K.A., Birin D.S., Dereabin B.O., Tsenda S.O., Tabakov K.V. Application of the logical-probabilistic method in information security (part 1). *Voprosy kiberbezopasnosti*. 2023;4(56):23–32. (in Russ.) DOI:10.21681/2311-3456-2023-4-23-32. EDN:GIHSBN
23. Zefferer T., Prunster B., Kollmann C., Corici A.A. A Security-Evaluation Framework for Mobile Cross-Border e-Government Solutions. *Proceedings of the 24th Annual International Conference on Digital Government Research, 11–14 July 2023, Gdansk, Poland*. New York: Association for Computing Machinery; 2023. p.536–543. DOI:10.1145/3598469.359852
24. Phiayura P., Teerakanok S. A Comprehensive Framework for Migrating to Zero Trust Architecture. *IEEE Access*. 2023;11:19487–19511. DOI:10.1109/ACCESS.2023.3248622
25. Ahmadi S. Zero Trust Architecture in Cloud Networks: Application, Challenges and Future Opportunities. *Journal of Engineering Research and Reports*. 2024;26(2):215–228. DOI:10.9734/jerr/2024/v26i21083
26. Khan M.J. Zero trust architecture: Redefining network security paradigms in the digital age. *World Journal of Advanced Research and Reviews*. 2023;19(3):105–116. DOI:10.30574/wjarr.2023.19.3.1785
27. Gryzunov V.V. FIST geoinformation system model using fog computing in destabilization. *Herald of Dagestan State Technical University. Technical Sciences*. 2021;48(1):76–89. (in Russ.) DOI:10.21822/2073-6185-2021-48-1-76-89. EDN:IDEYPX
28. Gryzunov V.V. Model of Purpose Aggressive Actions on the Information-Computing System. *Proceedings of the 3rd International Conference on Human Factors in Complex Technical Systems and Environments, ERGO, 04–07 July 2018, St. Petersburg, Russia*. IEEE; 2018. p.119–121. DOI:10.1109/ERGO.2018.8443814

29. Gryzunov V.V., Bondarenko I.Yu. A Social Engineer in Terms of Control Theory. *Proceedings of the 3rd International Conference on Human Factors in Complex Technical Systems and Environments, ERGO, 04–07 July 2018, St. Petersburg, Russia*. IEEE; 2018. p.202–204. DOI:10.1109/ERGO.2018.8443835
30. Gryzunov V., Gryzunova D. Problems of Providing Access to a Geographic Information System Processing Data of Different Degrees of Secrecy. In: *Khanna K., Estrela V.V., Rodrigues J.J.P.C. (eds.) Cyber Security and Digital Forensics. Lecture Notes on Data Engineering and Communications Technologies, vol.73*. Singapore: Springer; 2022. p.191–198. DOI:10.1007/978-981-16-3961-6_17
31. Anan'ev I.V., Serova E.G. Areas of effective application of IDEF0 notation for tasks of describing business processes. *Vestnik of Saint Petersburg University. Management*. 2008;2:161–172. (in Russ.) EDN:JUBTXH
32. Kanev S.A. Determination of characteristics of a company's informational assets use efficiency. *Creative Economy*. 2010;8(44):42–47. (in Russ.) EDN:MSVWFV
33. Manzhosov A.V., Bolodurina I.P., Saburov V.S., Dolgushev N.A. Development of a special classification of information assets in the information security field. *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*. 2022;4(59):54–60. (in Russ.) DOI:10.17072/1993-0550-2022-4-54-60. EDN:ZHZWNB

Статья поступила в редакцию 14.06.2024; одобрена после рецензирования 15.07.2024; принята к публикации 25.07.2024.

The article was submitted 14.06.2024; approved after reviewing 15.07.2024; accepted for publication 25.07.2024.

Информация об авторах:

ГРЫЗУНОВ Виталий Владимирович	доктор технических наук, доцент, профессор кафедры прикладной информатики и информационных технологий Санкт-Петербургского университета ГПС МЧС России  https://orcid.org/0000-0003-4866-217X
КРЮКОВ Александр Сергеевич	преподаватель кафедры информационного и интеллектуального права, цифровых технологий и инноватики Российского государственного университета правосудия  https://orcid.org/0000-0002-4633-8635
ШЕСТАКОВ Александр Викторович	доктор технических наук, старший научный сотрудник, ведущий научный сотрудник отдела информационного обеспечения населения и технологий информационной поддержки РСЧС и пожарной безопасности Санкт-Петербургского университета ГПС МЧС России  https://orcid.org/0000-0002-8462-6515
ЗИКРАТОВ Игорь Алексеевич	доктор технических наук, профессор, профессор кафедры информационных управляющих систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича  https://orcid.org/0009-0000-2939-1971

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

Обзорная статья

УДК 004.056(075.58)

<https://doi.org/10.31854/1813-324X-2024-10-4-126-141>

Прогресс в теории прикладной криптографии: обзор и некоторые новые результаты. Часть 1. Ключевая криптография

✉ Валерий Иванович Коржик¹, val-korzhik@yandex.ru✉ Виктор Алексеевич Яковлев¹, yakovlev.va@sut.ru✉ Борис Викторович Изотов², izotov.b@yandex.ru✉ Владимир Сергеевич Старостин¹, vm.ffp@sut.ru✉ Михаил Викторович Буйневич³, bmv1958@yandex.ru

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

²ЗАО «Научные приборы»,
Санкт-Петербург, 198095, Российская Федерация

³Санкт-Петербургский университет ГПС МЧС России,
Санкт-Петербург, 196105, Российская Федерация

Аннотация

В настоящей работе, состоящей из двух частей, представлены как уже опубликованные ранее (но труднодоступные) результаты, так и новые. **Актуальность** данной работы заключается, во-первых, в том, что в последнее время получен ряд новых результатов в области прикладной криптографии, которые нуждаются как в разъяснении, так и в практическом применении. Именно это и является основной **целью** настоящей работы. **Постановка проблемы** в первой части статьи касается сложности взлома симметричных шифров, в то время как во второй части статьи обсуждается так называемая бесключевая криптография, а именно: концепция канала прослушивания, реализация каналов связи, позволяющих обеспечить информационную безопасность без процедуры обмена ключами между легитимными корреспондентами. В работе широко используются **методы** прикладной математики, а именно: алгебры, чисел, вероятностей и теории информации. Также используется компьютерное моделирование. **Новизна** первой части работы заключается в следующем: во-первых, проясняется смысл ограничения времени жизни ключа для различных режимов симметричного шифрования, во-вторых, поясняется подход взлома шифра с использованием квантовых компьютеров, в-третьих, подробно исследуется аутентификация ключа для протокола Диффи – Хеллмана на основе технологии «спаривания» мобильных устройств. Во второй части статьи представлена уязвимость криптосистемы Дина – Голдсмита при некотором расширении атак. **Основными результатами** данной статьи являются: оценка времени жизни ключа для симметричного шифрования в режиме СВС, прояснение алгоритма Гровера взлома симметричных шифров методом грубой силы, разработка метода аутентификации значений Диффи – Хеллмана на основе предварительно распределенных последовательностей, выбор шифров, позволяющих работать с протоколом Шамира без предварительного разделения ключа, взлом протокола Дина – Голдсмита при некоторых условиях, доказательство факта о возможной взламываемости протокола разделения ключа по бесшумным каналам связи. **Практическое применение** результатов статьи заключается в стимулировании правильного выбора шифров и их параметров с целью обеспечения их устойчивости к различным атакам и большего внимания к алгоритмам бесключевой криптографии.

Ключевые слова: срок действия ключей, алгоритмы Гровера и Шора, аутентификация ключей, квантовые компьютеры, кодовое зашумление, коммутативная криптография, бесключевая криптография

Ссылка для цитирования: Коржик В.И., Яковлев В.А., Изотов Б.В., Старостин В.С., Буйневич М.В. Прогресс в теории прикладной криптографии: обзор и некоторые новые результаты. Часть 1. Ключевая криптография // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 126–141. DOI:10.31854/1813-324X-2024-10-4-126-141. EDN:NNNFBU

Review research

<https://doi.org/10.31854/1813-324X-2024-10-4-126-141>

Advance in Applied Cryptography Theory: Survey and New Results. Part 1. Key Cryptography

 Valery I. Korzhik¹✉, val-korzhik@yandex.ru

 Viktor A. Yakovlev¹, yakovlev.va@sut.ru

 Boris V. Izotov², izotov.b@yandex.ru

 Vladimir S. Starostin¹, vm.ffp@sut.ru

 Mikhail V. Buinevich³, bmv1958@yandex.ru

¹The Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

²JSC Scientific Instruments,
St. Petersburg, 198095, Russian Federation

³Saint Petersburg University of State Fire Service of Emercom of Russia,
St. Petersburg, 196105, Russian Federation

Annotation

*In the current paper, consisting from two parts, are presented both results already published before (but hard for access) and new once. **Actuality** of this work is firstly in a fact that recently has been obtained a number of new results in area of applied cryptography that are needed both in a clarification and be put into practice. This is namely the main goal of the current paper. The **setting problem** in the first paper part concerns to a complexity of symmetric cipher breaking while in the second part of the paper is discussed, so called, keyless cryptography, namely: wiretap channel concept, execution of communication channels which allow to provide information security without of key exchange procedure between legal correspondences. In the part widely used **methods** of applied mathematics, namely: algebra, number, probability and information theories. Computer simulation also used there. A **novelty** of the first part of work consists in the following: first of all it is clarified the sense of a key lifetime limitation for different symmetric cipher modes, secondly, it is explained an approach of cipher breaking by the use of quantum computers, finely, the key authentication for the Diffie – Hellman protocol based on the mobile device pairing technology is investigated in detail. In the second part of the current paper has been presented a vulnerability of Dean – Goldsmith cryptosystem under some extension of attacks. The **main results** of this paper are: estimation of the key lifetime of single key for symmetric cipher in CBC mode, clarifying of Grover's algorithm breaking of symmetric ciphers by brute force attack, development of a method for authentication of Diffie – Hellman values based on pre-distributed sequences, selection of ciphers which allow to execute with Shamir's protocol without any key sharing in advance, breaking of Dean – Goldsmith protocol under some conclusions, proof the fact regarding of a possible breakability of the key sharing protocol over noiseless communication channels. **Practical application** of paper results consists in the fact of stimulation the correct choice of ciphers and their parameters in order to provide their resistance to different attacks and more attention to algorithms of keyless cryptography.*

Keywords: key expiration, Grover and Shor algorithms, Diffie – Hellman key authentication, quantum computers, code noising, commutative cryptography, keyless cryptography

For citation: Korzhik V.I., Yakovlev V.A., Izotov B.V., Starostin V.S., Buinevich M.V. Advance in Applied Cryptography Theory: Survey and New Results. Part 1. Key Cryptography. *Proceedings of Telecommunication Universities*. 2024;10(4):126–141. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-126-141. EDN:NNNFBU

ВВЕДЕНИЕ

Нет сомнения, что вопросы теории в прикладной криптографии имеют большое значение для изучения проблем информационной безопасности.

Причем это распространяется как на обучение в ВУЗах, так и на поисковые исследования в данной области. Настоящим прорывом здесь явилось создание ГОСТа (и последующая реализация отече-

ственных алгоритмов шифрования/дешифрования), впервые представленного в 1989 г. [1] и затем усовершенствованного в 2015 г. [2], а также появление стандартов на электронные (цифровые) подписи [3] и хэш-функции [4].

Значимость этих «государственных» документов состоит в том, что методы шифрования, наверное, впервые в СССР, а затем и в РФ, разрешались для использования не только специальным ведомствам в определенных областях, но и гражданским организациям и бизнес-сообществу, легитимизовав их права на конфиденциальную информацию. Что, конечно, не отрицает возможность и существование «закрытых» стандартов в этой области.

Весьма важным в этой связи является то обстоятельство, что появление «открытых» стандартов шифрования стимулировало издание целого ряда зарубежных и отечественных монографий по данному направлению науки, а также проведение исследований «гражданскими» специалистами [5–7]. Более того, в ряде университетов были созданы и начали читаться курсы по основам криптографии, обеспеченные соответствующими учебниками и учебными пособиями [8–9].

Однако некоторые теоретические достижения в области прикладной криптографии еще не заняли должного места в вузовской научно-педагогической деятельности и, тем более, не стали широким достоянием ученых и специалистов в области информационной безопасности; что, по мнению авторов, является серьезным упущением. Поэтому целью данной статьи является попытка представить в систематизированном виде как малоизвестные или редко применяемые на практике, так и отдельные авторские научные результаты, частично опубликованные в отечественных и зарубежных высокорейтинговых изданиях; некоторые из них являются впервые представленным на апробацию обобщением ранее известных результатов.

Статья состоит из двух частей, описывающих прогресс в области ключевой и бесключевой криптографии. Последний термин не придуман авторами настоящей статьи, а был предложен еще в 1983 г. Б. Алперном и Ф. Шнайдером [10]. В последующем по этой тематике проф. В.И. Коржилом вместе с соавторами было опубликовано свыше десятка статей в зарубежных журналах и трудах международных конференций, которые, к сожалению, не все легко доступны для отечественного читателя. Особого внимания, на наш взгляд, заслуживает глава с аутентичным названием «Advance in Keyless Cryptography» (*перев. на русск. «Прогресс в бесключевой криптографии»*) в монографии [11], изданной InterTech в 2022 г.

В первой части уточняются понятие стойкости для блочных шифров, описываются пути использования квантовых компьютеров для взлома симметричных и асимметричных криптосистем и представляются усовершенствованные протоколы аутентификации пользователей компьютерных сетей.

Во второй части рассматриваются так называемые методы обеспечения секретности информации на физическом уровне, то есть с использованием физических свойств каналов связи. К таким свойствам относятся: присутствие обратной связи, наличие шума, квантовые эффекты, многолучевость и использование смарт-антенн.

Авторы надеются, что предлагаемая вниманию читателей статья послужит важным стимулом для развития исследований в данных направлениях, причем как теории, так и в продвижении для практического использования содержащихся в ней основных научных результатов.

1. СТОЙКОСТЬ БЛОКОВЫХ ШИФРОВ ПРИ ИСПОЛЬЗОВАНИИ РАЗЛИЧНЫХ МОД ШИФРОВАНИЯ

Как уже упоминалось во введении, в настоящее время доступен для практического использования целый ряд «открытых» стандартов шифрования для симметричных блочных шифров. Напомним, что блочными называются такие шифры, которые однозначно отображают двоичные последовательности криптограмм такой же длины и используют во время шифрования определенное количество блоков открытого текста одного и того же секретного ключа; причем тот же самый ключ используется и для дешифрования криптограмм обратно в блок открытых сообщений. Отсюда и термин «симметричные шифры», что относится именно к ключам.

Очевидно, что основным требованием к таким, как, впрочем, и к другим шифрам, является их стойкость, то есть устойчивость к возможности дешифрования криптограмм без знания ключа. Зачастую такое воздействие называют «взломом шифра». В современной постановке этой задачи обычно предполагается, что при поиске ключа возможно знание не только блоков криптограммы, но и некоторого количества соответствующих им блоков открытых сообщений (так называемая, «атака с частично известным открытым сообщением»). Более того, иногда возможно навязывание «противником» при шифровании заранее выбранных блоков открытого сообщения («атака с частично выбранным открытым сообщением»). Есть версия, что именно такой атакой США взломали «Пурпурный шифр», использовавшийся японцами

во время морских сражений на Тихом океане во время Второй Мировой войны.

Что касается атаки, упомянутой первой, то одному из авторов настоящей статьи от криптографа, занимавшегося вскрытием немецких шифров периода ВОВ, стало известно, что это ему достаточно легко удавалось, поскольку почти любая военная криптограмма в Третьем Рейхе заканчивалась фразой «Neil Hitler!».

Попытка создания идеального секретного шифра официально датируется 1948 г., когда «вышла в свет» и была рассекречена работа К. Шеннона Communication Theory of Secrecy Systems [12], хотя такой алгоритм был запатентован еще раньше Г. Вернамом [13]; справедливости ради надо отметить, что шифр носит его имя – «шифр Вернама». Однако заслуга К.Э. Шеннона состояла в том, что он рассчитал минимально известную длину криптограммы (измеряемую, например, числом символов), для которой существует единственно возможное открытое сообщение. Эта величина была названа им «расстоянием единственности» (РЕ).

Заметим, что это единственное открытое сообщение, тогда может быть найдено при помощи полного перебора всех возможных ключей, называемой обычно «грубой лобовой атакой». Несколько неожиданным оказался тот факт, что для достаточно избыточных сообщений (таких, например, как текст на каком-либо естественном языке) РЕ удивительно мало – см. например, формулу Шеннона для РЕ в зависимости от величины энтропии источника и длины ключа в [8]. Поэтому, если необходимо зашифровать «длинные сообщения коротким ключом», следует выбирать такой ключ, который безусловно является непереборным, то есть все его варианты невозможно опробовать ни при каком разумном времени и/или реалистичном объеме вычислительного оборудования.

Однако «атака перебором» не является единственно возможной. Существует еще множество других атак (не экспоненциальной сложности), к которым относятся, например: линейный и дифференциальный криптоанализ, решение нелинейных уравнений, атака на основе принципа максимального правдоподобия, атака со связкой ключей и т. д.; см. некоторые примеры в [8]. Поэтому возникает естественный вопрос – можно ли вывести нижнюю границу для числа элементарных вычислений при наилучшей (пусть даже и неизвестной) атаке по взлому блочного шифра с известными алгоритмами шифрования/дешифрования и заданными параметрами (длиной блока шифра и его ключа)?

Более того, как известно, блочные шифры могут использоваться в различных модах, то есть, формировать каждый блок криптограмм с использованием нескольких блоков открытого текста.

Так, мода с зацеплением блоков (СВС-мода, аббр. от англ. Cipher Block Chaining) реализуется в виде:

$$\begin{aligned} E_i &= f_k(E_{i-1} \oplus M_i), i = 1, 2, \dots, \\ M_i &= E_{i-1} \oplus g_k(E_i), i = 1, 2, \dots, \end{aligned} \quad (1)$$

где E_i, E_{i-1} – i -й и $i-1$ -й блоки криптограмм; M_i – i -й блок сообщения; $f_k(\cdot)$ – функция шифрования на ключе K ; $g_k(\cdot)$ – функция дешифрования на ключе K ; $\oplus$ – операция побитового сложения по mod 2.

В [8] описаны также моды с обратной связью по криптограмме и по сообщению.

Поэтому возникает следующий естественный вопрос – как выбор моды влияет на стойкость блочного шифра? Интуитивно ясно, что чем больше блоков будет зашифровано без смены ключа, тем менее стойким может оказаться шифр, поскольку атакующим доступна возможность анализировать больший объем статистики. Отсюда логично возникает уточнение вопроса о стойкости шифра – сколько (максимум) блоков можно зашифровать без смены ключа, чтобы шифр оставался стойким?

В точной формулировке, при «атаке с частично известными сообщениями», эта задача может быть сформулирована следующим образом. Пусть $N = U + V$ блоков открытого текста были зашифрованы с использованием одного ключа. Обозначим эти блоки как $P_1, P_2, \dots, P_U, P'_1, P'_2, \dots, P'_V$ и соответствующие блоки криптограммы как $C_1, C_2, \dots, C_U, C'_1, C'_2, \dots, C'_V$, причем первые « U » блоков полагаются известными, а последние « V » блоков – неизвестными.

Обозначим через π максимальную приемлемую для легального собственника шифра величину вероятности получения дополнительной информации о неизвестной части открытого текста $P'_1, P'_2, \dots, P'_V$, то есть вероятность его риска. Максимальное количество блоков, которое нужно зашифровать на одном ключе без потери информации о блоках сообщений, обозначим через $N_{\text{макс}}$. Основная проблема состоит в том, чтобы найти функциональную зависимость между $N_{\text{макс}}$ и π , причем для различных мод шифрования.

В работе [14] доказано, что если блочный шифр является идеальным (то есть представляет собой случайные подстановки сообщений в криптограммы), то для максимального количества блоков, зашифрованных на одном ключе и при использовании, например, СВС-моды, будет справедливо следующее неравенство:

$$N_{\text{макс}} \leq 2^{\frac{n}{2}+1} \sqrt{\ln\left(\frac{1}{1-\pi}\right)}. \quad (2)$$

Для получения оценочных формул подобных (2), связывающих основные параметры мод шифрования $N_{\text{макс}}$ и π , использовались различные ва-

рианты метода, основанного на так называемом «парадоксе дня рождения» (в основе этого метода при использовании двоичного блочного шифра лежит тот факт, что при последовательном случайном и равномерном выборе блоков открытых сообщений длиной n произойдет совпадение (коллизия) двух таких блоков с вероятностью $\frac{1}{2}$ при проведении примерно $2^{n/2}$ испытаний).

Приведем пример расчета границы для $N_{\text{макс}}$. Пусть π задана как вероятность утечки любого количества информации о неизвестной части блоков открытого текста $P'_1, P'_2, \dots, P'_V$, то есть практически вероятность риска частичной компрометации. Тогда, если выбрать достаточно малую вероятность, допустим 10^{-6} , а длину блока $n = 64$, то по формуле (2) получим: $N_{\text{макс}} \leq 8 \cdot 10^6$ блоков. При задании же большей вероятности риска, например $\pi = 0,1$, по (2) получим: $N_{\text{макс}} \leq 2,6 \cdot 10^9$ блоков, – что значительно больше, чем в предыдущем примере.

Если же шифр не идеальный (то есть не является случайной подстановкой) и обладает лишь свойством биективности (то есть, однозначности шифрования/дешифрования, что обычно имеет место для таких известных стандартов, как ГОСТ, DES, AES), то при вычислении максимально возможного числа блоков при шифровании на одном ключе с отсутствием утечки информации необходимо, как отмечено в работе [14], учесть эту «неидеальность», что по мнению авторов этой статьи требует проведения дополнительных исследований.

Неравенства, аналогичные (2), были получены авторами [14] и для таких мод шифрования, как OFB (аббр. от англ. Output Feed Back, обратная связь по выходу внутреннего блочного шифра) и CFB (аббр. от англ. Cipher Feed Back, обратная связь по криптограмме). Что же касается моды ECB (аббр. от англ. Electronic CodeBook, электронная шифровальная книга), то для нее тривиально $N_{\text{макс}} = 1$, так как даже при использовании одного и того же ключа только для двух блоков, при повторении блоков открытого текста получаем одинаковые блоки криптограмм, то есть утечку дополнительной информации.

Итак, вопрос о нахождении строгой границы для $N_{\text{макс}}$ при неидеальном шифре, насколько известно авторам настоящей статьи, остается открытым.

2. ВОЗМОЖНОСТИ УПРОЩЕНИЯ КРИПТОАНАЛИЗА ПРИ ИСПОЛЬЗОВАНИИ КВАНТОВЫХ КОМПЬЮТЕРОВ

Прежде всего необходимо подчеркнуть, что, по крайней мере, по своему целевому предназначению квантовый компьютер (КК) коренным образом отличается от квантовой криптографии (ККР);

можно метафорически заметить это отличие на примере понятий «Государь» и «Милостивый государь». Сформулируем эти различные цели: КК, применяемый для решения проблем криптоанализа, предназначен для отыскания секретных ключей, как для симметричных, так и для несимметричных криптосистем, в то время как ККР – для построения метода защиты от перехвата конфиденциальной информации (прежде всего ключевой) при ее передаче по незащищенному от перехвата каналу связи.

Говоря о принципах построения КК, следует отметить среди их основных задач сокращение вычислительной сложности при выполнении криптоанализа. Здесь обычно рассматривают две ее разновидности: умеренное (квадратичное) уменьшение сложности вычислений – или алгоритм Гровера (АГ), и переход от экспоненциальной зависимости количества вычислений от размеров криптоключа к полиномиальной – или алгоритм Шора (АШ). Причем АГ рассмотрим более подробно, а для АШ представим лишь оценки для количества вычислений ввиду достаточной полноты его описания в доступных источниках.

Вообще говоря, АГ может использоваться далеко не только для криптоанализа, а например, для упрощенного решения задач поиска в любой неструктурированной базе данных. Постановка такой задачи имеет следующий вид: пусть имеется некоторое множество M элементов, снабженных нумерацией и, возможно еще, уникальными идентификаторами (ID), состоящими из N «неструктурированных» (то есть, которые нельзя как-то упорядочить) элементов; требуется определить номер (или ID) одного из элементов этого множества, который известен решающему задачу.

Математическая модель этой проблемы описывается следующим образом: пусть задана некоторая булева функция $f(x)$, $x \in M$, причем $f(x) = 1$ тогда и только тогда, когда элементы x удовлетворяют критерию поиска, и $f(x) = 0$ для всех остальных элементов. Пусть, не умаляя общности, $N = 2^n$, $n \geq 0$. Очевидно, что при «лобовом» переборе потребуется не более, чем 2^n операций для достоверного решения задачи или 2^{n-1} операций нахождения решения с вероятностью $\frac{1}{2}$. АГ позволяет, при использовании ККР, сократить этот перебор до $2^{n/2}$ операций, что в некоторых случаях является весьма существенным.

Обычно в качестве простейшего примера постановки такой задачи предлагается по номеру телефона найти фамилию абонента или в огромной базе неупорядоченных фотографий найти по заданной фотографии фамилию, и дополнительно ID, искомого лица.

В случае криптоанализа задача формулируется так: пусть имеется пара – сообщение x и криптограмма y , полученная с использованием неизвестного ключа K – то есть $y = f(x, K)$, где $f(\cdot)$ – алгоритм (функция) шифрования; требуется найти секретный ключ K , используя АГ. Полагается, что такой ключ должен быть единственным, что, в свою очередь, обеспечивается, если длина криптограммы оказывается не менее чем так называемое *расстояние единственности*.

Заметим, что АГ, открытый его автором в 1996 г., описан на сегодня, наверно, в нескольких десятках превосходных монографий, статей и лекций, которые можно свободно найти в интернете [15–19].

Однако для специалистов в области ИТ (включая и информационную безопасность), которые не знакомы профессионально с квантовой физикой, достаточно трудно сознательно воспринять факт возможности существенного сокращения количества вычислений в рассмотренной выше задаче «лобового» поиска ключа до $2^{n/2}$ по сравнению с традиционным методом тотального опробования всех ключей, требующим 2^n вычислений. Поэтому поясним это замечательное свойство ККР, проявляющееся при использовании АГ.

Главное, что требуется для выполнения АГ при решении рассматриваемой задачи – это задание функции шифрования $f(\cdot)$ в виде квантовой схемы. В этом случае схема криптоанализа может быть представлена, как показано на рисунке 1.

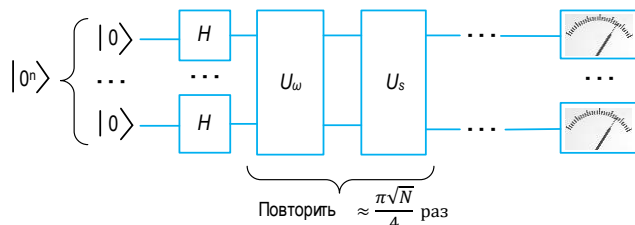


Рис. 1. Квантовая схема отыскания действующих ключей для симметричной криптосистемы при помощи АГ

Fig. 1. Quantum Scheme for Finding Valid Keys for a Symmetric Cryptosystem Using AG

Левый столбец схемы представляет собой так называемый квантовый регистр, состоящий из n кубитов, находящихся в нулевых состояниях. Напомним, что согласно [20], кубит представляет собой квантовое физическое устройство, например, электрон в квантовой точке или в ионной ловушке, которое может находиться в одном из двух базисных состояний. С физической точки зрения – это те и только те состояния системы, в одном из которых она будет обнаружена в результате измерения. В нотациях Дирака – это состояния обозначаются $|0\rangle, |1\rangle$. Первоначально все кубиты устанавливаются в состояние $|0\rangle$. Под действием оператора Адамара H на каждый кубит регистра состояние всего регистра a превращается в «равно-

мерную» линейную комбинацию всех его возможных состояний:

$$H^{\otimes n}|0^n\rangle = \frac{1}{\sqrt{2^n}} \times$$

$$\times (|00 \dots 0\rangle + |00 \dots 1\rangle + \dots + |01 \dots 1\rangle + |11 \dots 1\rangle) = |s\rangle.$$

При решении задач криптоанализа эти базисные состояния описывают все возможные ключи шифрования длины n . Это начальное состояние регистра перед запуском итераций Гровера. При измерении системы, находящейся в таком состоянии, она редуцируется в одно из всех возможных базисных состояний с одной и той же вероятностью $\frac{1}{2^n}$. Такое состояние регистра в точности соответствует состоянию неструктурированной базы данных из $N = 2^n$ элементов.

Оператор U_ω сохраняет без изменения все векторы состояний n -мерного квантового регистра x , кроме состояния $x = \omega$, при котором обеспечивается выполнение равенства $f(x, \omega) = y$, где $f(\cdot, \cdot)$ – функция шифрования, x – сообщение, а y – криптограмма, получаемая при данном сообщении x и ключе ω ; очевидно, что x и ω должны быть предварительно введены в квантовый шифратор.

Для выделенного состояния ω преобразование U_ω изменяет состояние квантового регистра на противоположное, то есть $U_\omega(|x\rangle) = -|x\rangle$.

Итак, на первом шаге итерации Гровера выполняется преобразование $U_\omega|s\rangle$. На втором шаге это состояние преобразуется с помощью так называемого оператора диффузии $U_s = 2|s\rangle\langle s| - I$: $U_s U_\omega|s\rangle$. Здесь I – тождественное преобразование.

Совместное применение операторов U_ω и U_s , называемое *оператором Гровера* $U_G = U_s U_\omega$, повторяется примерно $\frac{\pi}{4} 2^{n/2}$ раз (слово «примерно» означает, что оптимальное значение количества повторений колеблется вокруг данной величины, и оно подбирается экспериментально вокруг указанного значения).

После выполнения требуемого числа итераций производится измерение состояния кубитового регистра и, согласно доказательству, приведенному в работе самого Гровера [21], оно должно с вероятностью, близкой к 1, совпадать с искомым ключом ω . При прочтении описания АГ возникает естественный вопрос – зачем вообще нужны итерации оператора Гровера, если преобразование U_ω , казалось бы, находит ключ ω уже на первой итерации, поскольку выполняется равенство $f(x, \omega) = y$?

Ответ на этот вопрос заключается в том, что правильность измерения ключа после первых итераций имеет почти такую же вероятность, как и для неправильных решений; ситуация сходится к пра-

вильному решению с вероятностью, близкой к 1, лишь после проведения примерно $\frac{\pi}{4} 2^{n/2}$ итераций.

Для того, чтобы пояснить этот «трюк» со сходимостью по вероятности к истинному ключу, рассмотрим упрощенную схему, никак не связанную с КК и называемую *усилением амплитуды*.

Пусть имеется набор из N , например, вещественных чисел $a_1, a_2, \dots, a_N$, аналогов амплитуд квантовых состояний. Для полного сходства положим их равными и нормируем:

$$a_1 = a_2 = \dots = a_N = \frac{1}{\sqrt{N}}.$$

Такой набор чисел соответствует неструктурированной базе данных. Задача состоит в усилении некоторой выделенной амплитуды a_ω . Для этого введем операцию $\hat{U}_\omega$ преобразования числового набора по правилу $\hat{U}_\omega a_x = a_x \forall x \neq \omega$ и $\hat{U}_\omega a_\omega = -a_\omega$. Очевидно, что это преобразование – аналог оператора U_ω . Аналогом оператора диффузии U_S является преобразование числового набора по формуле $\hat{U}_S a_x = \frac{2}{N} \sum_y a_y - a_x$. Нетрудно сообразить, что это отражение значений a_x относительно их среднего $\bar{a} = \frac{1}{N} \sum_y a_y$. Последовательное применение операций $\hat{U}_S \hat{U}_\omega = \hat{U}_G$ действует на числовой набор также, как и оператор Гровера $U_G = U_S U_\omega$. В лекции [22] наглядно продемонстрировано, как происходит усиление выделенной амплитуды, соответствующей вероятности выделенного состояния и уменьшение всех остальных вероятностей.

Для примера выполним на классическом компьютере процедуру усиления «пяти-кубитного» набора чисел ($N = 32$). В качестве выделенной амплитуды возьмем a_2 . Результаты выполнения девяти итераций $\hat{U}_G$ приведены в таблице 1. Первый столбец содержит номер итерации k , второй – среднее значение амплитуд $\bar{a}$ на каждом шаге, третий – вторую амплитуду a_2 .

ТАБЛИЦА 1. Процедура усиления амплитуды ($N = 32$)

TABLE 1. Amplitude Enhancement Procedure ($N = 32$)

k	$\bar{a}$	a_2
0	0,17677	0,17677
1	0,16572	0,50823
2	0,13396	0,77616
3	0,08545	0,94707
4	0,02626	0,99959
5	0,03621	0,92716
6	0,09416	0,73884
7	0,14034	0,45817
8	0,16897	0,12022
9	0,17649	0,23275

Хорошо видно, как вторая амплитуда сходится по модулю к единице при $k \rightarrow 4$ и как деградируют все остальные амплитуды.

Так как с геометрической точки зрения оператор Гровера U_G , как и преобразование $\hat{U}_G$, это оператор поворота вектора состояния квантового регистра в одной гиперплоскости на фиксированный угол, то поведение амплитуд в зависимости от числа итераций носит периодический характер.

Если интерпретировать эти числа как вероятности найти текущий ключ криптосистемы (или как в первом примере – извлечь из памяти КК определенную фотографию), то вероятность нахождения истинного ключа (или искомой фотографии) будет приближаться к единице. Однако, конечно, для выполнения процедуры *усиления амплитуды* совершенно необходимо использовать ККР, где преобразования типа *вычисление среднего* и *инверсия выделенного элемента* должны выполняться не перебором, а одновременно!

Как было отмечено выше, для эффективной при $N = 2^n \gg 1$ реализации алгоритма кубиты нужны не только для «записи» амплитуд состояния квантового регистра, но и для работы квантового шифратора – составной части блока U_ω (рисунок 1), в том числе для хранения открытого текста и криптограммы. Таким образом, число кубит, необходимых для реализации алгоритма, заметно превышает длину n искомого ключа.

В работе [23] достаточно подробно описан взлом при помощи АГ криптосистемы AES с различными, допустимым по стандарту, длинами ключей – 128, 192 и 256. В таблице 2 приводятся результаты расчета требуемого числа кубит для такого квантового криптоанализа, взятые из этой работы.

ТАБЛИЦА 2. Требуемое количество кубитов для выполнения криптоанализа по АГ для стандарта AES с различными длинами ключей

TABLE 2. Number of Qubits Required to Perform AG Cryptanalysis for AES with Different Key Lengths

Длина ключа, n	128	192	256
Количество кубитов	984	1112	1336

В настоящее время лидерами в построении КК являются компании Google и IBM: так, первая еще в 2018 г. реализовала КК с процессором Bristlecone на 72 кубитов [24], вторая же представила в 2023 г. квантовый процессор Heron, имеющий уже 133 кубита и анонсировала создание 1121-кубитного процессора Condor в 2024 г. [25].

Важно также отметить, что с увеличением количества кубитов в КК в них возрастает и вероятность ошибок при выполнении на гейтах различных операций. Поэтому пока весьма проблематично создание КК, содержащих более 100 кубитов,

что подтверждается и авторитетным мнением М. Дьяконова, изложенным им в зарубежной монографии [26].

Кроме того, даже при успешной реализации многобитового КК, из-за нереализуемости слишком большого числа требуемых операций АГ, можно выразить надежду на взлом в будущем, разве лишь, AES-128. Важно также отметить, что в работах, выполненных уже после изобретения АГ, было строго доказано, что не существует другого, более эффективного, чем АГ, метода, требующего менее $2^{n/2}$ итераций при выполнении поиска заданного элемента на неструктурированном множестве.

Заметим, что при попытке понимания эффективного использования АГ для «лобового» криптоанализа, наиболее сложным представляется интуитивно согласиться с возможностью неперекрестного выполнения операции U_ω , то есть нахождения элемента « ω », обеспечивающего выполнение равенства $f(x, \omega) = y$. Интересно вспомнить, что один из профессоров, специализирующихся именно в квантовой физике, как-то сказал своему коллеге и тоже физики, что понять этот квантовый параллелизм невозможно, но, тем не менее, он работает. Можно привести один пример, который, хотя и не эквивалентен квантовой задаче, но, все же, немного поясняет невозможность полной замены физики математикой. Пусть имеется множество деревянных палочек различного, но близкого веса и похожих друг на друга. Лишь одна из этой кучи палочек сделана из железного дерева (с плотностью древесины $\geq 1000 \text{ кг/м}^3$), которое тонет в пресной воде. Требуется отыскать эту палочку. Простейший метод – это поочередно опускать палочки в стакан и найти ту, что тонет. Однако более быстрый метод состоит в том, чтобы вывалить кучу этих палочек в ванну и взять ту, которая пошла на дно.

Еще более впечатляющие результаты дает использование КК при взломе криптосистем с открытым ключом (РША, Эль-Гамала, Рабина и др.), где оно сводится к решению таких трудных задач теории чисел, как их факторизация и дискретное логарифмирование. Причем, если на обычном ПК их решение оказывается экспоненциально сложным, то КК решает эти задачи с полиномиальной зависимостью от длины ключа.

В работе П. Шора [27] приводится описание алгоритмов и оценка требуемого числа элементарных вычислений для решения этих задач. Поскольку описание данных алгоритмов оказывается достаточно сложным и требует знания, как минимум, основ квантовой физики и теории чисел, приведем здесь только оценки сложности вычислений, необходимых для взлома криптосистемы РША, представленные И. Голдовским в [28]:

- число операций для факторизации числа $N = O(n^2 \log_2 n (\log_2 \log_2 n))$, где $n = \log_2 N$;
- требуемое количество кубит КК – $O(n^3 \log_2 n)$.

Из этих границ следует, что для взлома шифра РША придется при использовании ключей длиной $N = 2048$ обеспечить количество кубит, равное 4099. Для криптосистем на основе использования эллиптических кривых (требующих более короткий ключ) потребуется 1500 кубит для ключа длиной $n = 256$ и 2500 – для ключа длиной $n = 512$.

Так что в настоящее время можно говорить о реализуемости взлома криптосистемы РША с длиной ключа не более 1024 бита. В настоящее время в зарубежной криптографии появилась даже аббревиатура SNDL (*от англ. Store Now-Decrypt Later*) в надежде на прогресс в будущем квантовых компьютеров.

В связи с опасностью взлома криптосистем с открытым ключом при помощи КК появились так называемые *постквантовые* криптосистемы. К ним относится криптосистема Мак-Элис, основанная на экспоненциальной сложности от длины ключа алгоритма исправления ошибок для случайных линейных кодов, а также криптосистемы на основе использования *числовых решеток*.

3. АУТЕНТИФИКАЦИЯ МЕТОДА РАСПРЕДЕЛЕНИЯ КЛЮЧЕЙ ПО ПРОТОКОЛУ ДИФФИ – ХЕЛЛМАНА НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ ТЕХНОЛОГИИ «СПАРИВАНИЯ» МОБИЛЬНЫХ УСТРОЙСТВ

3.1. Принцип формирования симметричного ключа по методу Диффи – Хеллмана

Метод Диффи – Хеллмана [29] позволяет использовать незащищенный от перехвата канал связи для формирования общего ключа шифрования с целью создания безопасного соединения между двумя корреспондентами. Этот метод нашел широкое применение в сетевых протоколах SSL/TLS, IPsec, PGP и др. [9]. Изложим кратко суть метода. Пользователи сети Алиса (*сокр. А*) и Боб (*сокр. В*) согласовывают параметры: p – простое число; g – элемент конечного поля $GF(p)$, порождающий группу, имеющий большой порядок, и выполняют следующий пошаговый протокол.

Шаг 1. Алиса генерирует элемент поля $x \in \{1, p-1\}$, вычисляет $X = g^x \pmod{p}$ и посылает его Бобу.

Шаг 2. Боб генерирует элемент поля $y \in \{1, p-1\}$, вычисляет $Y = g^y \pmod{p}$ и посылает его Алисе.

Шаг 3. Алиса вычисляет ключ $K_A = Y^x \pmod{p}$.

Шаг 4. Боб вычисляет ключ $K_B = X^y \pmod{p}$.

Несложно заметить, что ключи $K_A = K_B$. Величины X и Y , которыми обмениваются A и B по от-

крытым каналам, называются значениями *Диффи – Хеллмана* (сокр. ДН-значениями).

Однако данный метод подвержен атаке, известной как «человек посередине». Алиса и Боб при выполнении протокола могут оказаться в ситуации, когда они устанавливают связь со злоумышленником – Евой, которая Алисе выдает себя за Боба, а Бобу представляется Алисой. Поэтому Алисе и Бобу требуется аутентифицировать ключи, сгенерированные по методу Диффи – Хеллмана.

Одним из направлений решения задачи аутентификации ключа является сертификация ДН-значений удостоверяющим центром, что делается, например, в протоколе SSL/TLS. Проверка подлинности сертификата осуществляется с использованием открытого ключа, распространяемого в сети. Однако такой подход требует использования технологии PKI (аббр. от англ. Public Key Infrastructure, инфраструктура открытых ключей), что не всегда удобно для пользователей, особенно для пользователей мобильных сетей.

Другим подходом к решению задачи аутентификации ключа, распределяемого по методу Диффи – Хеллмана, является использование в целях аутентификации технологии *близкой аутентификации* (называемой иначе *спаррингом*). Для этого пользователи *A* и *B* предварительно выполняют процедуру сопряжения своих мобильных устройств во время личной встречи. В результате они формируют двоичные последовательности *a* и *b*, соответственно, которые впоследствии используют для взаимной аутентификации, в том числе и для аутентификации ДН-значений. Предполагается, что нарушитель в момент сопряжения устройств удален от пользователей и не получает доступа к сообщениям, которыми обмениваются пользователи. Специфика решения задачи аутентификации этим способом состоит в том, что пользователи не могут непосредственно использовать вырабатываемые последовательности ни для шифрования, ни для аутентификации, так как эти последовательности содержат определенный процент ошибок (то есть несовпадений бит).

Другой способ выработки аутентифицирующих последовательностей может заключаться в использовании *квантовой телепортации*.

Ниже приводятся результаты исследования аутентификации ДН-значений с использованием аутентифицирующих последовательностей независимо от способа их получения.

3.2. Сценарий спаривания мобильных устройств и аутентификации ДН-значений

Задача получения пользователями идентичных последовательностей может быть решена на основе создания между пользователями дополнитель-

ного канала: визуального, акустического, вибрационного, тактильного, магнитометрического или даже канала квантовой телепортации.

Дополнительный канал образуется между двумя мобильными устройствами при личной встрече пользователей, при этом не требуется передача какой-либо информации по каналу связи, что затрудняет проведение атак со стороны злоумышленника. Для краткости будем называть обмен данными по дополнительному каналу с целью выработки аутентифицирующих последовательностей *близкой аутентификацией* на основе *сопряжения мобильных устройств*.

Поскольку участие пользователей при сопряжении неизбежно, решающими факторами в выборе способа сопряжения являются удобство использования, помехоустойчивость и минимальный объем передаваемых данных; сравнительная характеристика дополнительных каналов по этим критериям приведена в работах [30, 31]. На основе этих данных в наибольшей степени для решения задачи аутентификации подходят вибрационный и магнитометрический каналы.

Вибрационный канал. Его использование представляется возможным для устройств, содержащих в себе датчики акселерометра. Два содержащих их мобильных устройства необходимо встряхивать в одной руке примерно в течение 5 с. В это время осуществляется считывание информации о положении мобильного устройства в пространстве и преобразование ее в цифровой код. Два устройства (*A* и *B*), которые трясли вместе, на выходе получают схожие последовательности.

Магнитометрический канал. Пользователям необходимо удерживать два устройства вблизи друг к другу несколько секунд без выполнения каких-либо дополнительных операций. Устройства считывают собственные показания датчиков магнитометра и обмениваются ими. Сопряжение мобильных устройств с использованием магнитометрического канала, по сравнению с вибрационным, обеспечивает более высокую скорость работы (4,5 с), низкую вероятность ошибочной аутентификации и позволяет минимизировать участие пользователя в процессе сопряжения устройств, что указывает на преимущество магнитометрического канала для выработки случайных последовательностей [30].

Применение магнитометрического канала с целью выработки последовательностей, их использования для построения аутентифицирующих помехоустойчивых кодов (сокр. АП-кодов) и аутентификации ключей Диффи – Хеллмана достаточно полно было исследовано в работе [31].

В [32] был рассмотрен способ аутентификации ДН-значений на основе использования универсальных хэш-функций, построенных путем ис-

пользования заранее выработанных пользователями аутентифицирующих последовательностей. В работе [33] аутентификация ДН-значений осуществлялась на основе алгоритма Картера – Вермана с одноразовым ключом, в качестве которого использовались аутентифицирующие последовательности, выработанные пользователями во время процедуры спаринга.

Перспективным способом получения аутентифицирующих последовательностей является, на наш взгляд, применение современных квантовых технологий и, в частности, технологии квантовой телепортации, которая интенсивно развивается в последние годы на основе использования теории ЭПР (аббр. от Эйнштейн, Подольский, Розен – авторы теории) [34]. Суть этого подхода состоит в осуществлении следующей процедуры.

Во время сопряжения мобильных устройств пользователи, используя специальное устройство, вырабатывают на ЭПР-пару «запутанных» кубит. Один кубит пары записывает в свою квантовую память пользователь А, а другой кубит – пользователь В. Во время процедуры телепортации, используя эти «запутанные» кубиты и дополнительный канал обмена (не квантовый), пользователи формируют аутентифицирующие последовательности. Перспективность данного подхода нам видится в том, что получение пар кубитов может быть выполнено пользователями на большом расстоянии между ними, и в этом случае необходимость в близкой аутентификации вообще отпадает. В настоящее время достигнуты дальности в процедуре телепортации: в свободном пространстве 1200 км [35], а по оптоволоконному кабелю – 44 км [36]. Однако нужно отметить, что полученные таким образом последовательности полностью не совпадают друг с другом. Точность совпадения квантовых состояний, переданных по волоконно-оптическому кабелю ~90 %. Конечно, многие вопросы телепортации, в частности хранение бит «запутанных» пар, пока окончательно не решены.

Рассмотрим более подробно сценарий сопряжения мобильных устройств, для формирования случайных последовательностей и аутентификации ДН-значений с использованием магнитометрического канала, и смартфонов. Сценарий включает четыре этапа (рисунок 2).

1-й этап – этап сопряжения. Алиса и Боб располагают устройства достаточно близко друг к другу (практически касаются друг друга) и удерживают несколько секунд. В это время в каждом смартфоне производится измерение магнитного поля Земли по трем пространственным осям. Значения магнитного поля квантуются и из них формируются двоичные последовательности.

В процессе измерений поля между устройства-

ми осуществляется локальный обмен данными, например, по каналу Bluetooth, для коррекции измерительных базисов устройств с целью обеспечения их идентичности. В итоге каждое устройство формирует случайную двоичную последовательность **a** или **b** необходимой длины. Эти последовательности пользователи записывают в память своих устройств и будут использовать по мере необходимости как ключи в процедуре аутентификации ДН-значений. Обозначим $p_m = P(a_i \neq b_i)$, где $i = 1, 2 \dots L$ – номер символа в последовательностях, p_m – вероятность несовпадения бит в последовательностях **a** и **b**.

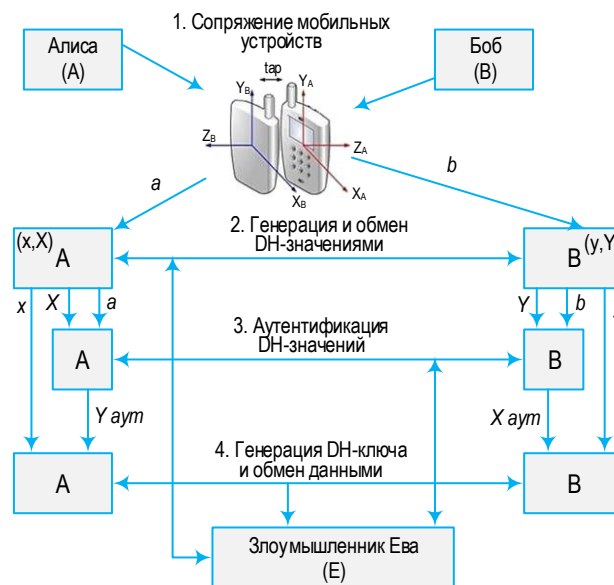


Рис. 2. Сценарий спаривания мобильных устройств и формирования ключа Диффи – Хеллмана с его аутентификацией

Fig. 2. Scenario of Pairing Mobile Devices and Formation of Diffie – Hellman Key with Its Authentication

В работе [30] подробно описан процесс сбора и коррекции данных магнитометров, в частности, предложен протокол Magpairing для создания корреспондентами А и В случайных последовательностей на основе магнитометрического канала и их использования для аутентификации ДН-значений. Нами были проведены аналогичные исследования [37], которые подтвердили результаты этой работы; вероятность совпадения символов последовательностей, полученных от двух устройств в наших экспериментах составила 0,94.

Предполагается, что нарушитель (посторонний пользователь) также имеет магнитометр и может быть расположен достаточно близко к сопрягаемым устройствам, но не в точке сопряжения, поэтому сформировать последовательность, совпадающую с большей вероятностью с **a** и **b** он не может. Будем полагать, что вероятность несовпадения символов в последовательности, которую может сформировать нарушитель, с символами по-

следовательности легальных пользователей равна $p_e = P(a_i \neq e_i) = 1/2, i = 1, 2 \dots L$, что соответствует обрыву канала.

2-й этап. Алиса и Боб вырабатывают секретные числа x и y , соответственно, вычисляют ДН-значения и обмениваются ими, используя канал связи между A и B , контролируемый нарушителем, который может выполнять пассивные и активные атаки на передаваемую информацию.

3-й этап – аутентификация переданных по каналу ДН-значений на основе использования последовательностей a и b в качестве ключа аутентификации для формирования и проверки аутентификаторов, передаваемых от A к B и обратно.

4-й этап. Если аутентификация ДН-значений в обе стороны прошла успешно, то Алиса и Боб, используя имеющиеся у них случайные числа x и y , соответственно, вычисляют ключи по методу Диффи – Хеллмана.

Однако, как показано в [31], способ с использованием магнитометрического канала подвержен атаке «человек посередине» (см. п. 3.1); там также предложен способ устранения этой уязвимости на основе использования аутентифицирующих помехоустойчивых кодов. Дальнейшее развитие этот способ получил в работах [32, 33], где были предложены способы применения аутентифицирующих последовательностей a и b для формирования аутентификаторов по алгоритму Вегмана – Картера с использованием универсальных хэш-функций (WCA UH, аббр. от англ. Wegman-Carter Algorithm Universal Hash Function) и/или с одноразовым ключом (WCA OTP, аббр. от англ. Wegman-Carter Algorithm One Time Password) [38].

Способ аутентификации на основе алгоритмов WCA UH и WCA OTP схожи и отличаются эффективностью в плане расхода аутентифицирующих последовательностей a и b , которые выработали корреспонденты на этапе близкой аутентификации. В то же время WCA UH представляет основу для обоих алгоритмов, поэтому рассмотрим далее детально способ аутентификации ДН-значений на основе универсальных хэш-функций.

3.3. Аутентификация ДН-значений на основе предварительно распределенных случайных последовательностей и алгоритма WCA UH

Рассмотрим этап аутентификации ДН-значений согласно сценарию, показанному на рисунке 1.

Пользователь A разделяет ДН-значение X на N блоков длиной m бит. Для каждого блока u_i вычисляется аутентификатор w_i длиной v бит ($i = 1, 2, \dots, N$) на основе универсального класса хэш-функций и использования последовательности a в качестве ключа. Для аутентификации каждого следующего блока выбираются новый блок в по-

следовательности a .

Корреспондент B проводит аутентификацию ДН-значения следующим образом. Для каждого принятого блока u_i вычисляется аутентификатор w_i' с использованием последовательности b , который сравнивается с аутентификатором w_i , полученным по каналу связи. Если $w = w_i'$, то блок u_i аутентифицирован, и не аутентифицирован в противном случае. ДН-значение считается аутентифицированным в целом, если среди N принятых блоков окажутся не аутентифицированными не более Δ блоков $1 \leq \Delta \leq N$.

Аутентификатор для каждого блока вычисляется согласно алгоритму WCA UH. Представим подблок u , как элемент поля Галуа $GF(2^m)$. Тогда аутентификатор $w = [u \times k_0 + k_1]_v$, где $k_0, k_1 \in GF(2^m)$ – ключи аутентификации. Знаки $\times, +$ обозначают, соответственно, умножение и сложение в конечном поле $GF(2^m)$, а $[\cdot]_v$ – «усечение», то есть выбор v левых или правых элементов последовательности в квадратных скобках. Блоки длиной m бит выбираются поочередно из последовательности a . Для каждого очередного блока u_i формируется новый аутентификатор w_i с использованием новой пары ключей k_0 и k_1 . Вероятность навязывания ложного блока при известных u и w для данного способа аутентификации определяется соотношением $P_s = 1/2^v$ [38].

Длина ключа аутентификации (минимальная длина последовательности a), необходимая для аутентификации ДН-значения: $L = 2mN$.

Будем считать, что нарушитель Ева может осуществить следующие атаки.

Атака подмены. Нарушитель перехватывает ДН-значение $X = g^x$ длиной n_0 бит, передаваемое в виде подблоков u_i и аутентификаторов $w_i, i = 1, 2, \dots, N$. Генерирует ложное сообщение $X' = g^{x'}$, отличающееся от исходного в D блоках и формирует аутентификаторы к нему по следующему правилу:

- если блоки u_i в исходном и ложном сообщении совпали, он использует перехваченные аутентификаторы;
- если блоки не совпали, нарушитель формирует аутентификаторы случайным образом.

Атака имперсонализации. Ложное ДН-значение $X' = g^{x'}$ создается нарушителем без предварительного приема истинного ДН-значения от корреспондента A .

Оценку эффективности аутентификации ДН-значения будем осуществлять по следующим параметрам:

P_f – вероятность ложного отклонения ДН-значения в отсутствие навязывания; событие наступает, когда число неправильно аутентифицированных блоков равно Δ и более из-за несогласованности аутентифицирующих последовательностей a и b ;

P_i – вероятность имперсонализации; наступает, когда нарушитель создает ложное ДН-значение (без предварительного приема истинного ДН-значения), которое принимается как истинное;

P_s – вероятность подмены ДН-значения; наступает, когда нарушитель создает ложное ДН-значение, отличающегося от истинного в D блоках размерности m (обозначим вероятность этого события – $P_g(D)$), и навязывает его корреспонденту (вероятность обмана – $P_r(D)$); поскольку величина D определяется злоумышленником, то в худшем случае:

$$P_s = \max_D (P_g(D) \times P_r(D)); \quad (3)$$

P_d – вероятность навязывания ложного ДН-значения:

$$P_d = \max(P_i, P_s); \quad (4)$$

L – длина ключа аутентификации (длина последовательностей **a** и **b**), необходимая для аутентификации ДН-значения длиной n_0 с заданными значениями $\tilde{P}_f, \tilde{P}_d$;

$W = vN$ – суммарная длина всех аутентификаторов ДН-значения.

Вероятность P_f может быть оценена, как вероятность суммы событий, состоящих в том, что в принятой последовательности X из-за ошибок в аутентифицирующих последовательностях окажется не аутентифицированными $\Delta + 1$ и более блоков:

$$P_f(\Delta) = \sum_{i=\Delta+1}^N C_N^i p_b^i \times (1 - p_b)^{N-i}, \quad (5)$$

где p_b – вероятность несовпадения ключей (k_0, k_1) равна вероятности несовпадения блоков длиной $2m$ бит, выбранных их последовательностей **a** и **b** и равняется $p_b = 1 - (1 - p_m)^{2m}$ в предположении, что ошибки в битах распределены по закону Бернулли.

Вероятность успешной атаки имперсонализации можно получить, как оценку количества обнаруженных и не обнаруженных ложных аутентификаторов в N блоках при их случайном генерировании. Учитывая, что вероятность успешного навязывания ложного аутентификатора длиной v символов при использовании универсальных хэш-функций равна $1/2^v$, можно записать:

$$P_f(\Delta) = \sum_{i=0}^{\Delta} C_N^i \left(\frac{1}{2^v}\right)^{(N-i)} \times \left(1 - \frac{1}{2^v}\right)^i. \quad (6)$$

Для оценки вероятности подмены P_s рассмотрим сомножители в (3).

Очевидно, что чем меньше отличаются истинное ДН-значение X и ложное X' , тем легче наруши-

телю реализовать атаку подмены. Заметим, что непосредственно значение X' нарушитель выбрать не может. Сначала он выбирает число x' , затем находит значение экспоненты $X' = g^{x'} \bmod p$. Если x' выбирать случайно из множества чисел от 0 до $p-1$, то число X' будет также случайным числом из множества чисел от 0 до $p-1$. При большом значении модуля $p \approx 2^{256}$ нахождение отображения $x' \Leftrightarrow X'$ требует необозримо больших вычислительных затрат. Поэтому разумной стратегией для нарушителя будет случайный выбор числа x' , что равносильно случайному выбору X' .

Вероятность случайного формирования нарушителем ложного ДН-значения X' , отличающегося от перехваченного значения X в D блоках, может быть найдена на основе комбинаторных рассуждений о выпадении одинаковых значений двух 2^m -гранных игральные костей при N бросаниях и записана как:

$$P_g(D) = C_N^D \left(\frac{1}{2^m}\right)^{N-D} \times \left(1 - \frac{1}{2^m}\right)^D. \quad (8)$$

Вероятность навязывания сформированного ложного ДН-значения, отличающегося от истинного в D блоках, в зависимости от порога принятия решения Δ , можно оценить как:

$$P_r(\Delta) \leq \sum_{i=0}^{\Delta} C_D^i \left(\frac{1}{2^v}\right)^{(D-i)} \times \left(1 - \frac{1}{2^v}\right)^i \sum_{j=0}^t C_{(N-D)}^j p_b^j \times (1 - p_b)^{(N-D-j)}, \quad (9)$$

где $t = \begin{cases} N - D, & \text{если } \Delta - i \geq N - D \\ \Delta - i, & \text{если } \Delta - i < N - D \end{cases}$

Первая сумма в (9) характеризует вероятность фиксации на приеме i ложных аутентификаторов ($0 \leq i \leq \Delta$), которые обнаруживаются, и $D-i$ ложных аутентификаторов, которые не обнаруживаются. Вторая сумма – это вероятность несовпадения j ($0 \leq j \leq t$) из $N-D$ аутентификаторов исходного сообщения, которые нарушителем были переданы без изменений, а несовпадение произошло за счет несоответствия ключей в аутентифицирующих последовательностях **a** и **b**.

Рассмотрим пример выбора параметров системы аутентификации, реализующей описанный выше способ. Пусть ДН-значение, которое необходимо аутентифицировать, имеет длину 256 бит, вероятность несовпадения аутентифицирующих последовательностей $p_m = 0,05$. К системе аутентификации предъявлены требования: $\tilde{P}_f = \tilde{P}_d = 10^{-6}$.

Выберем несколько длин аутентифицируемых блоков $m = 1, 2, 4, 8, 16, 32$, среди которых будем искать удовлетворяющие заданным требованиям.

Используя соотношение (6), построим зависимость $P_f(\Delta)$ для выбранных длин блока аутентификации (рисунок 3). На пересечении кривых $P_f(\Delta, m)$ и прямой $P_f = 10^{-6}$ находим значения

порогов ($\Delta_{\min} = 50, 46, 40, 30$ для $m = 1, 2, 4, 8$, соответственно), при которых выполняется неравенство $P_f \leq \tilde{P}_f$. Для $m = 16$ и 32 такие условия не выполняются, поэтому для дальнейшего исследования оставляем блоки с длинами $1, 2, 4, 8$.

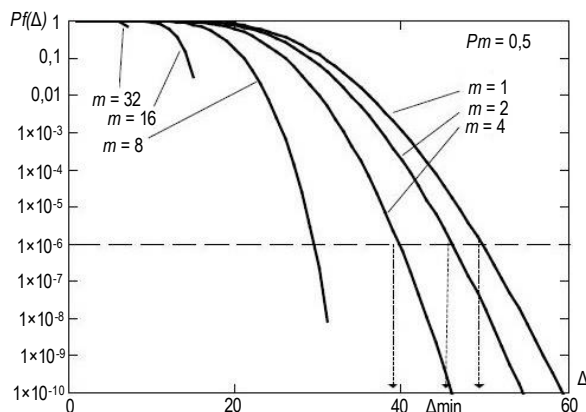


Рис. 3. Зависимость вероятности ложного отклонения ДН-значения от порога Δ для разных длин x_0 аутентифицируемых блоков

Fig. 3. False Rejection Probability of the DH Value Dependence on the Threshold Δ for Different Lengths x_0 of Authenticated Blocks

Используя найденные значения $\Delta_{\min}$, построим зависимости $P_s(D) = P_g(D) \times P_r(D)$ для выбранных блоков длины m и всех аутентификаторов длины $v \leq m$. Зависимости $P_g(D)$ и $P_r(D)$ рассчитываются согласно (8) и (9) (рисунок 4). По полученным зависимостям находим значения $P_s = \max_D P_s(D)$.

Максимумы кривых на рисунке 4 соответствуют вероятностям подмены в соответствии с соотношением (3) для разных значений длин аутентифицируемого блока (m) и аутентификатора (v). Вероятности ложного отклонения, подмены, имперсонализации, навязывания рассчитанные для этих значений представлены в таблице 3. Видим, что условию $\tilde{P}_f = \tilde{P}_d = 10^{-6}$ удовлетворяют только пары (m, v): (2,2) и (4,4).

В данном примере значение вероятности подмены, удовлетворяющее требованию $P_f \leq 10^{-6}$, выполняется только при $m = v$, поэтому суммарная длина всех аутентификаторов равна длине ДН-значения и равна $W = 256$ бит, а требуемая для

аутентификации длина последовательностей a и b (длина ключа аутентификации) равна $L = 512$ бит.

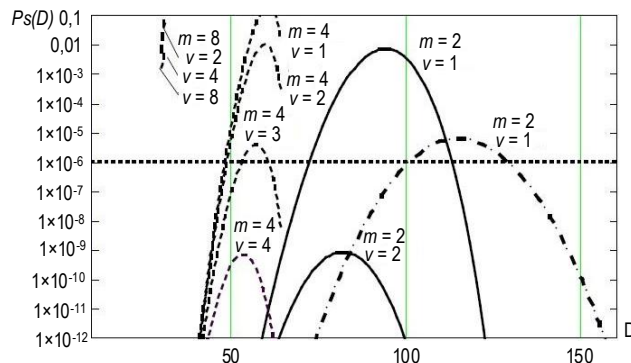


Рис. 4. Зависимость вероятности подмены от количества блоков, в которых различаются истинное и ложное ДН-значения (D) для отобранных пар (m, v)

Fig. 4. Dependence of Substitution Probability on the Blocks Number in Which True and False DH-Values (D) Differ for the Selected Pairs (m, v)

ТАБЛИЦА 3. Значения вероятности ложного отклонения (P_f), подмены (P_s), имперсонализации (P_i) и навязывания (P_d) для пар (m, v)

TABLE 3. Values of the Probability of False Rejection (P_f), Substitution (P_s), Impersonation (P_i) and Imposition (P_d) for Pairs (m, v)

(m, v)	P_f	P_s	P_i	$P_d = \max(P_s', P_i)$
(1,1)	$7,1 \times 10^{-7}$ (+)	$6,2 \times 10^{-6}$ (-)	$5,7 \times 10^{-24}$ (+)	$6,2 \times 10^{-6}$ (-)
(2,2)	1×10^{-6} (+)	$8,1 \times 10^{-10}$ (+)	$1,4 \times 10^{-20}$ (+)	$8,1 \times 10^{-10}$ (+)
(2,1)	1×10^{-6} (+)	$7,4 \times 10^{-3}$ (-)	$9,3 \times 10^{-4}$ (-)	$7,4 \times 10^{-3}$ (-)
(4,4)	$6,5 \times 10^{-7}$ (+)	$6,8 \times 10^{-10}$ (+)	$2,7 \times 10^{-13}$ (+)	$6,8 \times 10^{-10}$ (+)
(4,3)	$6,5 \times 10^{-7}$ (+)	$3,8 \times 10^{-6}$ (-)	$3,3 \times 10^{-7}$ (+)	$3,8 \times 10^{-6}$ (-)
(4,2)	$6,5 \times 10^{-7}$ (+)	1×10^{-2} (-)	$1,8 \times 10^{-2}$ (-)	1×10^{-2} (-)
(4,1)	$6,5 \times 10^{-7}$ (+)	$1,4 \times 10^{-1}$ (-)	$9,8 \times 10^{-1}$ (-)	$1,4 \times 10^{-1}$ (-)

Таким образом, приведенное выше исследование показало, что, используя процедуру аутентификации спариванием на основе магнитометрического или вибрационного каналов, содержащих небольшой процент ошибок $\approx 5\%$, можно достаточно надежно осуществить аутентификацию секретного ключа, распределяемого между мобильными устройствами по алгоритму Диффи – Хеллмана.

Продолжение следует...

Список источников

- ГОСТ 28147-89 Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования.
- ГОСТ Р 34.12-2015 Информационная технология. Криптографическая защита информации. Блочные шифры. М.: Стандартинформ, 2015.
- ГОСТ Р 34.10-2012 Информационная технология. Криптографическая защита информации. Процессы формирования и проверки электронной цифровой подписи. М.: Стандартинформ, 2012.
- ГОСТ Р 34.11-2012 Информационная технология. Криптографическая защита информации. Функция хэширования. М.: Стандартинформ, 2012.
- Бабаш А.В., Шанкин Г.П. Криптография. М.: Изд-во Солон-Р, 2002. 511 с.
- Korjik V.I., Mukherjee A., Ereemeev M.A., Moldovyan N.A. Fault-based analysis of flexible ciphers // Computer Science Journal of Moldova. 2002. Vol. 10. Iss. 2. PP. 223–236.

7. Korzhik V., Yakovlev V., Kovajkin Yu., Morales-Luna G. Secret Key Agreement Over Multipath Channels Exploiting a Variable-Directional Antenna // International Journal of Advanced Computer Science and Applications. 2012. Vol. 3. Iss. 1. PP. 172–178. DOI:10.14569/IJACSA.2012.030127
8. Коржик В.И., Яковлев В.А. Основы криптографии: учебное пособие. СПб.: Издательский центр Интермедиа, 2016. 312 с. EDN:WEQWMN
9. Коржик В.И., Кушнир Д.В. Основы защиты информации в компьютерных системах. Методические указания к лабораторным работам. Часть 1. СПб: СПбГУТ, 1999. 17 с.
10. Alpern B., Schneider F.B. Key exchange using 'keyless cryptography' // Information Processing Letters. 1983. Vol. 16. Iss. 2. PP. 79–81. DOI:10.1016/0020-0190(83)90029-7
11. Korzhik V., Starostin V., Yakovlev V., Kabardov M., Krasov A. and Adadurov S. Advance in Keyless Cryptography // In: Ramakrishnan S., ed. Lightweight Cryptographic Techniques and Cybersecurity Approaches. IntechOpen, 2022. DOI:10.5772/intechopen.104429
12. Shannon C.E. Communication theory of secrecy systems // Bell System Technical Journal. 1949. Vol. 28. Iss. 4. PP. 656–715. DOI:10.1002/j.1538-7305.1949.tb00928.x
13. Vernam G.S. Secret signaling System. Patent US, no. 1310719A, 22.07.1919.
14. Лавриков И.В., Шишкин В.А. Какой объем данных можно безопасно обрабатывать на одном ключе в разных режимах? // Математические вопросы криптографии. 2019. Т. 10. № 2. С. 125–134 DOI:10.4213/mvk290
15. Williams C.P. Explorations in Quantum Computing. Texts in Computer Science. 2011. DOI:10.1007/978-1-84628-887-6.
16. Алгоритм Гровера. URL: <http://www.youtube.com/watch?v=cQDpimNzKMo> (дата обращения 10.05.2024)
17. Денисенко Д.В., Никитенкова М.В. Применение Квантового Алгоритма Гровера в задаче поиска ключа блочно-го шифра SDES // Журнал экспериментальной и теоретической физики. 2019. Т. 155. № 1. С. 32–53. DOI:10.1134/S0044451019010036. EDN:VRNRNG
18. Нильсен М., Чанг И. Квантовые вычисления и квантовая информация. Пер. с англ. М.: Мир, 2006. 824 с.
19. Кайе Ф., Лафлам Р., Моска М. Введение в квантовые вычисления. М. – Ижевск: НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2012. 360 с.
20. Nielsen M.A., Chuang I.L. Quantum Computation and Quantum Information // Contemporary Physics. 2011. Vol. 52. Iss. 6. PP. 604–605 DOI:10.1080/00107514.2011.587535
21. Grover L.K. A fast quantum mechanical algorithm for database search // Proceedings of the twenty-eighth annual ACM symposium on Theory of Computing (STOC'96, Philadelphia, USA, 22–24 May 1996). New York: ACM, 1996. DOI:10.1145/237814.237866
22. Запругаев А. Квантовые компьютеры. URL: <http://www.youtube.com/watch> (дата обращения 10.05.2024)
23. Grassl M., Langenberg B., Roetteler M., Steinwand R. Applying Grover's algorithm to AES: quantum resource estimates // arXiv:1512.04965v1 [quantum-ph]. 2015. DOI:10.48550/arXiv.1512.04965
24. A Preview of Bristlecone, Google's New Quantum Processor // Google Research. 2018. URL: <https://research.google/blog/a-preview-of-bristlecone-googles-new-quantum-processor> (Accessed 10.05.2024)
25. IBM представила свой мощнейший квантовый процессор Heron и первый модульный квантовый компьютер // 3Dnews. 2023. URL: <https://3dnews.ru/1096936/ibm-predstavila-133kubitniy-kvantoviy-protsessor-heron-i-perviy-modulniy-kvantoviy-kompyuter> (дата обращения 10.05.2024)
26. Dyakonov M.I. Is Fault-Tolerant Quantum Computation Really Possible? In: Future Trends in Microelectronics // arXiv:quant-ph/0610117v1. 2006. DOI:10.48550/arXiv.quant-ph/0610117
27. Shor P.W., Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer // SIAM Journal on Computing. 1997. Vol. 26. Iss. 5. PP. 1484–1509. DOI:10.1137/S0097539795293172
28. Голдовский И. Постквантовая криптография. Готовимся сегодня? // ПЛАС. 2022. № 2(288). URL: <https://plusworld.ru/journal/2022/plus-2-2022/postkvantovaya-kriptografiya-gotovimsya-segodnya> (дата обращения 10.05.2024)
29. Diffe M., Hellman M. New directions in cryptography // IEEE Transactions on Information Theory. 1976. Vol. 22. Iss. 6. PP. 644–654. DOI:10.1109/TIT.1976.1055638
30. Jin R., Shi L., Zeng K., Pande A., Mohapatra P. MagPairing: Pairing Smartphones in Close Proximity Using Magnetometer // IEEE Transactions on Information Forensics and Security. 2016. Vol. 1. Iss. 6. PP. 1304–1319. DOI:10.1109/TIFS.2015.2505626
31. Яковлев В.А. Аутентификация ключей, распределяемых методом Диффи – Хеллмана, для мобильных устройств на основе аутентифицирующих помехоустойчивых кодов и магнитометрических данных // Труды СПИИРАН. 2019. Т. 18. № 3. С. 705–740. DOI:10.15622/sp.2019.18.3.705-740. EDN:PRNILE
32. Yakovlev V., Korzhik V., Adadurov S. Authentication of Diffie-Hellman Protocol for Mobile Units Executing a Secure Device Pairing Procedure in Advance // Proceedings of the 29th Conference of Open Innovations Association (FRUCT, Tampere, Finland, 12–14 May 2021). 2021. PP. 385–392. EDN:DWHDGP
33. Яковлев В.А. Способ аутентификации значений Диффи – Хеллмана на основе предварительно распределенных случайных последовательностей и алгоритма аутентификации Вермана – Картера с одноразовым ключом. // Труды учебных заведений связи. 2021. Т. 7. № 3. С. 79–90. DOI:10.31854/1813-324X-2021-7-3-79-90. EDN:TBVSMO
34. Нильсен М., Чанг И. Квантовые вычисления и квантовая информация. Пер. с англ. М.: Мир, 2006. 824 с.
35. Castelvetti D. China's quantum satellite clears major hurdle on way to ultrasecure communications // Nature. 2017. DOI:10.1038/nature.2017.22142
36. Nield D. Quantum Teleportation Was Just Achieved With 90% Accuracy Over a 44km Distance // ScienceAlert. 2020. <https://www.sciencealert.com/scientists-achieve-sustained-high-fidelity-quantum-teleportation-over-44-km> (Accessed 10.05.2024)
37. Корпусов В.Д., Ольховой О.О., Яковлев В.А. Исследование датчика случайных чисел на основе магнитометра // VII Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (Санкт-Петербург, Россия, 28 февраля–1 марта 2018). СПб.: СПбГУТ, 2018. С. 488–494. EDN:OVMOXN

38. Wegman M.N., Carter J.L. New Hash Functions and their Use in Authentication and Set Equality // Journal of Computer and System Sciences. 1981. Vol. 22. Iss. 3. PP. 265–279. DOI:10.1016/0022-0000(81)90033-7

References

1. GOST 28147-89 *Information processing systems. Cryptographic protection. Algorithm of cryptographic transformation.* (in Russ.)
2. GOST P 34.12-2015 *Information technology. Cryptographic data security. Block ciphers.* Moscow: Standartinform Publ.; 2015. (in Russ.)
3. GOST P 34.10-2012 *Information technology. Cryptographic data security. Generation and verification processes of electronic digital signature.* Moscow: Standartinform Publ.; 2012. (in Russ.)
4. GOST P 34.11-2012 *Information technology. Cryptographic data security. Hash-function.* Moscow: Standartinform Publ.; 2012. (in Russ.)
5. Babash A.V., Shankin G.P. *Cryptography.* Moscow: Solon-R Publ.; 2002. 511 p. (in Russ.)
6. Korzhik V.I., Mukherjee A., Ereemeev M.A., Moldovyan N.A. Fault-based analysis of flexible ciphers. *Computer Science Journal of Moldova.* 2002;10(2):223–236.
7. Korzhik V., Yakovlev V., Kovajkin Yu., Morales-Luna G. Secret Key Agreement Over Multipath Channels Exploiting a Variable-Directional Antenna. *International Journal of Advanced Computer Science and Applications.* 2012;3(1):172–178. DOI:10.14569/IJACSA.2012.030127
8. Korzhik V.I., Yakovlev V.A. *Fundamentals of Cryptography.* St. Petersburg: Intermedia Publ.; 2016. 312 p. (in Russ.) EDN:WEQWMN
9. Korzhik V.I., Kushnir D.V. *Fundamentals of information protection in computer systems. Methodical instructions for laboratory works. Part 1.* St. Petersburg: The Bonch-Bruевич Saint Petersburg State University of Telecommunications Publ.; 1999. 17 p. (in Russ.)
10. Alpern B., Schneider F.B. Key exchange using 'keyless cryptography'. *Information Processing Letters.* 1983;16(2):79–81. DOI:10.1016/0020-0190(83)90029-7
11. Korzhik V., Starostin V., Yakovlev V., Kabardov M., Krasov A. and Adadurov S. Advance in Keyless Cryptography. In: Ramakrishnan S., ed. *Lightweight Cryptographic Techniques and Cybersecurity Approaches.* IntechOpen; 2022. DOI:10.5772/intechopen.104429
12. Shannon C.E. Communication theory of secrecy systems. *Bell System Technical Journal.* 1949;28(4):656–715. DOI:10.1002/j.1538-7305.1949.tb00928.x
13. Vernam G.S. *Secret signaling System.* Patent US, no. 1310719A, 22.07.1919.
14. Lavrikov L.V., Shishkin V.A. How much data may be safely processed on one key in different modes? *Mathematical Aspects of Cryptography.* 2019;10(2):125–134 (in Russ.) DOI:10.4213/mvk290
15. Williams C.P. *Explorations in Quantum Computing. Texts in Computer Science.* 2011. DOI:10.1007/978-1-84628-887-6.
16. *Grover algorithm.* (in Russ.) URL: <http://www.youtube.com/watch?v=cQDpimNzKMo> [Accessed 10.05.2024]
17. Denisenko D.V., Nikitenkova M.V. Application of Grover's Quantum Algorithm for SDES Key Searching. *Journal of Experimental and Theoretical Physics.* 2019;128(1):25–44. (in Russ.) DOI:10.1134/S1063776118120142. EDN:GGWKYZ
18. Nielsen M.A., Chuang I.L. *Quantum Computation and Quantum Information.* Cambridge Universities Press; 2001.
19. Kaye F., Laflamme R., Mosca M. *Introduction to quantum computing.* Moscow – Izhevsk: Regul'yarnaya i haoticheskaya dinamika Publ.; Institut komp'yuternykh Issledovaniy Publ.; 2012. 360 p. (in Russ.)
20. Nielsen M.A., Chuang I.L. Quantum computation and Quantum Information. *Contemporary Physics.* 2011;52(6):604–605. DOI:10.1080/00107514.2011.587535
21. Grover L.K. A fast quantum mechanical algorithm for database search. *Proceedings of the twenty-eighth annual ACM symposium on Theory of Computing, STOC'96, 22–24 May 1996, Philadelphia, USA.* New York: ACM; 1996. DOI:10.1145/237814.237866
22. Zapryagaev A. *Quantum computers.* (in Russ.) URL: <http://www.youtube.com/watch> [Accessed 10.05.2024]
23. Grassl M., Langenberg B., Roetteler M., Steinwandt R. Applying Grover's algorithm to AES: quantum resource estimates. *arXiv:1512.04965v1 [quantum-ph].* 2015. DOI:10.48550/arXiv.1512.04965
24. *Google Research.* A Preview of Bristlecone, Google's New Quantum Processor. 2018. URL: <https://research.google/blog/a-preview-of-bristlecone-googles-new-quantum-processor> [Accessed 10.05.2024]
25. *3Dnews.* IBM unveiled its most powerful quantum processor, Heron, and the first modular quantum computer. 2023. URL: <https://3dnews.ru/1096936/ibm-predstavila-133kubitniy-kvantoviy-protsessor-heron-i-perviy-modulniy-kvantoviy-kompyuter> [Accessed 10.05.2024]
26. Dyakonov M.I. Is Fault-Tolerant Quantum Computation Really Possible? In: Future Trends in Microelectronics. *arXiv:quant-ph/0610117v1.* 2006. DOI:10.48550/arXiv.quant-ph/0610117
27. Shor P.W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing.* 1997;26(5):1484–1509. DOI:10.1137/S0097539795293172
28. Goldovsky I. Post-quantum cryptography. Are we preparing today? *PLUS.* 2022;2(288). (in Russ.) URL: <https://plusworld.ru/journal/2022/plus-2-2022/postkvantovaya-kriptografiya-gotovimsya-segodnya> [Accessed 10.05.2024]
29. Diffie M., Hellman M. New directions in cryptography. *IEEE Transactions on Information Theory.* 1976;22(6):644–654. DOI:10.1109/TIT.1976.1055638
30. Jin R., Shi L., Zeng K., Pande A., Mohapatra P. MagPairing: Pairing Smartphones in Close Proximity Using Magnetometer. *IEEE Transactions on Information Forensics and Security.* 2016;1(6):1304–1319. DOI:10.1109/TIFS.2015.2505626
31. Yakovlev V.A. Authentication of Keys Distributed by the Diffie – Hellman Method for Mobile Devices Based on Authentication Codes and Magnetometric Data. *SPIRAS Proceedings.* 2019;18(3):705–740. (in Russ.) DOI:10.15622/sp.2019.18.3.705-740. EDN:PRNILE

32. Yakovlev V., Korzhik V., Adadurov S. Authentication of Diffie-Hellman Protocol for Mobile Units Executing a Secure Device Pairing Procedure in Advance. *Proceedings of the 29th Conference of Open Innovations Association, FRUCT, 12–14 May 2021, Tampere, Finland*. 2021. p.385–392. EDN:DWHDGP
33. Yakovlev V. Method for Authentication of Diffie – Hellman Values Based on Pre-Distributed Random Sequences and Wegman – Carter One-Time Pad Algorithm. *Proceedings of Telecommunication Universities*. 2021;7(3):79–90. (in Russ.) DOI:10.31854/1813-324X-2021-7-3-79-90. EDN:TBVSMD
34. Nielsen M., Chang I. *Quantum computing and quantum information*. Moscow: Mir Publ.; 2006. 824 p. (in Russ.)
35. Castelvechi D. China's quantum satellite clears major hurdle on way to ultrasecure communications. *Nature*. 2017. DOI:10.1038/nature.2017.22142
36. Nield D. Quantum Teleportation Was Just Achieved With 90% Accuracy Over a 44km Distance. *ScienceAlert*. 2020. <https://www.sciencealert.com/scientists-achieve-sustained-high-fidelity-quantum-teleportation-over-44-km> [Accessed 10.05.2024]
37. Korpusov V., Olkhovoy O., Yakovlev V. The Research of the Random Number Generator Based on the Magnetometer. *Proceedings of the VIIth International Conference on Infotelecommunications in Science and Education, 28 February – 1 March 2018, St. Petersburg, Russia*. St. Petersburg: The Bonch-Bruevich Saint-Petersburg State University of Telecommunications Publ.; 2018. p.488–494. (in Russ.) EDN:OVMQXN
38. Wegman M.N., Carter J.L. New Hash Functions and their Use in Authentication and Set Equality. *Journal of Computer and System Sciences*. 1981;22:265–279. DOI:10.1016/0022-0000(81)90033-7

Статья поступила в редакцию 12.05.2024; одобрена после рецензирования 14.06.2024; принята к публикации 01.07.2024.

The article was submitted 12.05.2024; approved after reviewing 14.06.2024; accepted for publication 01.07.2024.

Информация об авторах:

КОРЖИК
Валерий Иванович

доктор технических наук, профессор, профессор кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0000-0002-8347-6527>

ЯКОВЛЕВ
Виктор Алексеевич

доктор технических наук, профессор, профессор кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0007-2861-9605>

ИЗОТОВ
Борис Викторович

кандидат технических наук, заместитель технического директора, начальник сектора защиты информации ЗАО «Научные приборы»
 <https://orcid.org/0009-0004-7081-3610>

СТАРОСТИН
Владимир Сергеевич

кандидат-физико-математических наук, доцент, доцент кафедры высшей математики Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0000-2939-1971>

БУЙНЕВИЧ
Михаил Викторович

доктор технических наук, профессор, профессор кафедры прикладной математики и информационных технологий Санкт-Петербургского университета ГПС МЧС России
 <https://orcid.org/0000-0001-8146-0022>

Коржик В.И. и Буйневич М.В. являются членами редакционного совета журнала «Труды учебных заведений связи» с 2016 г., но не имеют никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.

Korzhik V.I. and Buinevich M.V. have been a members of the journal "Proceedings of Telecommunication Universities" Editorial Council since 2016, but have 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-2024-10-4-142-148>

Интеллектуальный метод мутации входных корпусов с обратной связью

Николай Николаевич Самарин ✉, samarin_nik@mail.ru

Анастасия Викторовна Тулинова, yarmak.av@ibks.spbstu.ru

Научно-исследовательский институт «Квант»,
Москва, 125438, Российская Федерация

Аннотация

Фаззинг-тестирование представляется одним из эффективных способов повышения надежности программного обеспечения и входит в обязательный перечень мероприятий, проводимых на этапе квалификационного тестирования согласно национальному стандарту ГОСТ Р 56939-2016. Использование штатных мутаторов сводит реализацию данной задачи практически к полному перебору, что негативно сказывается на времени обнаружения некорректного поведения программы. В этой связи актуальным является вопрос рационализации подбора входных данных, при котором учитывается специфика корпуса данных, а также контекст, описывающий реакцию тестируемого программного обеспечения и позволяющий определить область и метод мутации на следующей итерации тестирования.

Цель настоящей работы – повышение эффективности фаззинг-тестирования за счет интеллектуализации штатного мутатора с использованием аппарата нейронных сетей, предполагающей учет синтаксических и семантических особенностей входного корпуса и использующей обратную связь от тестируемой программы.

Методы исследования. При выполнении работы использовались методы анализа и синтеза, теории алгоритмов, дискретной и вычислительной математики, машинного обучения.

Результаты. Рассмотрены преимущества и недостатки штатного модуля модификации входных корпусов – мутатора – фаззера AFL. Приведено обоснование выбора искусственной нейронной сети на базе архитектуры долгой краткосрочной памяти в качестве механизма, реализующего интеллектуальное управление процессом генерации и преобразования входных корпусов. Описан предлагаемый метод мутации, подразумевающий интеграцию в работу штатного мутатора механизма принятия решения об объеме и формате необходимых мутаций для увеличения покрытия кода, а также последующее уточнение входных данных шелл-кодом для проверки работоспособности фрагмента, вызвавшего нештатное поведение программного обеспечения. Представлена схема работы модуля мутации, включающего в себя компонент преобразования входных корпусов для генерации трасс исполнения программы и компонент, направленный на подтверждение концепта и повторный вызов нештатного поведения программного обеспечения с использованием сформированного шелл-кода.

Научная новизна. В отличие от известных предлагаемый метод использует обратную связь, фиксирующую реакцию программного обеспечения, при формировании стратегии мутации данных, что определяет научную новизну полученных результатов.

Значимость. Предложенное решение позволяет уменьшить время тестирования программы при сохранении уровня покрытия кода. Полученные в работе результаты являются универсальными и, в перспективе, могут быть использованы при фаззинг-тестировании методами белого, черного и серого ящика.

Ключевые слова: кибербезопасность, фаззинг-тестирование, мутация входного корпуса, искусственные нейронные сети, сеть с долгой краткосрочной памятью

Ссылка для цитирования: Самарин Н.Н., Тулинова А.В. Интеллектуальный метод мутации входных корпусов с обратной связью // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 142–148. DOI:10.31854/1813-324X-2024-10-4-142-148. EDN:GKUGLY

Original research

<https://doi.org/10.31854/1813-324X-2024-10-4-142-148>

Intelligent Method for Mutation of Input Cases with Feedback

 **Nikolay N. Samarin** ✉, samarin_nik@mail.ru

 **Anastasia V. Tulinova**, yarmak.av@ibks.spbstu.ru

Research Institute «Kvant»,
Moscow, 125438, Russian Federation

Annotation

Relevance. Fuzzing is one of the effective ways to improve the software reliability and is included in the mandatory list of research carried out at the stage of qualification testing according to national standard GOST R 56939-2016. The use of standard mutators reduces the fuzzing process to brute force, which negatively affects the time of incorrect program behavior detection. In this regard, it is important to rationalize the selection of input data, which takes into account the data corpus specifics, as well as the context describing the software response under test and allowing to determine the mutations at the next iteration of testing.

Purpose of the research is to increase the efficiency of fuzzing by intellectualizing the standard mutator using neural networks, which takes into account the syntactic and semantic features of the input corpus and uses program feedback.

Methods. The methods of analysis and synthesis, theory of algorithms, discrete and computational mathematics, machine learning were used.

Result. The advantages and disadvantages of the standard module for AFL fuzzer's input corpus mutation are considered. The justification of neural network choice with LSTM-architecture as a mechanism that realizes the intelligent control of input corpora's generation and transformation is given. The proposed mutation method is described, which implies the integration of decision making mechanism on the amount and format of necessary mutations to increase the code coverage into the standard mutator, as well as the subsequent refinement of input data by shell-code to check the operability of the fragment that caused abnormal software's behavior. The scheme of the mutation module is presented, which includes a component of input corporas conversion for generation of program execution traces and a component aimed at concept confirmation and re-call of abnormal software behavior using the generated shell-code.

Novelty. Unlike the known ones, the proposed method uses feedback, fixing the software reaction, when forming the data mutation strategy, which determines the scientific novelty of the obtained results.

Significance. The proposed solution allows reducing the program testing time while maintaining the code coverage. The results obtained in the research are universal and, in the future, can be used in white, black and gray box fuzzing methods.

Keywords: cybersecurity, fuzzing testing, input corpus mutation, artificial neural networks, long short-term memory network

For citation: Samarin N.N., Tulinova A.V. Intelligent Method for Mutation of Input Cases with Feedback. *Proceedings of Telecommunication Universities*. 2024;10(4):142–148. (in Russ.) DOI:10.31854/1813-324X-2024-10-4-142-148. EDN:GKUGLY

Введение

Для автоматизации динамического анализа исполняемых файлов широко применяется фаззинг-тестирование. В зависимости от полноты знаний об исполняемом файле, варьируется тип тестирования: методом черного, серого или белого ящика [1]. Однако, даже в условиях исчерпывающей для проведения фаззинг-тестирования полноты знаний, одной из проблем является длительное время

получения результата ввиду высокого показателя случайности мутации входных данных.

Наиболее широко используемым на сегодняшний день является фаззер AFL (аббр. от англ. American Fuzzy Lop) [1–3]. Как любое комплексное программное обеспечение (ПО), AFL состоит из набора модулей, которые могут быть изменены пользователем в зависимости от их цели тестирования [3]. Модуль, функционал которого рассмат-

ривается в рамках настоящего исследования, называется мутатором. Его задачей является подготовка нового входного корпуса на основе того, который подавался на вход ПО на прошлой итерации. Штатный мутатор не обладает возможностями анализа входных корпусов [2, 4]. В нем реализованы только базовые операции, такие как инвертирование бита, подмена строки или подстроки, сокращение длины корпуса или его увеличение произвольными данными и т. д.

Таким образом, применение штатного мутатора в ходе фаззинг-тестирования, по сути, мало отличается от метода перебора, что приводит к увеличению времени обнаружения случаев некорректного поведения тестируемого ПО [2]. Кроме того, полученный корпус, для которого выявлено некорректное поведение ПО, не может быть доработан до полноценного подтверждения концепта. Вместо этого пользователь должен самостоятельно реализовать его, опираясь на характеристики корпуса, такие как его размер, семантические особенности и пр.

Перспективным механизмом генерации входных корпусов на основе предыдущего опыта, то есть, с удержанием контекста, является искусственная нейронная сеть (ИНС) на базе архитектуры долгой краткосрочной памяти (LSTM, аббр. от англ. Long Short Term Memory). Данный тип ИНС имеет ключевую особенность – вентили забывания, которыми может быть задано окно удержания опыта. Реализация мутатора с использованием ИНС с долгой краткосрочной памятью обеспечивает возможность обработки обратной связи, поступающей в качестве реакции тестируемого ПО на входной корпус, и его преобразование – мутацию – в соответствии с полученной реакцией.

Процесс мутации входных корпусов

Фаззинг-тестирование с применением мутации – один из нескольких видов техник фаззинга. При таком тестировании фаззеру предоставляются сведения о полном наборе начальных входных корпусов: формат, максимальная и минимальная длина, наличия закономерностей или последовательностей в корпусе и т. д. На основе полученной информации мутатор – модуль преобразования – создает новые образцы входных корпусов, которые изменены в соответствии с указанными правилами мутации [3, 5]. Штатный мутатор поддерживает базовые преобразования, например, инвертирование произвольного бита входного корпуса, дополнение входного корпуса под последовательностью, перестановка слов и групп слов входного корпуса. Общий принцип работы мутатора представлен на рисунке 1.

Совокупное применение данных преобразований позволяет получить огромное количество вариантов входных корпусов, однако базовый мутатор не учитывает их синтаксические и семантические особенности, вследствие чего уровень покрытия на протяжении долгого времени может оставаться низким [4]. Прежде всего это связано с тем, что для запуска ПО чаще всего требуется ввести команду с заданными параметрами, которые стандартизированы разработчиком, а при получении в качестве аргументов сырых данных, ПО, вероятнее всего, будет выводить сообщение с поддерживаемыми командами.

В зависимости от целей фаззинг-тестирования создаются собственные мутаторы, расширяющие функционал штатных. Основной задачей мутатора является создание разнообразного множества входных корпусов с целью обеспечения прироста покрытия на каждой итерации фаззинг-тестирования.

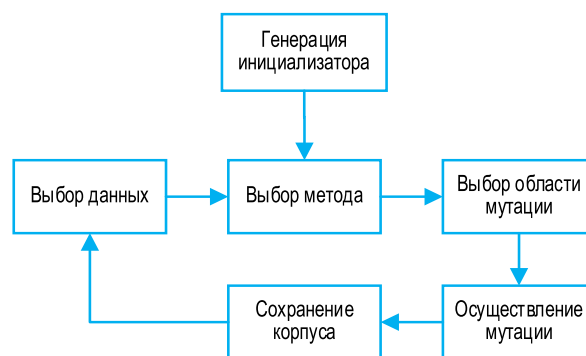


Рис. 1. Общий принцип работы мутатора

Fig. 1. The General Principle of Operation of the Mutator

Таким образом, для выполнения, например, фаззинг-тестирования сетевого сокета необходимо реализовать мутатор сетевого трафика. Мутатор сырых данных будет модифицировать сетевой пакет без учета требований к его формату, что приведет к отбрасыванию сетевого пакета в стеке TCP/IP до его попадания в сетевой сокет и ПО.

Несмотря на такие преимущества мутационных фаззеров, как простота настройки и скорость создания новых входных корпусов, они имеют два существенных недостатка, обусловленных самим механизмом мутации:

1) мутационные фаззеры могут пропустить комплексные уязвимости, требующие значительных изменений данных ввиду простоты выполняемых в отношении входных корпусов преобразований [3].

2) мутационные фаззеры не смогут воспроизвести все возможные формы протокола или формата файла в соответствии со спецификацией ввиду отсутствия правил для модификации файла или кадра протокола [1].

Данные недостатки могут быть компенсированы разработкой собственного мутатора с заданным набором правил и применением комплексных операций в отношении корпусов, однако такой мутатор может быть применим только в отношении конкретного экземпляра или нескольких схожих экземпляров ПО, он не обеспечивает полной унификации. Для унификации мутатора необходимо использование механизма, который обеспечивает мутацию на основе контекста (например, принимаемых на вход типов данных) и обратной связи – реакции программы на входной корпус.

Наиболее перспективны для решения такой задачи технологии машинного обучения и ИНС. Более того, обработку обратной связи с использованием ИНС возможно реализовать с использованием рекуррентного механизма в сетях с архитектурой LSTM.

Возможности искусственной нейронной сети с долгой краткосрочной памятью

Архитектура ИНС с долгой краткосрочной памятью основана на принципе удержания предыдущего опыта, полученного нейронной сетью, с помощью вентиля забывания. Такие сети являются разновидностью рекуррентных нейронных сетей, то есть, сетей, имеющих обратную связь [6]. Благодаря этому LSTM-сети решают проблему долговременной зависимости – запоминания в рамках достаточно больших окон [7]. Традиционно LSTM-сети используются в обработке естественного языка, распознавании речи, а также в генерации музыки и текста [6, 7].

Как любая рекуррентная сеть, LSTM-сеть состоит из цепочки повторяющихся модулей обычной нейронной сети, включающей в себя четыре слоя. Ключевым компонентом модуля LSTM является состояние ячейки – аналога конвейерной ленты, которая проходит через все модули сети. В тракте состояния ячейки данные передаются свободно и в отношении них могут быть применены различного рода линейные операции.

Для контроля состояния ячейки применяются фильтры – вентили забывания, традиционно представляющие собой слой сигмоидальной нейронной сети и операции поточечного умножения. Данный слой возвращает значение в интервале от 0 до 1 включительно, означающее, какой объем данных пропустить в следующий модуль. В LSTM-сети реализовано три таких фильтра, которые позволяют защищать данные от забывания и контролировать состояния ячейки.

В ходе обучения для минимизации ошибки на всем множестве тренировочных последовательностей применяется метод обратного распростране-

ния ошибки, развернутый во времени. Это, в свою очередь, позволяет корректировать веса пропорционально производной градиентного спуска в зависимости от величины ошибки.

Несмотря на один из важнейших недостатков применения градиентного спуска для обучения стандартных рекуррентных ИНС – градиенты ошибок уменьшаются с экспоненциальной скоростью по мере увеличения временной задержки между важными событиями – его применение для обучения LSTM-сетей достаточно эффективно [8]. Это обусловлено механизмом контроля состояния ячейки. Когда величины ошибки распространяются в обратном направлении от выходного слоя, ошибка оказывается «заперта» в памяти блока, что создает ситуацию «карусели ошибок», которая непрерывно передает ошибку обратно каждому из вентилях, пока они не будут натренированы отбрасывать значение.

Принцип работы сетей с долгой краткосрочной памятью состоит из четырех этапов [7, 9].

Этап 1. Данные, поступающие в модуль сети, передаются на слой фильтра забывания, определяющий объем информации, который пропускается в следующий слой [8]. Значение предыдущего выхода h_{t-1} и текущего входа x_t передаются в сигмоидальный слой и приобретают коэффициенты от 0 до 1. Значения с коэффициентами ближе к 0 забываются, ближе к 1 – передаются далее. Данный этап описывается формулой:

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f). \quad (1)$$

Этап 2. Определяется, какая информация будет храниться в состоянии ячейки [7]. Данный этап делится на два подэтапа. Сначала сигмоидальный слой – слой входного фильтра – определяет, какие значения должны быть обновлены:

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i). \quad (2)$$

После чего слой гиперболического тангенса строит вектор новых значений кандидатов C'_t , которые могут быть добавлены в состояние ячейки:

$$C'_t = \tanh(W_c * [h_{t-1}, x_t] + b_c). \quad (3)$$

Этап 3. Далее для того, чтобы изменить старое состояние ячейки C_{t-1} на C_t , необходимо умножить его на значение f_t и прибавить $i_t * C'_t$ (4). Таким образом, применяется механизм забывания в отношении состояния ячейки [9, 10]. Фактически, полученные значения будут являться кандидатами, умноженными на число необходимых изменений:

$$C_t = f_t * C_{t-1} + i_t * C'_t. \quad (4)$$

Этап 4. Последний этап заключается в определении объема выходной информации [10]. Значе-

ния предыдущего выхода h_{t-1} и текущего входа x_t проходят сигмоидальный слой, который определяет объем вывода из состояния ячейки:

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o). \quad (5)$$

После этого значения состояния ячейки проходят слой гиперболического тангенса для получения на выходе значений в диапазоне от 1 до -1, которые, в свою очередь, перемножаются с выходными значениями сигмоидального слоя:

$$h_t = o_t * \tanh(C_t). \quad (6)$$

Полученные значения h_t и C_t передаются в следующий модуль цепочки.

Описание метода мутации

В рамках предлагаемого метода предполагается модификация штатного мутатора механизмом, получающим на вход предыдущий корпус I_{t-1} и трассу исполнения T_t , которую этот корпус прошел в тестируемой программе. На основании значений кортежа входных данных (I_{t-1}, T_t) вычисляются необходимые для увеличения покрытия мутации M_1 , входящие во множество возможных операций преобразования:

$$M_1 = [F_b, F_s, S_a, S_d], \quad (7)$$

где $F_b(v) = \neg v$ – инвертирование значения произвольного бита v ; $F_s(s) = \neg s$ – инвертирование значения произвольной подстроки s ; $S_a(s, d) = s + d$ – добавление к исходной строке s другой подстроки d ; $S_d(s, d) = s - d$ – удаление из исходной строки s подстроки d .

Объем необходимых мутаций при этом оценивается, исходя из длины трассы, фиксируемой в рамках заданного временного окна: если длина трассы не изменяется на протяжении N запусков, количество мутаций входного корпуса, осуществляемых в рамках одной итерации, увеличивается до тех пор, пока не будет зафиксирован прирост по покрытию. При обнаружении прироста количество мутаций снижается до исходного значения.

После того, как в ходе фаззинг-тестирования установлено нештатное завершение работы тестируемого ПО, мутация с целью увеличения покрытия прекращается. Вместо этого происходит загрузка второго блока входных корпусов, содержащих исполняемый код для создания подтвержденного концепта. Входные корпуса второго блока состоят из двух частей: входного корпуса, вызвавшего нештатное завершение работы, и шелл-кода.

Мутация на данном этапе выполняется только в отношении шелл-кода в соответствии с заданными правилами M_2 :

$$M_2 = [N_a, N_d, A_m], \quad (8)$$

где $N_a(n)$ – увеличение NOP-дорожки (NOP, аббр. от англ. No OPeration) на значение n ; $N_d(n)$ – уменьшение NOP-дорожки на значение n ; $A_m(a_{\min}, a_{\max})$ – изменение адреса возврата на произвольный, в диапазоне $[a_{\min}, a_{\max}]$.

Данный этап выполняется до тех пор, пока шелл-код не будет выполнен, то есть пока не прекратится нештатное завершение тестируемого ПО. После этого фаззер переходит к новой итерации и продолжает искать новые трассы в ПО. Наглядно действие описанного метода представлено на рисунке 2.

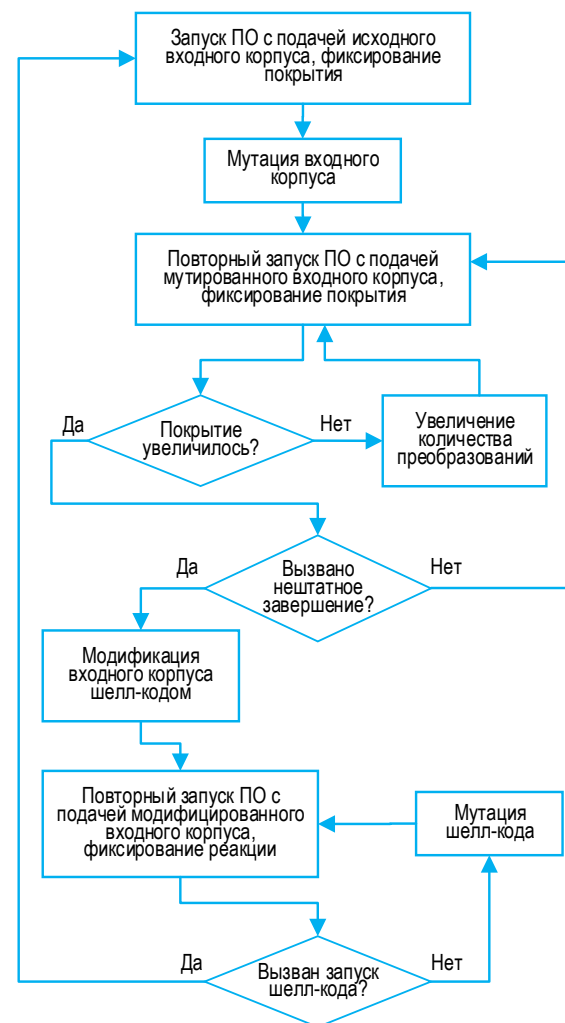


Рис. 2. Схема работы интеллектуального метода мутации входных данных

Fig. 2. The Schematic Diagram of the Input Mutation Intelligent Method Operation

Заключение

Фаззинг-тестирование является основным подходом при проведении динамического анализа программного обеспечения на предмет наличия уязвимостей, архитектурных или логических ошибок. Основным недостатком данного подхода является время его проведения ввиду хаотичности генера-

ции входных корпусов, подаваемых тестируемой программе [3, 4, 11]. В рамках проведенных исследований создан интеллектуальный метод мутации входных корпусов, который использует обратную связь, получаемую от объекта исследований и на ее основе принимает решение о необходимости увеличения количества модификаций. Метод включает в себя два этапа, имеющие различные цели.

На первом этапе фаззеру требуется вызвать нештатное завершение работы программы и найти как можно большее количество трасс для дальнейшего анализа. Возможные в программе трассы используются после работы второго этапа в качестве отправной точки для новой итерации.

На втором этапе, который наступает после обнаружения нештатного завершения работы про-

граммы, осуществляется дополнение входного корпуса шелл-кодом. На данном этапе целью является запуск этого шелл-кода в контексте тестируемой программы для подтверждения концепта выявленной уязвимости. Таким образом, на втором этапе мутации подвергается не исходная часть входного корпуса, а сам шелл-код.

Дальнейшие исследования в данной области должны быть направлены на создание интеллектуального метода мутации шелл-кода на основе реакции ПО. Такой метод позволит повысить производительность за счет формализации правил модификации шелл-кода и снижения количества некорректных с точки зрения синтаксиса и семантики исполняемого кода мутаций [12].

Список источников

1. Muduli S.K., Roy S. Satisfiability modulo fuzzing: a synergistic combination of SMT solving and fuzzing // *Proceedings of the ACM on Programming Languages*. 2022. Vol. 6. Iss. OOPSLA2. PP. 1236–1263. DOI:10.1145/3563332. EDN:VUQLTQ
2. Liu Z., Qian P., Yang J., Liu L., Xu X., He Q., et. al. Rethinking Smart Contract Fuzzing: Fuzzing with Invocation Ordering and Important Branch Revisiting // *IEEE Transactions on Information Forensics and Security*. 2023. Vol. 18. PP. 1237–1251. DOI:10.1109/tifs.2023.3237370. EDN:ZWKGJS
3. Ерышов В.Г. Фаззинг тестирование. Классификация современных средств фаззинга // Сборник избранных статей по материалам научных конференций ГНИИ "Нацразвитие" Международные научные конференции (Санкт-Петербург, Россия, 26–31 августа 2021 года). СПб.: ГНИИ «Нацразвитие», 2021. С. 287–289. DOI:10.37539/AUG298. 2021.94.77.007. EDN:QZILPI
4. Situ L.-Y., Zuo Z.-Q., Guan L., Wang L.-Z., Li X.-D., Shi J., et. al. Vulnerable Region-Aware Greybox Fuzzing // *Journal of Computer Science and Technology*. 2021. Vol. 36. Iss. 5. PP. 1212–1228. DOI:10.1007/s11390-021-1196-0. EDN:PAPPKT
5. Kim S.J., Shon T. Field classification-based novel fuzzing case generation for ICS protocols // *The Journal of Supercomputing*. 2018. Vol. 74. Iss. 9. PP. 4434–4450. DOI:10.1007/s11227-017-1980-3. EDN:TLEZVS
6. Wei W., Li X., Zhang B., Li L., Damaševičius R., Scherer R. LSTM-SN: complex text classifying with LSTM fusion social network // *The Journal of Supercomputing*. 2023. Vol. 79. Iss. 9. PP. 9558–9583. DOI:10.1007/s11227-022-05034-w. EDN:TSXGJI
7. Bayram F., Aupke P., Bestoun S.A., Kassler A., Theocharis A., Forsman J. DA-LSTM: A dynamic drift-adaptive learning framework for interval load forecasting with LSTM networks // *Engineering Applications of Artificial Intelligence*. 2023. Vol. 123. P. 106480. DOI:10.1016/j.engappai.2023.106480. EDN:WOXZSB
8. Pierre A.A., Akim S.A., Semeno A.K., Babiga B. Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches // *Energies*. 2023. Vol. 16. Iss. 12. P. 4739. DOI:10.3390/en16124739. EDN:RKROHA
9. Singh P., Kumar Ch., Kumar A. Next-LSTM: a novel LSTM-based image captioning technique // *International Journal of System Assurance Engineering and Management*. 2023. Vol. 14. Iss. 4. PP. 1492–1503. DOI:10.1007/s13198-023-01956-7. EDN:QUOVUU
10. Wen X., Li W. Wen X. Time Series Prediction Based on LSTM-Attention-LSTM Model // *IEEE Access*. 2023. Vol. 11. PP. 48322–48331. DOI:10.1109/access.2023.3276628. EDN:ZAABGZ
11. Ding T., Fu J., Shen R. Research on Multidimensional Mutation Strategy Method of Fuzzing Test // *Proceedings of the 40th Chinese Control Conference (CCC, Shanghai, China 26–28 July 2021)*. Shanghai: IEEE, 2021. PP. 8639–8644. DOI:10.23919/CCC52363.2021.9550435. EDN:MJGBJG
12. Manès V.J.M., Han H., Han C., Cha S.K., Egele M., Woo M. The Art, Science, and Engineering of Fuzzing: A Survey // *IEEE Transactions on Software Engineering*. 2021. Vol. 47. Iss. 11. PP. 2312–2331. DOI:10.1109/TSE.2019.2946563. EDN:ZDDKFN

References

1. Muduli S.K., Roy S. Satisfiability modulo fuzzing: a synergistic combination of SMT solving and fuzzing. *Proceedings of the ACM on Programming Languages*. 2022;6(OOPSLA2):1236–1263. DOI:10.1145/3563332. EDN:VUQLTQ
2. Liu Z., Qian P., Yang J., Liu L., Xu X., He Q., et. al. Rethinking Smart Contract Fuzzing: Fuzzing with Invocation Ordering and Important Branch Revisiting. *IEEE Transactions on Information Forensics and Security*. 2023;18:1237–1251. DOI:10.1109/tifs.2023.3237370. EDN:ZWKGJS

3. Yeryshov V.G. Fuzzing testing. Classification of modern fuzzing tools. *Proceedings of International Scientific Conferences of the National Research Institute "National Development"*, 26–31 august 2021, St. Petersburg, Russia. St. Petersburg: GNII "National Development" Publ.; 2021. p.287–289. DOI:10.37539/AUG298.2021.94.77.007. EDN:QZILPI
4. Situ L.-Y., Zuo Z.-Q., Guan L., Wang L.-Z., Li X.-D., Shi J., et. al. Vulnerable Region-Aware Greybox Fuzzing. *Journal of Computer Science and Technology*. 2021;36(5):1212–1228. DOI:10.1007/s11390-021-1196-0. EDN:PAPPKT
5. Kim S. J., Shon T. Field classification-based novel fuzzing case generation for ICS protocols. *The Journal of Supercomputing*. 2018;74(9):4434–4450. DOI:10.1007/s11227-017-1980-3. EDN:TLEZVS
6. Wei W., Li X., Zhang B., Li L., Damaševičius R., Scherer R. LSTM-SN: complex text classifying with LSTM fusion social network. *The Journal of Supercomputing*. 2023;79(9):9558–9583. DOI:10.1007/s11227-022-05034-w. EDN:TSXGJI
7. Bayram F., Aupke P., Bestoun S.A., Kassler A., Theocharis A., Forsman J. DA-LSTM: A dynamic drift-adaptive learning framework for interval load forecasting with LSTM networks. *Engineering Applications of Artificial Intelligence*. 2023; 123:106480. DOI:10.1016/j.engappai.2023.106480. EDN:WOXZSB
8. Pierre A.A., Akim S.A., Semeny A.K., Babiga B. Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches. *Energies*. 2023;16(12):4739. DOI:10.3390/en16124739. EDN:RKROHA
9. Singh P., Kumar Ch., Kumar A. Singh P. Next-LSTM: a novel LSTM-based image captioning technique. *International Journal of System Assurance Engineering and Management*. 2023;14(4):1492–1503. DOI:10.1007/s13198-023-01956-7. EDN:QUOVUU
10. Wen X., Li W. Wen X. Time Series Prediction Based on LSTM-Attention-LSTM Model. *IEEE Access*. 2023;11: 48322–48331. DOI:10.1109/access.2023.3276628. EDN:ZAABGZ
11. Ding T., Fu J., Shen R. Ding T. Research on Multidimensional Mutation Strategy Method of Fuzzing Test *Proceedings of the 40th Chinese Control Conference, CCC, 26–28 July 2021, Shanghai, China*. Shanghai: IEEE; 2021. p.8639–8644. DOI:10.23919/CCC52363.2021.9550435. EDN:MJGBJG
12. Manès V.J.M., Han H., Han C., Cha S.K., Egele M., Woo M. The Art, Science, and Engineering of Fuzzing: A Survey. *IEEE Transactions on Software Engineering*. 2021;47(11):2312–2331. DOI:10.1109/TSE.2019.2946563. EDN:ZDDKFN

Статья поступила в редакцию 27.05.2024; одобрена после рецензирования 10.08.2024; принята к публикации 17.08.2024.

The article was submitted 27.05.2024; approved after reviewing 10.08.2024; accepted for publication 17.08.2024.

Информация об авторах:

САМАРИН
Николай Николаевич

кандидат технических наук, начальник научно-исследовательского отделения
№ 6 «Научно-исследовательского института «Квант»
 <https://orcid.org/0009-0007-4911-8471>

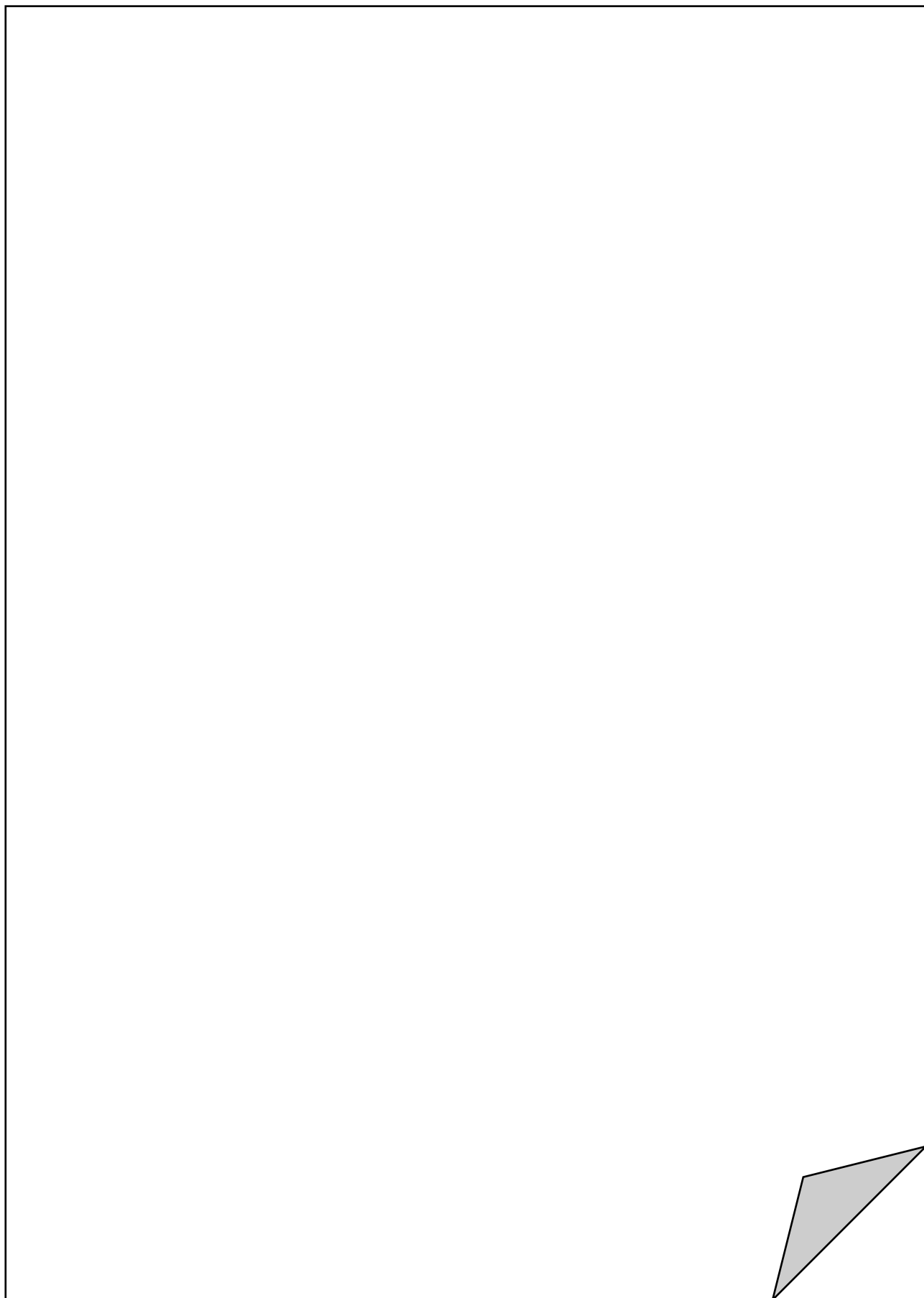
ТУЛИНОВА
Анастасия Викторовна

кандидат технических наук, инженер 1 категории научно-исследовательского
отделения № 6 «Научно-исследовательского института «Квант»
 <https://orcid.org/0000-0002-7121-6031>

Авторы сообщают об отсутствии конфликтов интересов.

The authors declare no conflicts of interests.

ДЛЯ ЗАМЕТОК



План издания научной литературы 2024 г., п. 9

Усл.-печ. л.
16,0

Формат
60×84_{1/8}

Заказ
№ 1606

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

