



R U S S I A N

T E C H N O L O G I C A L J O U R N A L

**РОССИЙСКИЙ
ТЕХНОЛОГИЧЕСКИЙ
ЖУРНАЛ**

*Информационные системы.
Информатика.
Проблемы информационной безопасности*

*Роботизированные комплексы и системы.
Технологии дистанционного зондирования и неразрушающего контроля*

Современные радиотехнические и телекоммуникационные системы

*Микро- и нанoeлектроника.
Физика конденсированного состояния*

Аналитическое приборостроение и технологии

Математическое моделирование

*Экономика наукоемких и высокотехнологичных
предприятий и производств. Управление в организационных системах*

Управление качеством продукции. Стандартизация

Мировоззренческие основы технологии и общества



RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

- Информационные системы. Информатика. Проблемы информационной безопасности
 - Роботизированные комплексы и системы. Технологии дистанционного зондирования и неразрушающего контроля
 - Современные радиотехнические и телекоммуникационные системы
 - Микро- и нанoeлектроника. Физика конденсированного состояния
 - Аналитическое приборостроение и технологии
 - Математическое моделирование
 - Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах
 - Управление качеством продукции. Стандартизация
 - Мировоззренческие основы технологии и общества
- Information systems. Computer sciences. Issues of information security
 - Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing
 - Modern radio engineering and telecommunication systems
 - Micro- and nanoelectronics. Condensed matter physics
 - Analytical instrument engineering and technology
 - Mathematical modeling
 - Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems
 - Product quality management. Standardization
 - Philosophical foundations of technology and society

Russian Technological Journal
2024, том 12, № 3

Russian Technological Journal
2024, Vol. 12, No. 3

Russian Technological Journal 2024, том 12, № 3

Дата опубликования 31 мая 2024 г.

Научно-технический рецензируемый журнал освещает вопросы комплексного развития радиотехнических, телекоммуникационных и информационных систем, электроники и информатики, а также результаты фундаментальных и прикладных междисциплинарных исследований, технологических и организационно-экономических разработок, направленных на развитие и совершенствование современной технологической базы.

Периодичность: один раз в два месяца.

Журнал основан в декабре 2013 года. До 2016 г. издавался под названием «Вестник МГТУ МИРЭА» (ISSN 2313-5026), а с января 2016 г. по июль 2021 г. под названием «Российский технологический журнал» (ISSN 2500-316X).

Учредитель и издатель:

федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет» 119454, РФ, г. Москва, пр-т Вернадского, д. 78.

Журнал входит в Перечень ведущих рецензируемых научных журналов ВАК РФ, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук и доктора наук, индексируется в РГБ, РИНЦ, eLibrary, Соционет, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

Главный редактор:

Сигов Александр Сергеевич, академик РАН, доктор физ.-мат. наук, профессор, президент ФГБОУ ВО МИРЭА – Российский технологический университет (РТУ МИРЭА), Москва, Россия.
Scopus Author ID 35557510600, ResearcherID L-4103-2017, sigov@mirea.ru.

Редакция:

Зав. редакцией	к.т.н. Г.Д. Середина
Научный редактор	д.т.н., проф. Г.В. Куликов
Выпускающий редактор	А.С. Алексеенко
Технический редактор	Д.В. Трофимова

119571, г. Москва, пр-т Вернадского, 86, оф. Л-119.
Тел.: +7 (499) 600-80-80 (#31288).
E-mail: seredina@mirea.ru.

Регистрационный номер и дата принятия решения о регистрации СМИ ПИ № ФС 77 - 81733 от 19.08.2021 г. СМИ зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Индекс по объединенному каталогу «Пресса России» 79641.

<https://www.rtlj-mirea.ru>

Russian Technological Journal 2024, Vol. 12, No. 3

Publication date May 31, 2024.

The peer-reviewed scientific and technical journal highlights the issues of complex development of radio engineering, telecommunication and information systems, electronics and informatics, as well as the results of fundamental and applied interdisciplinary researches, technological and economical developments aimed at the development and improvement of the modern technological base.

Periodicity: bimonthly.

The journal was founded in December 2013. The titles were «Herald of MSTU MIREA» until 2016 (ISSN 2313-5026) and «Rossiiskii tekhnologicheskii zhurnal» from January 2016 until July 2021 (ISSN 2500-316X).

Founder and Publisher:

Federal State Budget
Educational Institution of Higher Education
«MIREA – Russian Technological University»
78, Vernadskogo pr., Moscow, 119454 Russia.

The journal is included into the List of peer-reviewed science press of the State Commission for Academic Degrees and Titles of Russian Federation. The Journal is included in Russian State Library (RSL), Russian Science Citation Index, eLibrary, Socionet, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

Editor-in-Chief:

Alexander S. Sigov, Academician at the Russian Academy of Sciences, Dr. Sci. (Phys.–Math.), Professor, President of MIREA – Russian Technological University (RTU MIREA), Moscow, Russia.
Scopus Author ID 35557510600, ResearcherID L-4103-2017, sigov@mirea.ru.

Editorial staff:

Managing Editor	Cand. Sci. (Eng.) Galina D. Seredina
Scientific Editor	Dr. Sci. (Eng.), Prof. Gennady V. Kulikov
Executive Editor	Anna S. Alekseenko
Technical Editor	Darya V. Trofimova

86, Vernadskogo pr., Moscow, 119571 Russia.
Phone: +7 (499) 600-80-80 (#31288).
E-mail: seredina@mirea.ru.

The registration number ПИ № ФС 77 - 81733 was issued in August 19, 2021 by the Federal Service for Supervision of Communications, Information Technology, and Mass Media of Russia.

The subscription index of *Pressa Rossii*: 79641.

Редакционная коллегия

- Кудж Станислав Алексеевич** д.т.н., профессор, ректор РТУ МИРЭА, Москва, Россия. Scopus Author ID 56521711400, ResearcherID AAG-1319-2019, <https://orcid.org/0000-0003-1407-2788>, rector@mirea.ru
- Банис Юрас Йонович** хабилитированный доктор наук, профессор, проректор Вильнюсского университета, Вильнюс, Литва. Scopus Author ID 7003687871, juras.banys@ff.vu.lt
- Бетелин Владимир Борисович** академик Российской академии наук (РАН), д.ф.-м.н., профессор, научный руководитель Федерального научного центра «Научно-исследовательский институт системных исследований» РАН, Москва, Россия. Scopus Author ID 6504159562, ResearcherID J-7375-2017, betelin@niisi.msk.ru
- Боков Алексей Алексеевич** д.ф.-м.н., старший научный сотрудник, химический факультет и 4D LABS, Университет Саймона Фрейзера, Ванкувер, Британская Колумбия, Канада. Scopus Author ID 35564490800, ResearcherID C-6924-2008, <http://orcid.org/0000-0003-1126-3378>, abokov@sfu.ca
- Вахрушев Сергей Борисович** д.ф.-м.н., профессор, заведующий лабораторией нейтронных исследований Физико-технического института им. А.Ф. Иоффе РАН, профессор кафедры Физической электроники СПбГПУ, Санкт-Петербург, Россия. Scopus Author ID 7004228594, ResearcherID A-9855-2011, <http://orcid.org/0000-0003-4867-1404>, s.vakhrushev@mail.ioffe.ru
- Гуляев Юрий Васильевич** академик РАН, д.ф.-м.н., профессор, научный руководитель Института радиотехники и электроники им. В.А. Котельникова РАН, Москва, Россия. Scopus Author ID 35562581800, gulyaev@cplire.ru
- Жуков Дмитрий Олегович** д.т.н., профессор, заведующий кафедрой интеллектуальных технологий и систем РТУ МИРЭА, Москва, Россия. Scopus Author ID 57189660218, zhukov_do@mirea.ru
- Кимель Алексей Вольдемарович** к.ф.-м.н., профессор, Университет Радбауд, г. Наймеген, Нидерланды. Scopus Author ID 6602091848, ResearcherID D-5112-2012, a.kimel@science.ru.nl
- Крамаров Сергей Олегович** д.ф.-м.н., профессор, Сургутский государственный университет, Сургут, Россия. Scopus Author ID 56638328000, ResearcherID E-9333-2016, <https://orcid.org/0000-0003-3743-6513>, mavoo@yandex.ru
- Новиков Дмитрий Александрович** академик РАН, д.т.н., директор Института проблем управления им. В.А. Трапезникова РАН, Москва, Россия. Scopus Author ID 7102213403, ResearcherID Q-9677-2019, <https://orcid.org/0000-0002-9314-3304>, novikov@ipu.ru
- Перно Филипп** Dr. Sci. (Electronics), профессор, Центральная Школа г. Лилль, Франция. Scopus Author ID 7003429648, philippe.pernod@ec-lille.fr
- Романов Михаил Петрович** д.т.н., профессор, директор Института искусственного интеллекта РТУ МИРЭА, Москва, Россия. Scopus Author ID 14046079000, <https://orcid.org/0000-0003-3353-9945>, m_romanov@mirea.ru
- Савиных Виктор Петрович** академик РАН, Дважды Герой Советского Союза, д.т.н., профессор, президент Московского государственного университета геодезии и картографии, Москва, Россия. Scopus Author ID 56412838700, vp@miigaik.ru
- Соболевский Андрей Николаевич** д.ф.-м.н., директор Института проблем передачи информации им. А.А. Харкевича, Москва, Россия. Scopus Author ID 7004013625, ResearcherID D-9361-2012, <http://orcid.org/0000-0002-3082-5113>, sobolevski@iitp.ru
- Сюй Ли Да** академик Европейской академии наук, Российской инженерной академии и Инженерной академии Армении, Dr. Sci. (Systems Science), профессор, Университет Олд Доминион, Норфолк, Соединенные Штаты Америки. Scopus Author ID 13408889400, <https://orcid.org/0000-0002-5954-5115>, lxu@odu.edu
- Харин Юрий Семенович** академик Национальной академии наук Беларуси, д.ф.-м.н., профессор, директор НИИ прикладных проблем математики и информатики Белорусского государственного университета, Минск, Беларусь. Scopus Author ID 6603832008, <http://orcid.org/0000-0003-4226-2546>, kharin@bsu.by
- Чаплыгин Юрий Александрович** академик РАН, д.т.н., профессор, член Отделения нанотехнологий и информационных технологий РАН, президент Института микроприборов и систем управления им. Л.Н. Преснухина НИУ «МИЭТ», Москва, Россия. Scopus Author ID 6603797878, ResearcherID B-3188-2016, president@miet.ru
- Шпак Василий Викторович** к.э.н., зам. министра промышленности и торговли Российской Федерации, Министерство промышленности и торговли РФ, Москва, Россия; доцент, Институт микроприборов и систем управления им. Л.Н. Преснухина НИУ «МИЭТ», Москва, Россия, mishinevaiv@minprom.gov.ru

Editorial Board

- Stanislav A. Kudzh** Dr. Sci. (Eng.), Professor, Rector of RTU MIREA, Moscow, Russia. Scopus Author ID 56521711400, ResearcherID AAG-1319-2019, <https://orcid.org/0000-0003-1407-2788>, rector@mirea.ru
- Juras Banys** Habilitated Doctor of Sciences, Professor, Vice-Rector of Vilnius University, Vilnius, Lithuania. Scopus Author ID 7003687871, juras.banys@ff.vu.lt
- Vladimir B. Betelin** Academician at the Russian Academy of Sciences (RAS), Dr. Sci. (Phys.-Math.), Professor, Supervisor of Scientific Research Institute for System Analysis, RAS, Moscow, Russia. Scopus Author ID 6504159562, ResearcherID J-7375-2017, betelin@niisi.msk.ru
- Alexei A. Bokov** Dr. Sci. (Phys.-Math.), Senior Research Fellow, Department of Chemistry and 4D LABS, Simon Fraser University, Vancouver, British Columbia, Canada. Scopus Author ID 35564490800, ResearcherID C-6924-2008, <http://orcid.org/0000-0003-1126-3378>, abokov@sfu.ca
- Sergey B. Vakhrushev** Dr. Sci. (Phys.-Math.), Professor, Head of the Laboratory of Neutron Research, A.F. Ioffe Physico-Technical Institute of the RAS, Department of Physical Electronics of St. Petersburg Polytechnic University, St. Petersburg, Russia. Scopus Author ID 7004228594, ResearcherID A-9855-2011, <http://orcid.org/0000-0003-4867-1404>, s.vakhrushev@mail.ioffe.ru
- Yury V. Gulyaev** Academician at the RAS, Dr. Sci. (Phys.-Math.), Professor, Supervisor of V.A. Kotelnikov Institute of Radio Engineering and Electronics of the RAS, Moscow, Russia. Scopus Author ID 35562581800, gulyaev@cplire.ru
- Dmitry O. Zhukov** Dr. Sci. (Eng.), Professor, Head of the Department of Intelligent Technologies and Systems, RTU MIREA, Moscow, Russia. Scopus Author ID 57189660218, zhukov_do@mirea.ru
- Alexey V. Kimel** PhD (Phys.-Math.), Professor, Radboud University, Nijmegen, Netherlands, Scopus Author ID 6602091848, ResearcherID D-5112-2012, a.kimel@science.ru.nl
- Sergey O. Kramarov** Dr. Sci. (Phys.-Math.), Professor, Surgut State University, Surgut, Russia. Scopus Author ID 56638328000, ResearcherID E-9333-2016, <https://orcid.org/0000-0003-3743-6513>, mavoo@yandex.ru
- Dmitry A. Novikov** Academician at the RAS, Dr. Sci. (Eng.), Director of V.A. Trapeznikov Institute of Control Sciences, Moscow, Russia. Scopus Author ID 7102213403, ResearcherID Q-9677-2019, <https://orcid.org/0000-0002-9314-3304>, novikov@ipu.ru
- Philippe Pernod** Dr. Sci. (Electronics), Professor, Dean of Research of Centrale Lille, Villeneuve-d'Ascq, France. Scopus Author ID 7003429648, philippe.pernod@ec-lille.fr
- Mikhail P. Romanov** Dr. Sci. (Eng.), Professor, Director of the Institute of Artificial Intelligence, RTU MIREA, Moscow, Russia. Scopus Author ID 14046079000, <https://orcid.org/0000-0003-3353-9945>, m_romanov@mirea.ru
- Viktor P. Savinykh** Academician at the RAS, Dr. Sci. (Eng.), Professor, President of Moscow State University of Geodesy and Cartography, Moscow, Russia. Scopus Author ID 56412838700, vp@miigaik.ru
- Andrei N. Sobolevski** Professor, Dr. Sci. (Phys.-Math.), Director of Institute for Information Transmission Problems (Kharkevich Institute), Moscow, Russia. Scopus Author ID 7004013625, ResearcherID D-9361-2012, <http://orcid.org/0000-0002-3082-5113>, sobolevski@iitp.ru
- Li Da Xu** Academician at the European Academy of Sciences, Russian Academy of Engineering (formerly, USSR Academy of Engineering), and Armenian Academy of Engineering, Dr. Sci. (Systems Science), Professor and Eminent Scholar in Information Technology and Decision Sciences, Old Dominion University, Norfolk, VA, the United States of America. Scopus Author ID 13408889400, <https://orcid.org/0000-0002-5954-5115>, lxu@odu.edu
- Yury S. Kharin** Academician at the National Academy of Sciences of Belarus, Dr. Sci. (Phys.-Math.), Professor, Director of the Institute of Applied Problems of Mathematics and Informatics of the Belarusian State University, Minsk, Belarus. Scopus Author ID 6603832008, <http://orcid.org/0000-0003-4226-2546>, kharin@bsu.by
- Yuri A. Chaplygin** Academician at the RAS, Dr. Sci. (Eng.), Professor, Member of the Departments of Nanotechnology and Information Technology of the RAS, President of the National Research University of Electronic Technology (MIET), Moscow, Russia. Scopus Author ID 6603797878, ResearcherID B-3188-2016, president@miet.ru
- Vasilii V. Shpak** Cand. Sci. (Econ.), Deputy Minister of Industry and Trade of the Russian Federation, Ministry of Industry and Trade of the Russian Federation, Moscow, Russia; Associate Professor, National Research University of Electronic Technology (MIET), Moscow, Russia, mishinevaiv@minprom.gov.ru

Содержание

Информационные системы. Информатика. Проблемы информационной безопасности

- 7** *А.С. Боронников, П.С. Цынгалёв, В.Г. Ильин, Т.А. Деменкова*
Оценка эффективности балансировщика соединений *PgBouncer* для оптимизации вычислительных ресурсов реляционных баз данных
- 25** *С.И. Смирнов, М.А. Еремеев, Ш.Г. Магомедов, Д.А. Изергин*
Критерии и показатели оценивания качества проведения расследования инцидента информационной безопасности при целевой кибератаке
- 37** *И.Е. Тарасов, П.Н. Советов, Д.В. Люлява, Д.И. Мирзоян*
Методика проектирования специализированных вычислительных систем на основе совместной оптимизации аппаратного и программного обеспечения

Современные радиотехнические и телекоммуникационные системы

- 46** *А. Андрун, О.В. Тихонова*
Влияние шумов и помех на вероятность битовых ошибок в системах 5G, использующих банк фильтров с несколькими несущими со смещенной квадратурной амплитудной модуляцией

Микро- и нанoeлектроника. Физика конденсированного состояния

- 55** *М.А. Мухутдинова, А.Н. Юрасов*
Моделирование магниторефрактивного эффекта в нанокompозитах $\text{Co-Al}_2\text{O}_3$ в рамках приближения Бруггемана

Математическое моделирование

- 65** *И.В. Артюхин*
Двумерный алгоритм с последовательной оценкой углов прихода сигналов в системе когерентных распределенных автомобильных радаров с несколькими приемными и передающими антеннами
- 78** *Ю.П. Перова, С.А. Лесько, А.А. Иванов*
Анализ и прогнозирование динамики настроек пользователей интернет-ресурсов на основе уравнения Фоккера – Планка
- 93** *У.С. Мохнаткина, Д.В. Парфенов, Д.А. Петрусевич*
Анализ подходов к определению тренда в структуре временного ряда

Contents

Information systems. Computer sciences. Issues of information security

- 7** Anton S. Boronnikov, Pavel S. Tsyngalev, Victor G. Ilyin, Tatiana A. Demenkova
Evaluation of connection pool *PgBouncer* efficiency for optimizing relational database computing resources
- 25** Stanislav I. Smirnov, Mikhail A. Ereemeev, Shamil G. Magomedov, Dmitry A. Izergin
Criteria and indicators for assessing the quality of the investigation of an information security incident as part of a targeted cyberattack
- 37** Ilya E. Tarasov, Peter N. Sovietov, Daniil V. Lulyava, Dmitry I. Mirzoyan
Method for designing specialized computing systems based on hardware and software co-optimization

Modern radio engineering and telecommunication systems

- 46** Abed Androon, Olga V. Tikhonova
Impacts of noise and interference on the bit error rate of the FBMC-OQAM modulation scheme in 5G systems

Micro- and nanoelectronics. Condensed matter physics

- 55** Muza A. Mukhutdinova, Alexey N. Yurasov
Modeling of the magnetorefractive effect in Co-Al₂O₃ nanocomposites in the framework of the Bruggeman approximation

Mathematical modeling

- 65** Igor V. Artyukhin
High-resolution 2D-DoA sequential algorithm of azimuth and elevation estimation in automotive distributed system of coherent MIMO radars
- 78** Julia P. Perova, Sergey A. Lesko, Andrey A. Ivanov
Analyzing and forecasting the dynamics of Internet resource user sentiments based on the Fokker–Planck equation
- 93** Ulyana S. Mokhnatkina, Denis V. Parfenov, Denis A. Petrusevich
Analysis of approaches to identification of trend in the structure of the time series

УДК 004.657

<https://doi.org/10.32362/2500-316X-2024-12-3-7-24>

EDN BNQNDI



НАУЧНАЯ СТАТЬЯ

Оценка эффективности балансировщика соединений *PgBouncer* для оптимизации вычислительных ресурсов реляционных баз данных

А.С. Боронников[@],
П.С. Цынгалёв,
В.Г. Ильин,
Т.А. Деменкова

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[@] Автор для переписки, e-mail: boronnikov-anton@mail.ru

Резюме

Цели. Целью работы является исследование возможностей использования балансировщика подключений *PgBouncer* с различными конфигурациями в современных инсталляциях баз данных (БД) путем проведения нагрузочного тестирования с различными сценариями, максимально приближенными к реальной нагрузке, определение критичных показателей, получение результатов тестирования и интерпретация их в виде графиков.

Методы. В ходе исследования использовались методы эксперимента, индукции, тестирования и статистического анализа.

Результаты. Рассмотрены основные возможности, архитектура и режимы работы сервиса *PgBouncer*. Проведено нагрузочное тестирование на виртуальной машине, развернутой на базе открытой облачной платформы, с различной конфигурацией затрачиваемых вычислительных ресурсов – центрального процессора (CPU), оперативной памяти (RAM) и использованием нескольких сценариев с разной конфигурацией и разным количеством подключений балансировщика к БД. В ходе тестирования были исследованы основные показатели: распределение использования процессора, утилизация оперативной памяти, дискового пространства и центрального процессора. Выполнены интерпретация полученных данных и анализ полученных результатов путем выделения критических параметров. Сформулированы выводы и рекомендации по использованию балансировщика подключения в реальных высоконагруженных инсталляциях для оптимизации утилизируемых ресурсов сервером, на котором расположена система управления базами данных (СУБД). Сформировано заключение о полезности использования балансировщика запросов *PgBouncer* и предложены варианты конфигурации для последующего использования в реальных инсталляциях.

Выводы. Исследована степень влияния использования балансировщика соединений *PgBouncer* на производительность системы в целом, развернутой в виртуализированной среде. Результаты работы показали, что применение *PgBouncer* позволяет существенно оптимизировать затрачиваемые вычислительные ресурсы вычислительного узла под сервер СУБД, а именно: уменьшилась нагрузка на CPU на 15%, на RAM – на 25–50%, на дисковую подсистему – на 20%, в зависимости от сценариев тестов, количества подключений к БД, конфигурации балансировщика подключений.

Ключевые слова: PgBouncer, PostgreSQL, пуллер, балансировщик, база данных, оптимизация, мониторинг, виртуальная машина, облачные технологии

• Поступила: 13.06.2023 • Доработана: 06.12.2023 • Принята к опубликованию: 09.04.2024

Для цитирования: Боронников А.С., Цынгалёв П.С., Ильин В.Г., Деменкова Т.А. Оценка эффективности балансировщика соединений *PgBouncer* для оптимизации вычислительных ресурсов реляционных баз данных. *Russ. Technol. J.* 2024;12(3):7–24. <https://doi.org/10.32362/2500-316X-2024-12-3-7-24>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Evaluation of connection pool *PgBouncer* efficiency for optimizing relational database computing resources

Anton S. Boronnikov [®],
Pavel S. Tsyngalev,
Victor G. Ilyin,
Tatiana A. Demenkova

MIREA – Russian Technological University, Moscow, 119454 Russia

[®] *Corresponding author, e-mail: boronnikov-anton@mail.ru*

Abstract

Objectives. The aim of the research is to investigate the possibilities of using the *PgBouncer* connection pool with various configurations in modern database installations by conducting load testing with diverse real-world like scenarios, identifying critical metrics, obtaining testing results, and interpreting them in the form of graphs.

Methods. The research utilized methods of experimentation, induction, testing, and statistical analysis.

Results. The main features, architecture and modes of operation of the *PgBouncer* service are considered. Load testing was carried out on a virtual machine deployed on the basis of an open cloud platform with different configurations of computing resources (CPU, RAM) and according to several scenarios with different configurations and different numbers of balancer connections to the database, during which the following main indicators were investigated: distribution of processor usage, utilization of RAM, disk space, and CPU. The interpretation of the data obtained and the analysis of the results obtained by highlighting critical parameters are performed. On the basis of results analysis, conclusions and recommendations are formulated on the use of a connection balancer in real high-load installations for optimizing the resources utilized by the server on which the database management system (DBMS) is located. A conclusion is presented on the usefulness of using the *PgBouncer* query balancer along with proposed configuration options for subsequent use in real installations.

Conclusions. The degree of influence of the use of the *PgBouncer* connection balancer on the performance of the system as a whole deployed in a virtualized environment is investigated. The results of the work showed that the use of *PgBouncer* allows significantly optimization of the computing resources of a computing node for a DBMS server, namely, load on the CPU decreased by 15%, RAM—by 25–50%, disk subsystem—by 20%, depending on the test scenarios, the number of connections to the database, and the configuration of the connection balancer.

Keywords: *PgBouncer*, PostgreSQL, connection pool, balancer, databases, optimization, monitoring, virtual machines, cloud technologies

• Submitted: 13.06.2023 • Revised: 06.12.2023 • Accepted: 09.04.2024

For citation: Boronnikov A.S., Tsyngalev P.S., Ilyin V.G., Demenkova T.A. Evaluation of connection pool PgBouncer efficiency for optimizing relational database computing resources. *Russ. Technol. J.* 2024;12(3):7–24. <https://doi.org/10.32362/2500-316X-2024-12-3-7-24>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Базы данных (БД) являются неотъемлемой частью современных приложений. Они используются для хранения, управления и обработки большого объема информации. Одной из главных проблем, с которыми сталкиваются приложения, является управление множественными подключениями к БД.

Подключение к БД – это процесс установления связи между клиентским приложением и сервером БД. Каждое клиентское приложение устанавливает свое подключение к БД. Это может привести к излишней нагрузке на сервер БД и снижению производительности приложения. Кроме того, каждое подключение к БД требует определенных ресурсов, таких как память и процессорное время. Если значительное количество клиентских приложений одновременно устанавливает подключение к БД, то это может привести к перегрузке сервера и снижению производительности приложения.

Для решения проблемы можно оптимизировать саму систему управления базами данных (СУБД) путем конфигурирования ее параметров на этапе запуска инфраструктуры [1–3] или же использовать сторонние сервисы – балансировщики подключений в БД. Они позволяют управлять подключениями клиентов, чтобы максимально использовать ресурсы сервера БД, тем самым повышая производительность приложения. Существует несколько типов балансировщиков [4]. В данной статье рассматриваются основные возможности инструмента, который относится к типу балансировки подключений к БД на прикладном уровне (application-level balancing) – *connection pool* (набор, пул соединений). В русскоязычной литературе не имеется точного понятия для обозначения таких агрегаторов запросов к БД, поэтому в рамках данной публикации введен термин «пуллер соединений» или просто «пуллер».

1. ПУТЬ ЗАПРОСОВ К БД

Внедрение пуллера приводит к значительным изменениям в работе с БД. Чтобы их заметить, необходимо предварительно изучить стандартный маршрут прохождения запросов. В обычной архитектуре клиент-серверных соединений (client-server) имеет

место следующий стандартный принцип взаимодействия, изображенный на рис. 1.

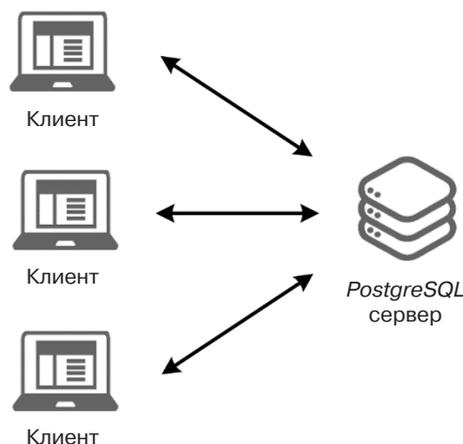


Рис. 1. Обычное клиент-серверное соединение¹

При установлении новой сессии клиентское приложение запрашивает соединение с сервером, и проходит процесс аутентификации. Сервер в ответе создает отдельный системный процесс для обработки соединения и работы сессии. Инициализация состояния сессии осуществляется на основе различных параметров конфигурации, определенных на уровне сервера, БД и пользователя. В рамках одной сессии клиент выполняет требуемые операции. Работа продолжается до тех пор, пока клиент не завершит сессию путем отключения. После завершения сессии сервер уничтожает соответствующий системный процесс, ответственный за обработку данной сессии.

Можно выделить недостатки обычного клиент-серверного соединения:

- 1) создание, управление и удаление процессов соединения занимает время и расходует ресурсы;
- 2) при увеличении числа соединений на сервере возрастает и потребность в ресурсах для их управления. Кроме того, использование памяти на сервере растет с выполнением операций клиентами;
- 3) поскольку одна сессия обслуживает только одного клиента, клиенты могут изменять состояние сессии БД и ожидать, что эти изменения сохранятся в последующих транзакциях.

¹ PostgreSQL. <https://www.postgresql.org/>. Дата обращения 18.04.2023. / Accessed April 18, 2023.

При использовании пуллера клиенты подключаются к нему, а он уже устанавливает соединение с сервером (рис. 2). Это изменяет модель стандартного принципа соединений на клиент-прокси-серверную архитектуру (client-proxy-server).

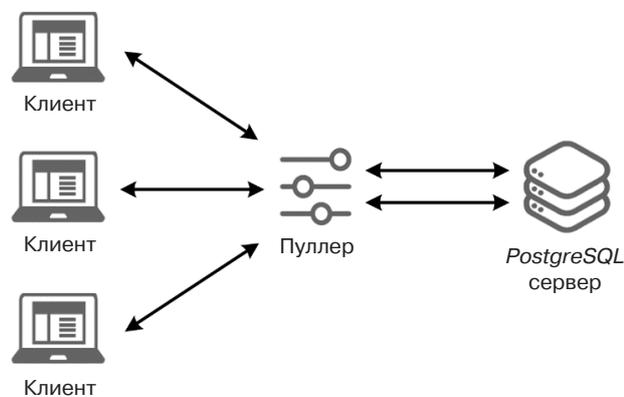


Рис. 2. Соединение «клиент-прокси-сервер»

Теперь подключение клиента к серверу не зависит от времени жизни соединения и процесса на сервере. Пуллер отвечает за принятие и управление соединениями от клиента, установление и поддержание соединений с сервером, а также за назначение серверных соединений клиентским соединениям.

Отличным выбором по качеству системы управления БД является *PostgreSQL* [5–9]. Одним из главных факторов является статус *PostgreSQL* как открытого программного обеспечения (open source). Существует несколько достойных пуллеров соединений с *PostgreSQL* [6, 10–12], таких как *PgBouncer*², *Pgpool-II*³ и *Odyssey*⁴. В данной работе рассмотрены основные возможности, архитектура и режимы работы *PgBouncer*.

2. ПУЛЛЕР СОЕДИНЕНИЙ *PgBouncer*

PgBouncer – это пуллер, позволяющий управлять соединениями с БД *PostgreSQL*. Он работает как прокси-сервер, который обрабатывает запросы на подключение к БД и перенаправляет их на соответствующий сервер. *PgBouncer* может быть установлен на той же машине, что и *PostgreSQL*, либо на отдельной.

Данный пуллер широко используется во многих приложениях на базе *PostgreSQL* и применяется для решения различных задач, связанных с производительностью, масштабируемостью и безопасностью.

Данный балансировщик активно используется в продуктах таких крупных компаний, как *Alibaba*⁵, *Huawei*⁶, *Instagram*⁷ (запрещена в Российской Федерации), *Skype*⁸, в т.ч. на российском рынке⁹ – *Яндекс*¹⁰, *Avito*¹¹, *Сбербанк*¹², *Газпромнефть*¹³ и др.

Одной из главных задач *PgBouncer* является управление соединениями. Он позволяет создавать пулы соединений, которые могут быть использованы несколькими клиентами. Это позволяет снизить нагрузку на сервер БД и повысить производительность приложения.

2.1. Архитектура

В официальной документации *PgBouncer* отсутствует описание архитектуры балансировщика. Был проведен анализ библиотек данного пуллера соединений и исследована его функциональность. На основе собственного реверс-инжиниринга предложена архитектура *PgBouncer*, изображенная на рис. 3.

Слушатель играет важную роль в обработке соединений клиентов с БД *PostgreSQL*. Он обеспечивает точку входа для клиентских подключений (так называемый *socket*) и выполняет роль посредника между клиентом и сервером. Слушатель включает в себя еще *протокол*, который определяет формат обмена данными между клиентом и сервером через *socket*. *PgBouncer* использует тот же протокол, что и *PostgreSQL*, но, помимо этого, имеет собственные расширения и дополнительные команды.

Аутентификация обеспечивает проверку подлинности клиента при попытке подключения к нему. Поддерживаются различные методы, такие как *md5*, *trust*, *plain*, *cert* и др.

⁵ <https://www.alibaba.com/>. Дата обращения 15.04.2023. / Accessed April 15, 2023.

⁶ <https://www.huawei.com/>. Дата обращения 15.04.2023. / Accessed April 15, 2023.

⁷ <https://www.instagram.com/>. Дата обращения 15.04.2023. / Accessed April 15, 2023.

⁸ <https://www.skype.com/ru/> (in Russ.). Дата обращения 15.04.2023. / Accessed April 15, 2023.

⁹ Почему крупнейшие компании России и мира выбирают Postgres. Итоги PgConf.Russia 2017. <http://www.interface.ru/home.asp?artId=39028>. Дата обращения 10.04.2023. [Why the largest companies in Russia and the world choose Postgres. Results of PgConf.Russia 2017. <http://www.interface.ru/home.asp?artId=39028> (in Russ.). Accessed April 10, 2023.]

¹⁰ <https://yandex.ru/> (in Russ.). Дата обращения 15.04.2023. / Accessed April 15, 2023.

¹¹ <https://www.avito.ru/> (in Russ.). Дата обращения 15.04.2023. / Accessed April 15, 2023.

¹² <http://www.sberbank.ru/> (in Russ.). Дата обращения 15.04.2023. / Accessed April 15, 2023.

¹³ <https://www.gazprom-neft.ru/> (in Russ.). Дата обращения 15.04.2023. / Accessed April 15, 2023.

² Официальная документация *PgBouncer* [Official documentation *PgBouncer*]. <https://www.pgouncer.org/>. Дата обращения 02.04.2023. / Accessed April 02, 2023.

³ *Pgpool* Wiki. https://pgpool.net/mediawiki/index.php/Main_Page. Дата обращения 15.04.2023. / Accessed April 15, 2023.

⁴ *Odyssey* – Yandex Technologies. <https://yandex.ru/dev/odyssey/> (in Russ.). Дата обращения 15.04.2023. / Accessed April 15, 2023.

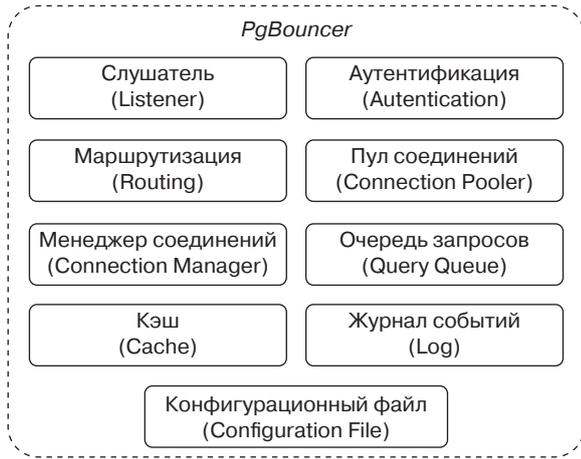


Рис. 3. Архитектура PgBouncer

Маршрутизация определяет режимы работы пуллера. Всего выделяют четыре типа: пул сеансов (session pooling), пул транзакций (transaction pooling), пул операторов (statement pooling) и комбинированный метод (combined pooling). Они подробнее рассматриваются в разделе 2.2.

Пул соединений представляет собой набор доступных соединений, который может быть использован клиентом для выполнения операций в БД. Если соединение свободно, то клиент может взять его из пула. Если же соединения нет, пуллер может создать новое.

Менеджер соединений отвечает за управление жизненным циклом каждого соединения в пуле. Он отслеживает состояние каждого активного соединения к БД, включая его открытие, закрытие и повторное использование свободных соединений.

Очередь запросов управляет запросами, поступающими от клиентов, когда все соединения заняты. Это позволяет обеспечить справедливую обработку запросов и избежать блокировок или перегрузки БД. Очередь запросов имеет настраиваемые параметры, которые позволяют контролировать количество и время ожидания запросов в очереди.

Кэш представляет собой дополнительный компонент, который хранит результаты предыдущих запросов, чтобы при повторном выполнении запроса не обращаться к БД, а возвращать результат непосредственно из кэша. Кэш позволяет сэкономить время и ресурсы на выполнение запросов, особенно если они являются часто повторяющимися и их результаты не изменяются.

Журнал событий ведет запись событий, таких как установление и разрыв соединений, выполнение запросов и других операций и т.д. Он позволяет анализировать и отслеживать работу пуллера, обнаруживать различные ошибки и предупреждения, мониторить производительность и проводить отладку.

Далее следует непосредственно *конфигурационный файл*, в котором можно настроить и указать

параметры всех компонентов архитектуры, рассмотренных в этом разделе.

При правильной настройке этих элементов обеспечивается эффективное использование ресурсов сервера, повышается производительность и улучшается работа приложений, использующих PostgreSQL.

2.2. Режимы работы

Режимы работы определяют, каким образом пуллер будет управлять соединениями. В разных режимах PgBouncer может оказывать различное влияние на производительность и функциональность системы.

Режим сеансов (session mode) – стандартный подход, который является наиболее корректным. Заключается в том, что каждому клиенту назначается одно серверное подключение на протяжении всего времени, пока клиент остается подключенным. При отключении клиента данное подключение к серверу возвращается обратно в пул. Этот метод работы используется по умолчанию. Режим может быть полезен для приложений, имеющих много клиентских запросов, которые не являются частыми, но выполняются в рамках долгих сессий.

Режим транзакций (transaction mode) – клиенту назначается подключение к серверу только на время выполнения транзакции. При обнаружении завершения транзакции PgBouncer возвращает данное подключение обратно в пул. Режим может быть полезен в приложениях, которые имеют много коротких транзакций.

Режим операторов (statement mode) – самый агрессивный подход, который предполагает, что подключение к серверу будет возвращаться в пул сразу после завершения каждого запроса. В этом режиме транзакции с несколькими операторами запрещены, поскольку они не будут работать. Данный режим может быть полезен для приложений, которые имеют много повторяющихся запросов или используют запросы с одинаковой структурой.

Комбинированный режим (combined mode) – этот подход объединяет режимы транзакций и операторов. Для запросов, которые не начинают новую транзакцию, PgBouncer будет использовать режим операторов, а при запросах, которые начинают новую транзакцию – режим транзакций. Этот метод может быть полезен для приложений, которые выполняют множество повторяющихся запросов и требуют выполнения транзакций. Кроме того, он может быть эффективен для приложений, которые имеют большое количество уникальных запросов, но в которых транзакции могут повторяться.

Проведено сравнение режимов работы PgBouncer по следующим критериям: изоляция транзакций, пул соединений и производительность. Результаты исследования отображены в табл. 1.

Таблица 1. Сравнение режимов работы *PgBouncer*

Режимы работы <i>PgBouncer</i>	Изоляция транзакций	Пул соединений	Производительность
Режим сеансов	Полная	Назначается на время сеанса	Снижение из-за создания и удаления подключений
Режим транзакций	Полная	Назначается на время транзакции	Снижение из-за создания и удаления подключений
Режим операторов	Частичная	Назначается на время запроса	Повышение благодаря переиспользованию подключений
Комбинированный режим	Баланс	В зависимости от типа запроса	Баланс между производительностью и изоляцией

2.3. Особенности использования

Необходимо сделать акцент на некоторых особенностях *PgBouncer*, которые впоследствии помогут избежать ошибок при работе с ним.

Первой особенностью является ограничение по типам запросов. Некоторые запросы, такие как создание или удаление БД, не могут быть маршрутизированы через пуллер и должны быть выполнены напрямую на сервере *PostgreSQL*. Также стоит учитывать конфигурацию самой СУБД, поскольку часть параметров, например, связанных с кэшированием, могут повлиять на производительность самого пуллера. В таком случае потребуется дополнительная настройка сервера *PostgreSQL* для оптимальной работы в связке с пуллером.

Для избежания перегрузок пуллера необходимо контролировать размер пула соединений балансирующего, поскольку это может привести к переполнению и исчерпанию ресурсов системы – утилизации центрального процессора (ЦП) и оперативного запоминающего устройства (ОЗУ). Переполненный пул может привести к снижению производительности или сбоям при работе с приложениями.

При настройке расширений необходимо учитывать, что некоторые из них могут быть несовместимы с *PgBouncer*, т.к. они могут создавать свои собственные соединения к серверу *PostgreSQL*, что будет вредить производительности системы в целом. Также при использовании данного балансирующего необходимо убедиться, что он совместим с другими инструментами и технологиями, используемыми в приложении, например, с используемым ORM-фреймворком¹⁴.

Не стоит забывать и о необходимости поддержки версии у СУБД при использовании с *PgBouncer*, поскольку некоторые версии пуллера могут не поддерживать последние версии *PostgreSQL*.

¹⁴ Object relation mapping – технология программирования, связывает базы данных с концепциями объектно-ориентированных языков программирования, создавая виртуальную объектную базу данных. [Object relation mapping is a programming technology that connects DBs with the concepts of object-oriented programming languages, creating a virtual object DB.]

3. ТЕСТИРОВАНИЕ И ОЦЕНКА ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

Тестирование проводилось на двух виртуальных машинах (ВМ) со следующими характеристиками:

- ВМ с СУБД: 8 виртуальных центральных процессоров (vCPU)¹⁵, 16 ГБ оперативной памяти (RAM)¹⁶, внешний накопитель типа SSD 32 ГБ (под систему), 500 Мб/с, 320 IOPS¹⁷, 120 ГБ (под БД), 500 Мб/с, 1200 IOPS, операционная система (ОС) Ubuntu 22.04;
- ВМ с пуллером: 2 vCPU, 4 ГБ RAM, внешний накопитель типа SSD 120 ГБ, 500 Мб/с, 1200 IOPS, ОС Ubuntu 22.04.

Для сбора метрик и их отображения использовалось программное обеспечение *pgwatch2*¹⁸, *Grafana*¹⁹ и *PostgreSQL*, которое было запущено в докер-контейнере [13–15]. Под работоспособность системы выделялось 2 ГБ ОЗУ, 1 vCPU. Также в ходе тестирования увеличивалась производительность ВМ тестового стенда, поскольку некоторые тесты утилизировали все доступные ресурсы.

Тестирование проводилось с использованием нескольких сценариев подключения: напрямую к БД и через пуллер. Пуллер был установлен в режим сеансов. Сами сценарии включали в себя постепенное увеличение числа подключений (100, 500, 1000) и сложность запроса с размером активной сессии 10 мин.

Запросы носили следующий характер:

- простые запросы к пустой БД;
- crud-запросы (создание, чтение, модификация, удаление) с применением временных таблиц к БД, содержащей тестовые данные.

¹⁵ vCPU – virtual central processing unit.

¹⁶ RAM – random access memory.

¹⁷ IOPS – input/output operations per second, количество операций ввода-вывода в секунду.

¹⁸ PGWatch: Optimized PostgreSQL monitoring. <https://pgwatch.com>. Дата обращения 15.04.2023. / Accessed April 15, 2023.

¹⁹ Grafana Labs. <https://grafana.com>. Дата обращения 12.04.2023. / Accessed April 12, 2023.

Для анализа выделялись следующие метрики:

1. Распределение использования ЦП:
 - idle – свободные ресурсы;
 - user – затраты на использование пользователями системы;
 - system – затраты на систему;
 - iowait – ожидание от дисковой подсистемы;
 - other – другие операции ЦП;
 - irqс – прерывания ядра ЦП.
2. Утилизация ЦП.
3. Утилизация ОЗУ.
4. Утилизация дисковой подсистемы (диска).

3.1. Запросы на чтение

В рамках данного тестирования параллельно выполнялся запрос на чтение к БД (получение версии СУБД).

Данный тест был выбран, т.к. позволяет максимально справедливо оценить влияние сессии на ресурсы БД (сильнее всего это отображается на утилизации ОЗУ).

Метрики при прямых подключениях простыми запросами изображены на рис. 4 (100 подключений), рис. 5 (500 подключений), рис. 6 (1000 подключений), где временные отрезки, показанные на графике,

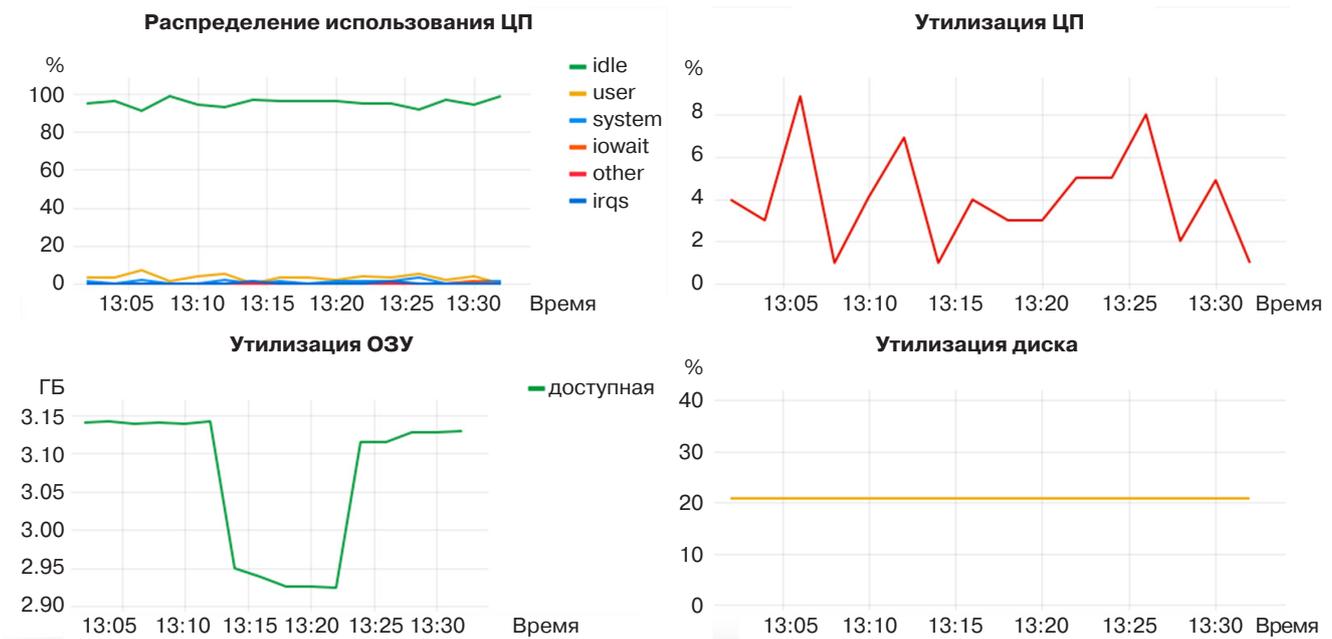


Рис. 4. Проведение нагрузочного тестирования простыми запросами с количеством подключений 100 напрямую к пустой БД

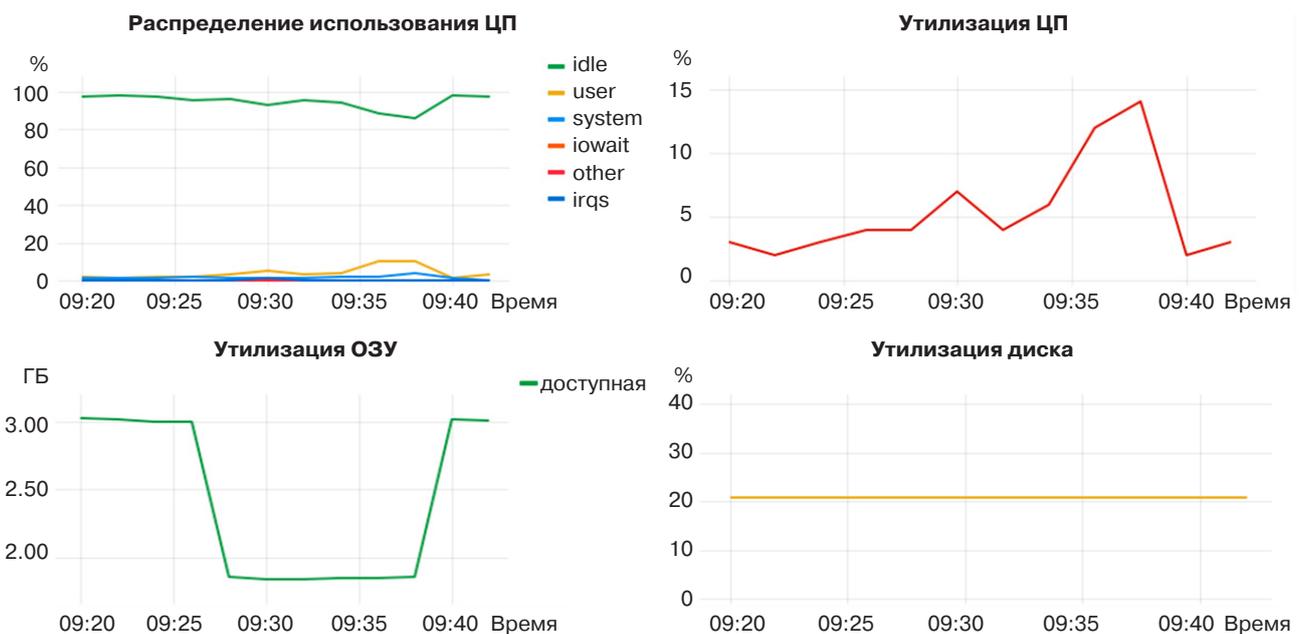


Рис. 5. Проведение нагрузочного тестирования простыми запросами с количеством подключений 500 напрямую к пустой БД

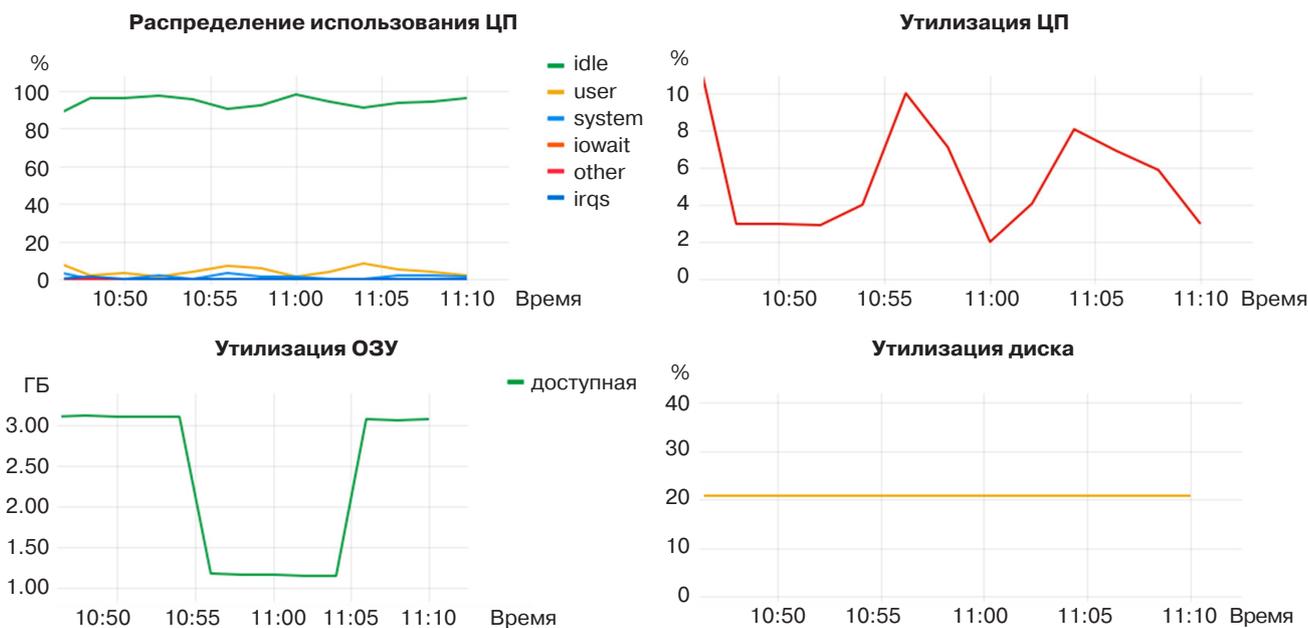


Рис. 6. Проведение нагрузочного тестирования простыми запросами с количеством подключений 1000 напрямую к пустой БД

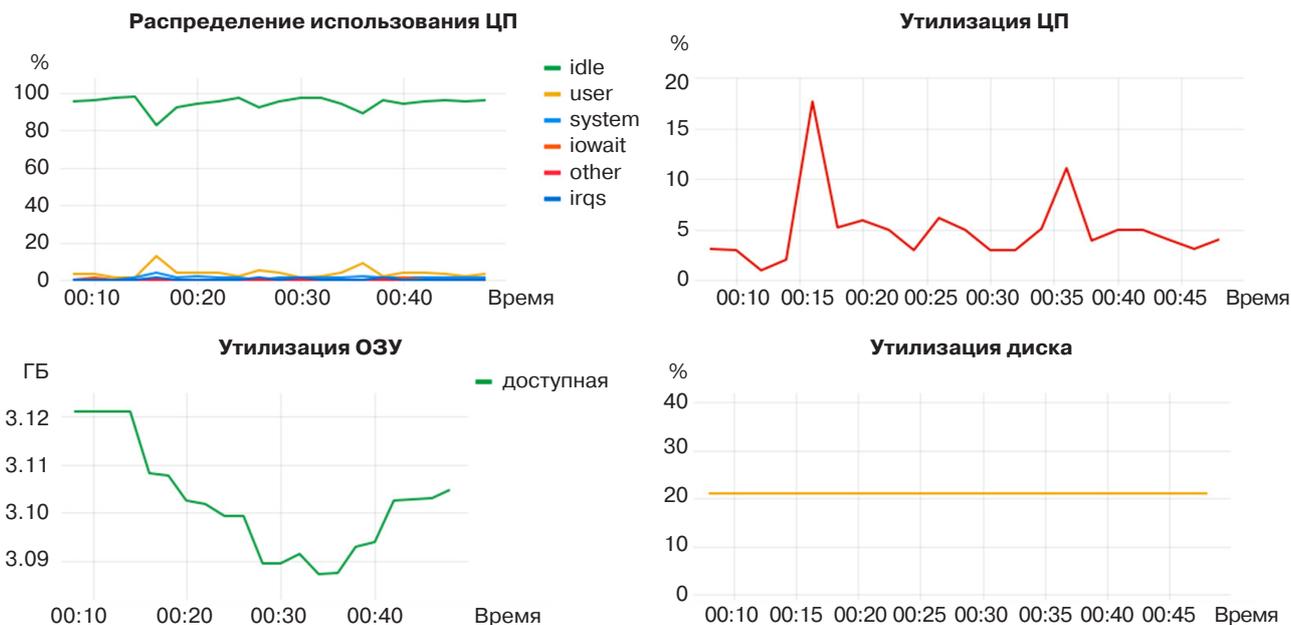


Рис. 7. Проведение нагрузочного тестирования простыми запросами с количеством подключений 100 через пуллер к пустой БД

отражают изменение параметров в период тестирования в реальном времени в формате «часы:минуты».

Метрики при подключениях через пуллер простыми запросами изображены на рис. 7 (100 подключений), рис. 8 (500 подключений), рис. 9 (1000 подключений).

В ходе тестирования ожидаемо было замечено сильное влияние простаивающих подключений на ресурсы, резервируемые под БД.

Если утилизацию ЦП и диска можно списать на погрешность и влияние внешних факторов, то ОЗУ стоит рассмотреть детальнее. Была получена следующая утилизация для тестов:

1. 100 подключений – 230 МБ (2.3 МБ/подключение);
2. 500 подключений – 1180 МБ (2.36 МБ/подключение);
3. 1000 подключений – 1810 МБ (1.81 МБ/подключение).

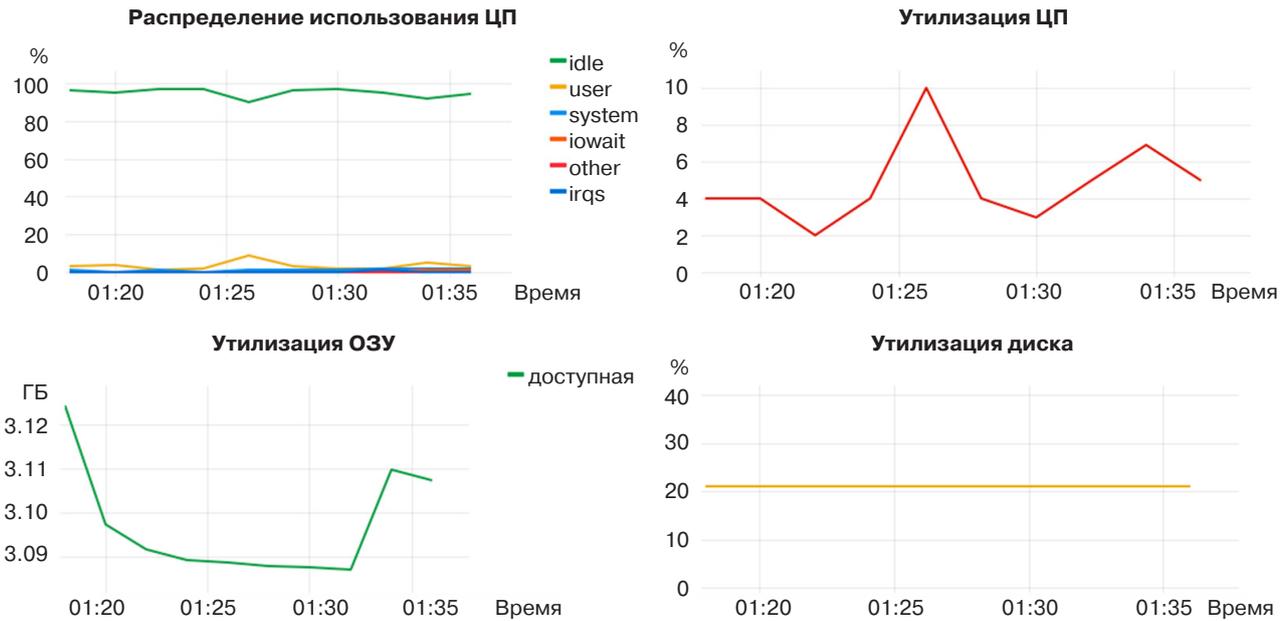


Рис. 8. Проведение нагрузочного тестирования простыми запросами с количеством подключений 500 через пуллер к пустой БД

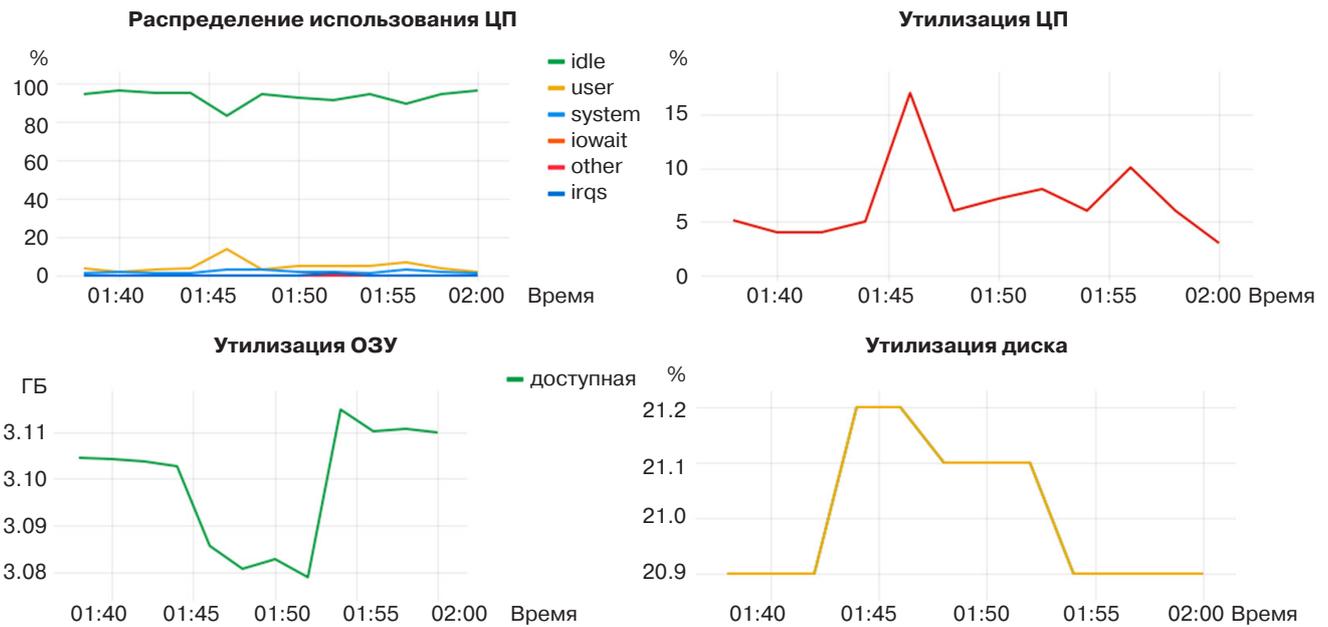


Рис. 9. Проведение нагрузочного тестирования простыми запросами с количеством подключений 1000 через пуллер к пустой БД

Показатели утилизации ОЗУ при тестировании через пуллер были крайне незначительны и не колебались в зависимости от количества подключений, оставаясь на крайне низком уровне (~30 МБ на всех тестах).

Также стоит отметить, что наибольшее влияние на потребляемые ресурсы единичное подключение оказывало при 500 параллельных подключений.

3.2. Усложненные запросы

В рамках данного тестирования параллельно выполнялся запрос на запись к БД, создание, наполнение, удаление таблицы (2 столбца, 1000000 рядов, с типом значения «text»).

Данный тест был выбран, т.к. позволяет оценить возможную экономию ресурсов ЦП и диска при использовании пуллера запросов.

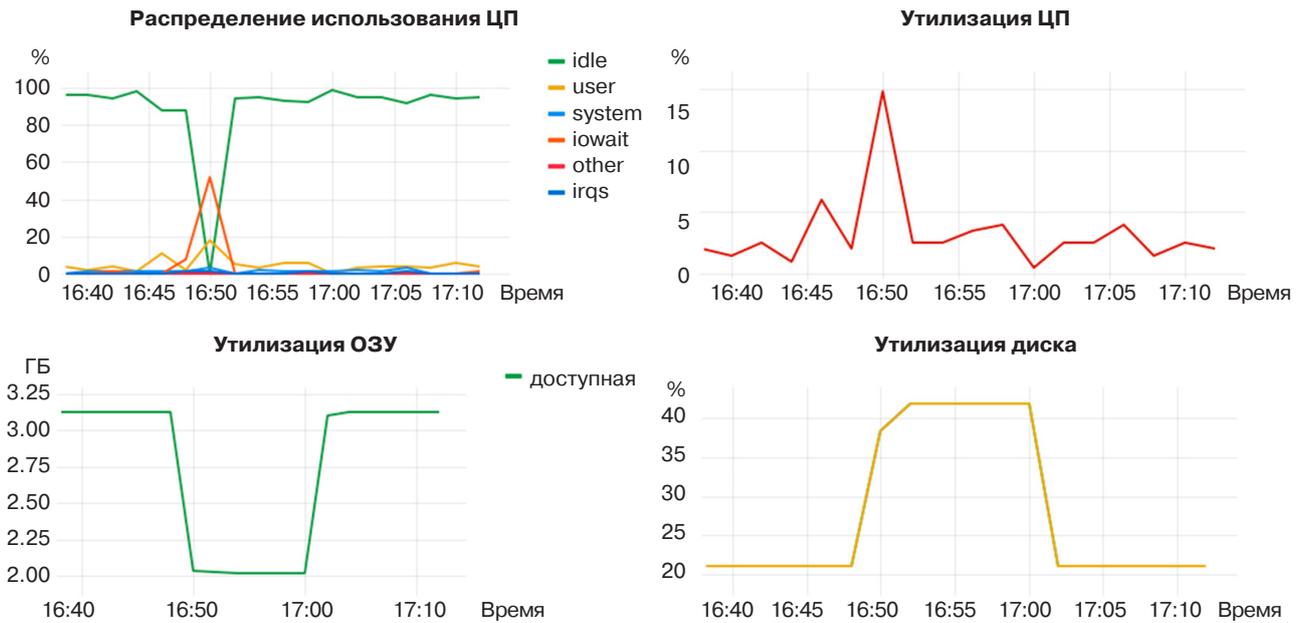


Рис. 10. Проведение нагрузочного тестирования усложненными запросами с использованием временных таблиц и количеством подключений 100 к БД, заполненной тестовыми данными

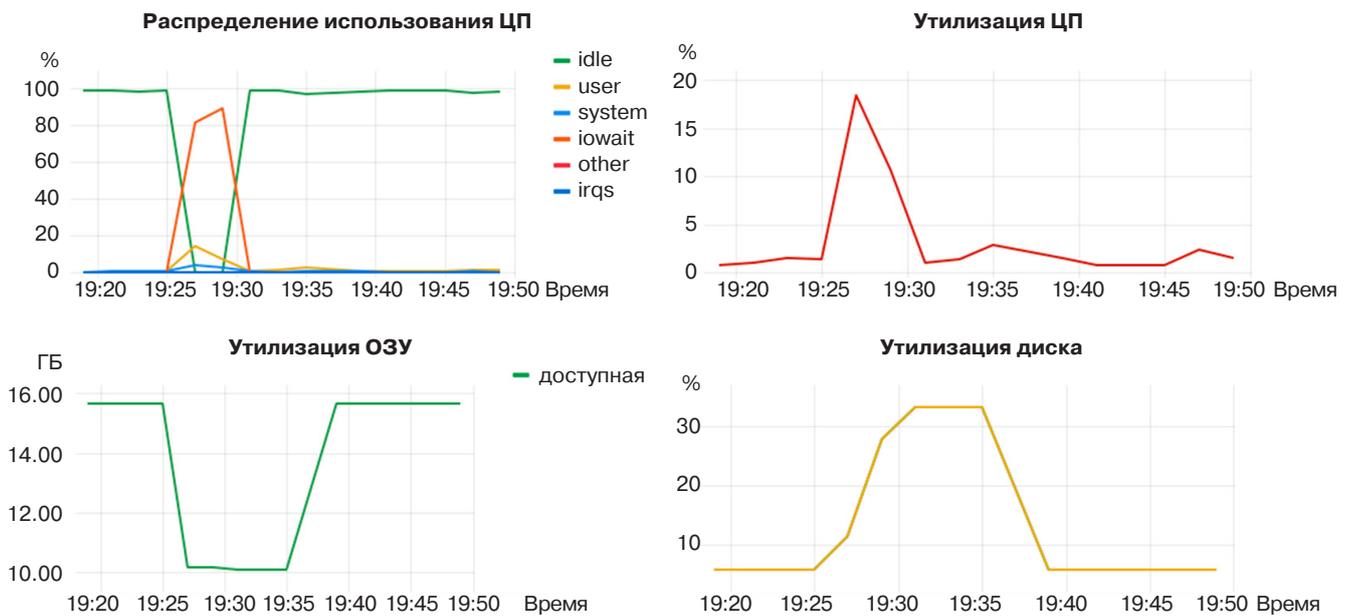


Рис. 11. Проведение нагрузочного тестирования усложненными запросами с использованием временных таблиц и количеством подключений 500 к БД, заполненной тестовыми данными

Метрики при прямых подключениях усложненными запросами изображены на рис. 10 (100 подключений), рис. 11 (500 подключений), рис. 12 (1000 подключений), где временные отрезки, показанные на графике, отражают изменение параметров в период тестирования в реальном времени в формате «часы:минуты».

Метрики при подключениях через пуллер усложненными запросами изображены на рис. 13 (100 подключений), рис. 14 (500 подключений), рис. 15 (1000 подключений).

Тестирование показало большое влияние параллельных операций на ресурсы БД и продемонстрировало возможность минимизировать их средствами пуллера.

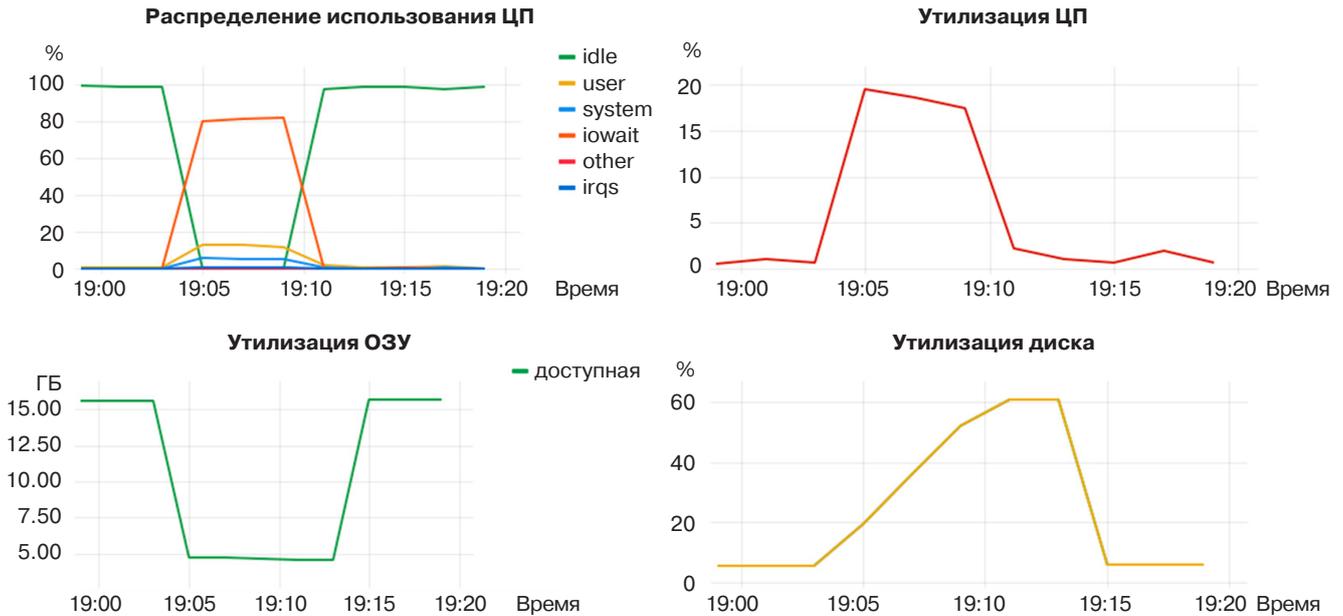


Рис. 12. Проведение нагрузочного тестирования усложненными запросами с использованием временных таблиц и количеством подключений 1000 к БД, заполненной тестовыми данными

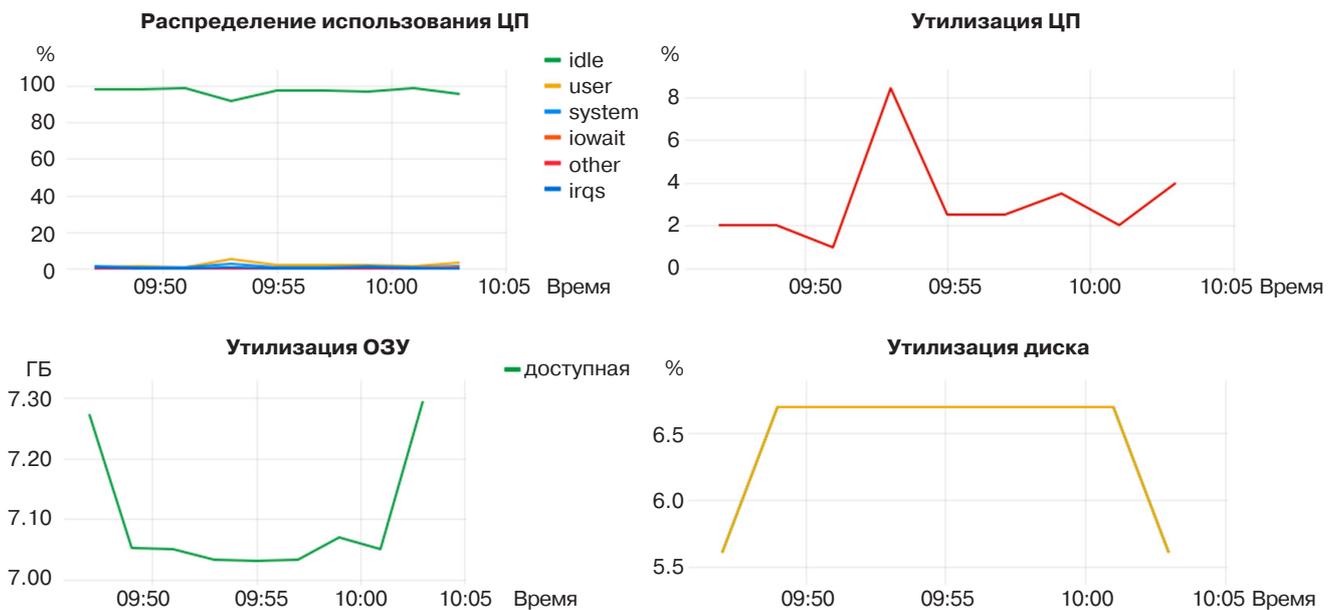


Рис. 13. Проведение нагрузочного тестирования усложненными запросами с использованием временных таблиц и количеством подключений 100 через пуллер к БД, заполненной тестовыми данными

Рассмотрим утилизацию ОЗУ при разном количестве подключений напрямую к БД:

1. 100 подключений – 1125 МБ (11.25 МБ/подключение);
2. 500 подключений – 5900 МБ (11.8 МБ/подключение);
3. 1000 подключений – 10250 МБ (10.25 МБ/подключение).

Показатели утилизации ОЗУ при тестировании через пуллер были крайне незначительны и не колебались в зависимости от количества подключений, оставаясь на крайне низком уровне (~250 МБ на всех тестах).

В сравнении с предыдущим тестом видны скачки утилизации по всем отслеживаемым метрикам. Тест на 500 подключений также, как и в прошлом случае, показывает самую высокую утилизацию памяти

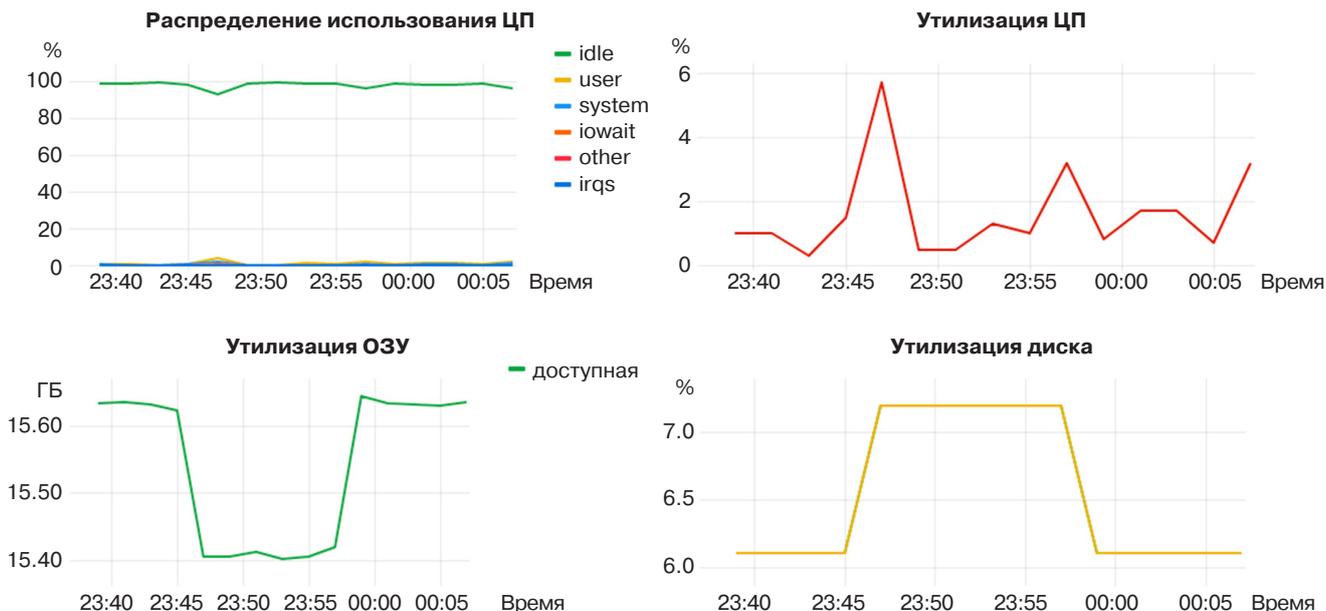


Рис. 14. Проведение нагрузочного тестирования усложненными запросами с использованием временных таблиц и количеством подключений 500 через пуллер к БД, заполненной тестовыми данными

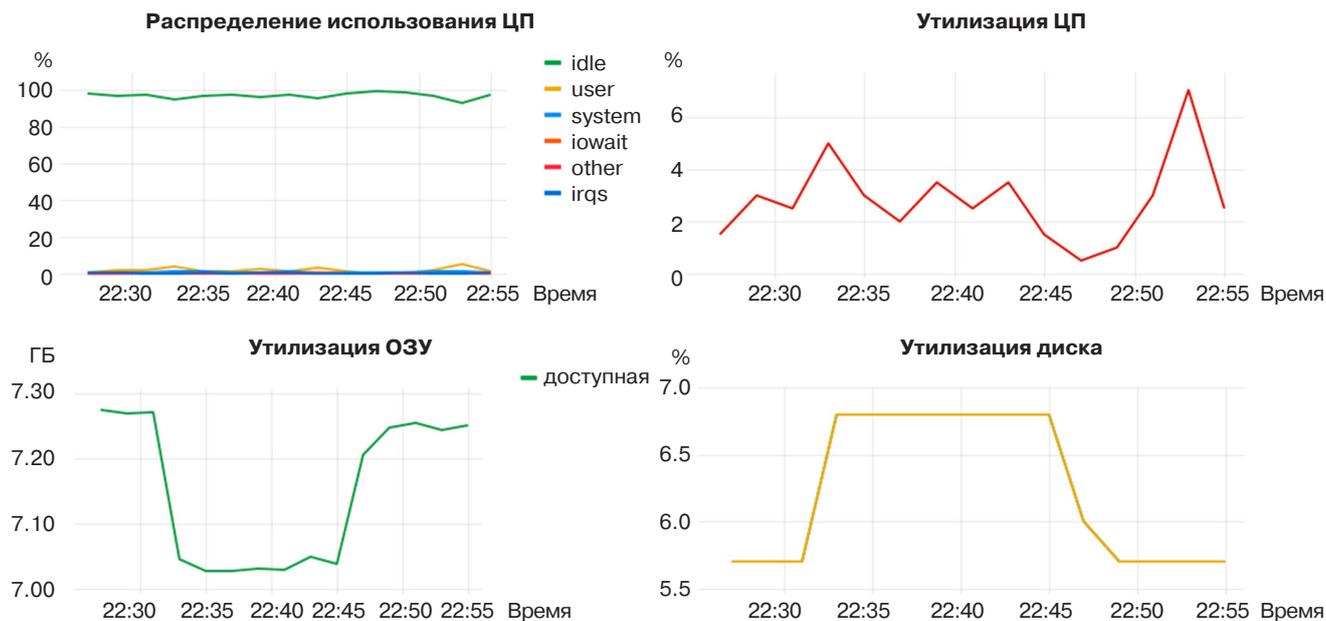


Рис. 15. Проведение нагрузочного тестирования усложненными запросами с использованием временных таблиц и количеством подключений 1000 через пуллер к БД, заполненной тестовыми данными

на единичное подключение. Тест на 1000 подключений показал, что СУБД не ориентирована на такое количество подключений, время повышенной утилизации ЦП и дисковой системы сравнительно выше остальных случаев.

Также можно заметить оптимизацию утилизации дисковой подсистемы. Это связано с возможностью

выноса обработки операций с временными таблицами на мощности пуллера (он заранее знает результат работы всех запросов).

Уменьшение загрузки ЦП при тестировании через пуллер связано с тем, что он берет управление сессиями на себя (самый затратный момент по ресурсам ЦП).

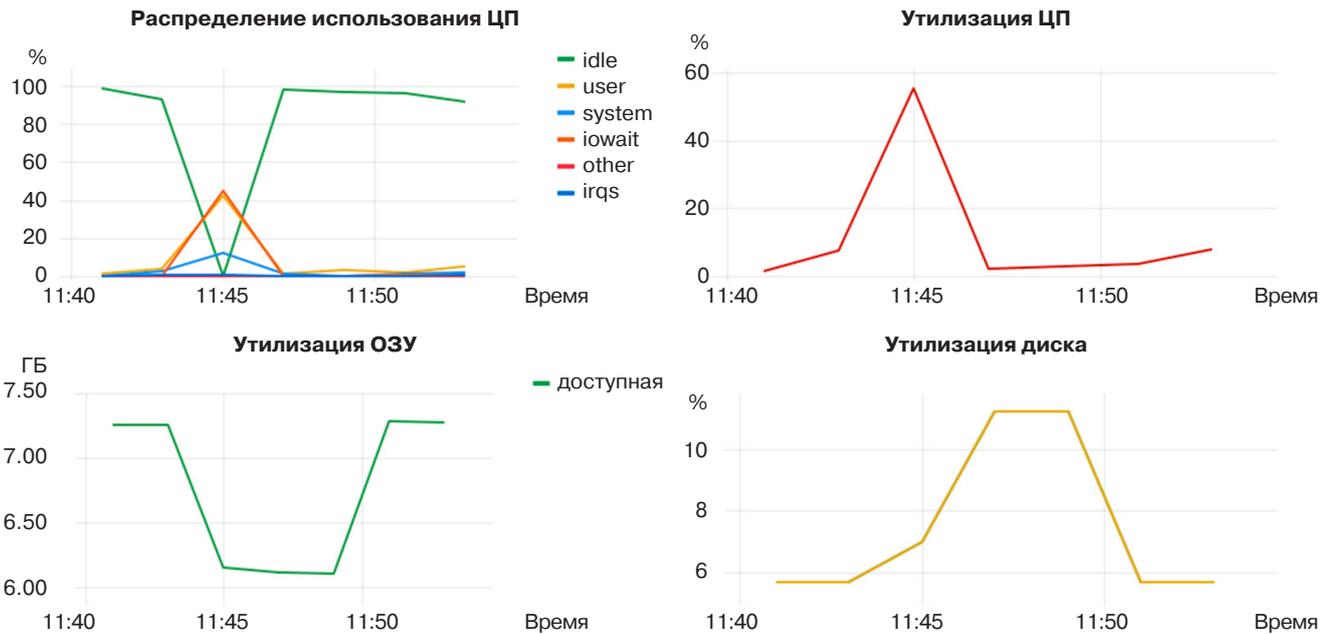


Рис. 16. Проведение нагрузочного тестирования комбинированными запросами с использованием временных таблиц и количеством подключений 100 напрямую к БД

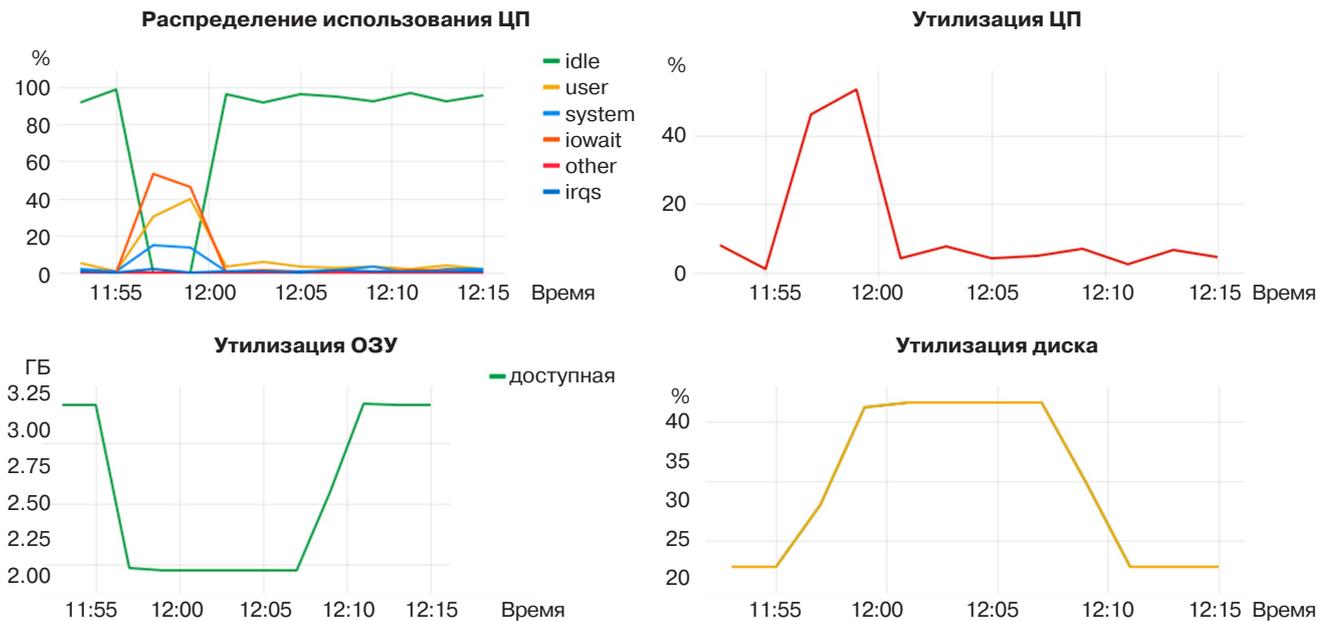


Рис. 17. Проведение нагрузочного тестирования комбинированными запросами с использованием временных таблиц и количеством подключений 500 напрямую к БД

3.3. Комбинированные запросы

Данный тест является комбинацией двух предыдущих. Он направлен на оценку более общего случая использования БД (параллельная запись и чтение).

Метрики при прямых подключениях комбинированными запросами изображены на рис. 16 (100 подключений), рис. 17 (500 подключений),

рис. 18 (1000 подключений), где временные отрезки, показанные на графике, отражают изменение параметров в период тестирования в реальном времени в формате «часы:минуты».

Метрики при подключениях через пуллер комбинированными запросами изображены на рис. 19 (100 подключений), рис. 20 (500 подключений), рис. 21 (1000 подключений).

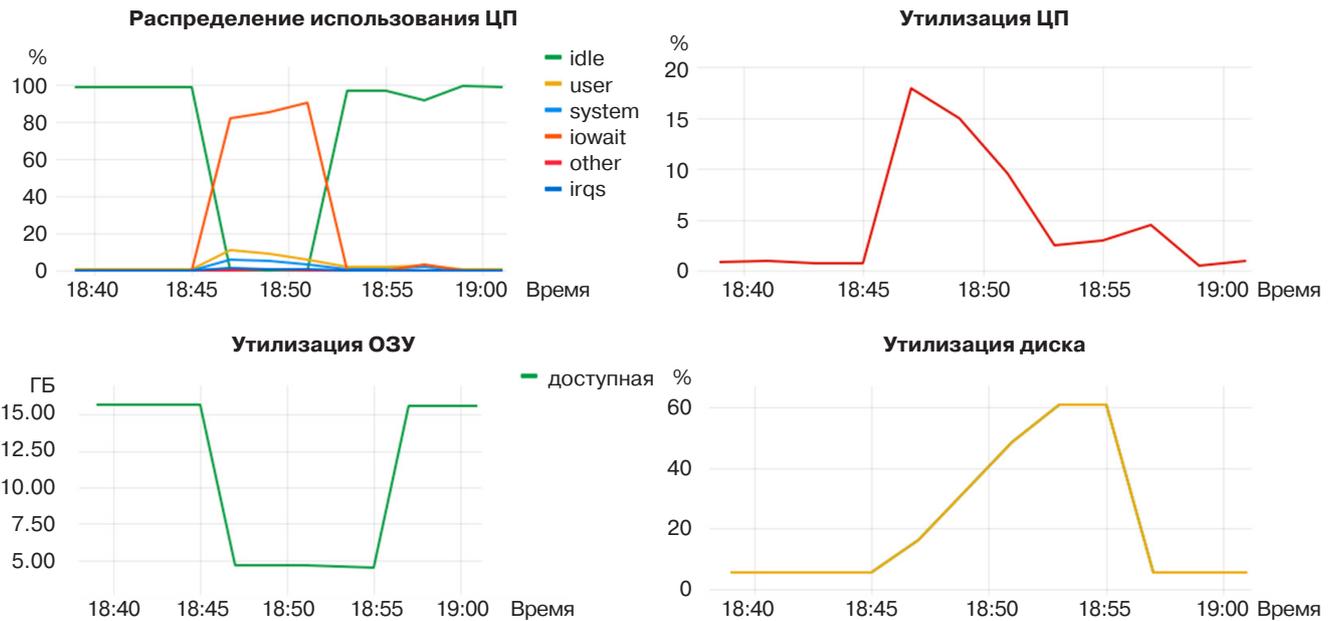


Рис. 18. Проведение нагрузочного тестирования комбинированными запросами с использованием временных таблиц и количеством подключений 1000 напрямую к БД

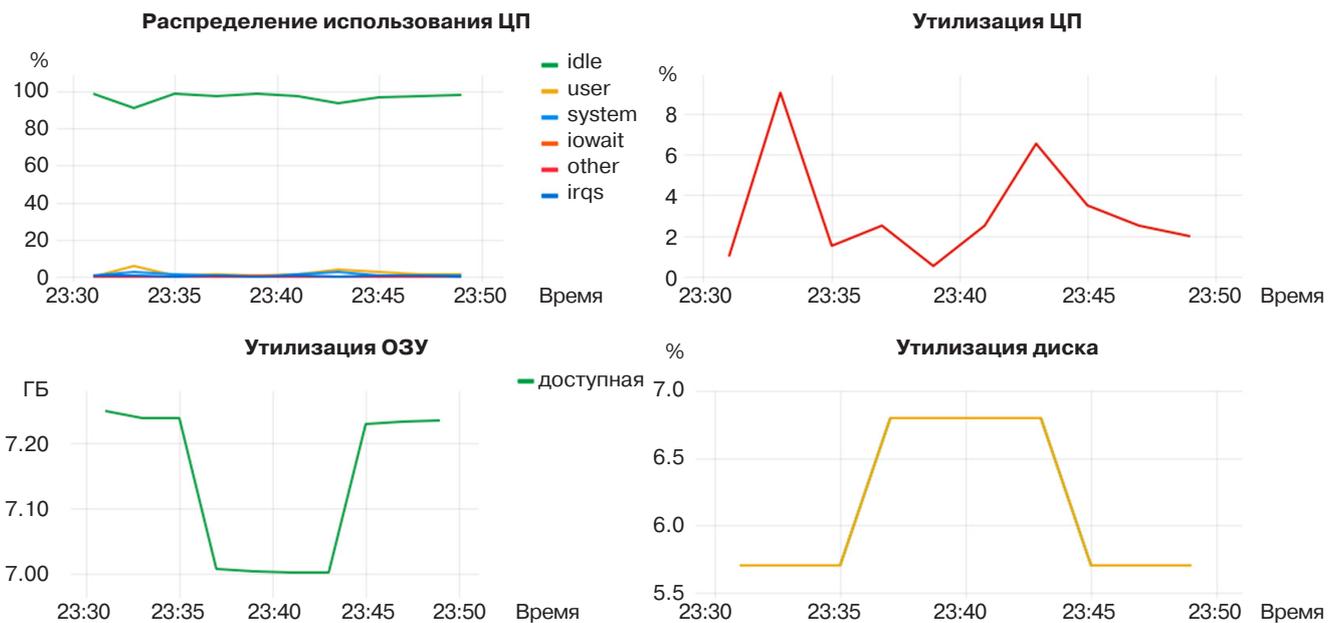


Рис. 19. Проведение нагрузочного тестирования комбинированными запросами с использованием временных таблиц и количеством подключений 100 через пуллер к БД

Тестирование не показало никаких аномалий при проведении тестов на чтение и запись одновременно. Это показывает, что тестирование с более реальными случаями никак не конфликтует с использованием балансировщика подключений.

Рассмотрим утилизацию ОЗУ при разном количестве подключений напрямую к БД:

1. 100 подключений – 1200 МБ (12.0 МБ/подключение);
2. 500 подключений – 6050 МБ (12.1 МБ/подключение);
3. 1000 подключений – 11150 МБ (11.15 МБ/подключение).

Показатели утилизации ОЗУ при тестировании через пуллер были крайне незначительны и не колебались в зависимости от количества подключений, оставаясь на крайне низком уровне (~250 МБ на всех тестах).

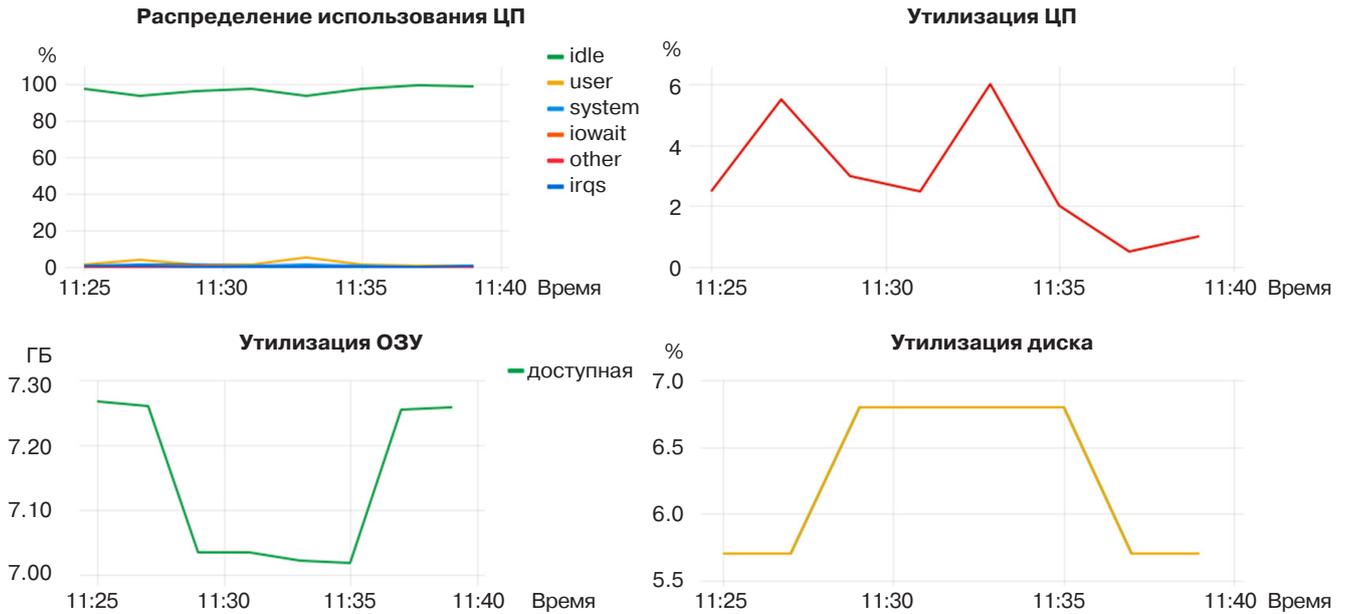


Рис. 20. Проведение нагрузочного тестирования комбинированными запросами с использованием временных таблиц и количеством подключений 500 через пуллер к БД

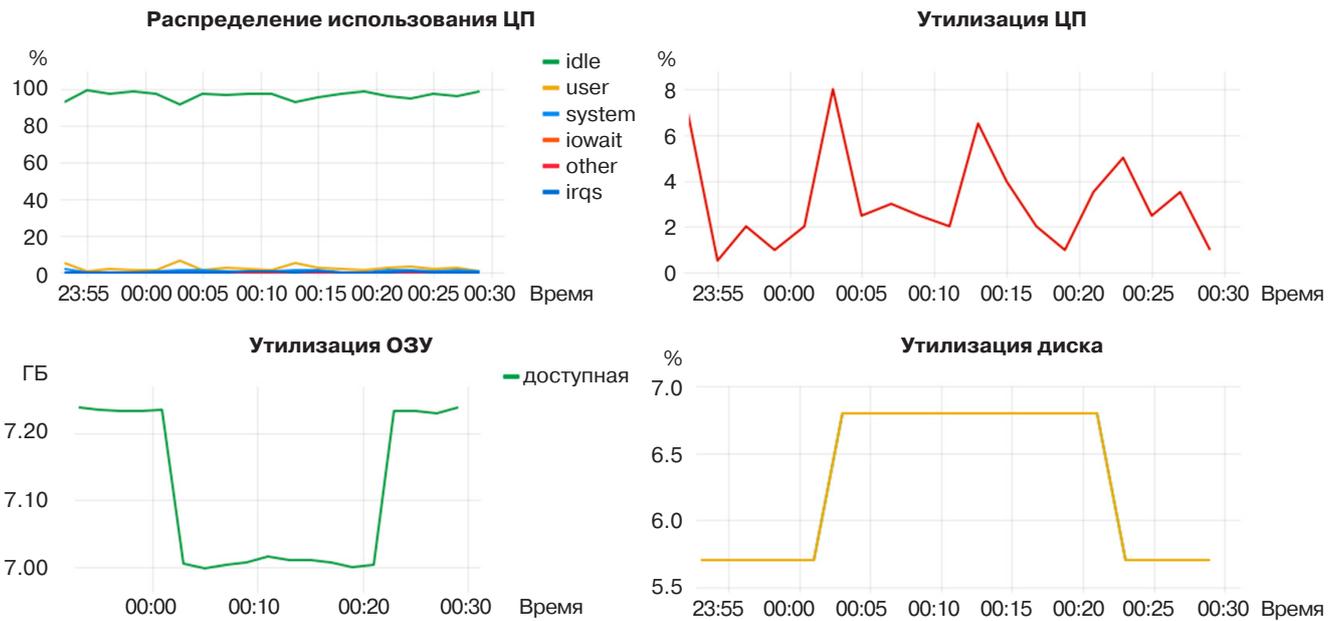


Рис. 21. Проведение нагрузочного тестирования комбинированными запросами с использованием временных таблиц и количеством подключений 1000 через пуллер к БД

3.4. Общая оценка результатов

В целом тестирование показало ожидаемые результаты, а именно снижение нагрузки на систему. Таким образом, использование балансировщика позволит не только снизить затраты на инфраструктуру, но и сильно оптимизировать саму систему, в особенности избежать «огромных» ВМ по несколько сотен гигабайт ОЗУ и ядер ЦП под СУБД. Тестирование показало, что в среднем оптимизация

подключений позволила освободить 25–50% ОЗУ, которые были предназначены для самой СУБД, учитывая размер ВМ, выделяемой под сам пуллер. Если же взять чистые вычисления, без учета ресурсов под сам балансировщик, то разница составила ~30 раз. При этом на разных тестах PgBouncer показывал примерно одни и те же значения, что свидетельствует о некоторой универсальности данного решения в отличии от корректировки параметров самой СУБД.

Таблица 2. Результаты нагрузочного тестирования простыми запросами

Количество запросов	Способ отправки запросов к БД	Утилизация ЦП, %	Утилизация ОЗУ, МБ	Утилизация диска, %
100 запросов	Напрямую к БД	7	230	0
	Через пуллер к БД	14	30	0
500 запросов	Напрямую к БД	11	1100	0
	Через пуллер к БД	8	30	0
1000 запросов	Напрямую к БД	8	1800	0
	Через пуллер к БД	11	20	1.4

Таблица 3. Результаты нагрузочного тестирования усложненными запросами

Количество запросов	Способ отправки запросов к БД	Утилизация ЦП, %	Утилизация ОЗУ, МБ	Утилизация диска, %
100 запросов	Напрямую к БД	25	1125	20
	Через пуллер к БД	8	260	1.5
500 запросов	Напрямую к БД	16	5900	30
	Через пуллер к БД	6	250	1.5
1000 запросов	Напрямую к БД	19	10250	55
	Через пуллер к БД	6	250	1.1

Таблица 4. Результаты нагрузочного тестирования комбинированными запросами

Количество запросов	Способ отправки запросов к БД	Утилизация ЦП, %	Утилизация ОЗУ, МБ	Утилизация диска, %
100 запросов	Напрямую к БД	53	1200	7
	Через пуллер к БД	8	250	1
500 запросов	Напрямую к БД	58	6050	32
	Через пуллер к БД	4	260	1
1000 запросов	Напрямую к БД	18	11150	55
	Через пуллер к БД	7	250	1

Стоит отметить снижение утилизации ресурсов дисковой подсистемы при использовании балансировщиков запросов. Данная оптимизация также позволит снизить затраты и уменьшить параметры диска, предназначенного под БД.

Также было замечено снижение утилизации ЦП при использовании пуллера. В среднем колебание составило 15–20%, что трудно назвать оптимизацией, поскольку нагрузка на ЦП имеет моментальный, пиковый характер, а система мониторинга собирает данные 1 раз в минуту, и, следовательно, могут наблюдаться резкие скачки показаний. Полученные результаты можно учитывать при проектировании системы.

Приведенные на рисунках графики «Распределение использования ЦП» при тестировании отображают информацию о том, что ресурсы непосредственно затрачиваются для работы СУБД *PostgreSQL*, а не для других процессов, например, *iowait*, который возникает при максимальной нагрузке на дисковую подсистему.

Сводные данные проведенного тестирования представлены в табл. 2–4.

ЗАКЛЮЧЕНИЕ

Исследование показало, что программное обеспечение *PgBouncer* представляет собой эффективный инструмент для управления пулом соединений с БД *PostgreSQL*. В ходе проведенного тестирования было выявлено улучшение производительности системы путем снижения затрачиваемых ресурсов на СУБД *PostgreSQL*, а именно нагрузка на ЦП уменьшилась на 15%, на ОЗУ – на 25–50%, на дисковую подсистему – на 20%.

Данный балансировщик имеет гибкую и легкую систему настройки режимов работы, позволяет выбрать наиболее подходящий вариант в зависимости от конкретных потребностей приложений и настроек БД.

Применение *PgBouncer* повышает надежность БД и сокращает время обработки запросов. Это особенно важно для приложений, работающих с большим объемом данных, которые обрабатывают множество запросов одновременно.

Таким образом, можно сделать вывод о том, что *PgBouncer* является полезным инструментом для

управления БД *PostgreSQL* и может быть успешно применен во многих приложениях и инфраструктурах, где требуется высокая производительность, масштабируемость и безопасность.

В дальнейшем планируется изучение способов развертывания данного архитектурного решения на базе российской платформы и проведение нагрузочного тестирования для определения целесообразности переноса систем, а также проведение тестирования в составе кластера высокой доступности.

Вклад авторов

А.С. Боронников – идея исследования, проведение исследования, написание текста статьи, интерпретация и обобщение результатов исследования, научное редактирование статьи.

П.С. Цынгальев – консультации по вопросам проведения исследования, написание текста статьи.

В.Г. Ильин – консультации по вопросам проведения исследования, написание текста статьи.

Т.А. Деменкова – идея исследования, планирование исследования, научное редактирование статьи.

Authors' contributions

A.S. Boronnikov – the research idea, conducting research, writing the text of the article, interpretation and generalization of the results of the research, scientific editing of the article.

P.S. Tsyngalev – consultations on research issues, writing the text of the article.

V.G. Ilyin – consultations on research issues, writing the text of the article.

T.A. Demenkova – the research idea, research planning, scientific editing of the article.

СПИСОК ЛИТЕРАТУРЫ

1. Шараев Е.В. Использование алгоритмических композиций при оптимизации PostgreSQL методами машинного обучения. *Научному Прогрессу – Творчество Молодых*. 2019;3:135–137.
2. Borodin A., Mirvoda S., Porshnev S., Kulikov I. Optimization of Memory Operations in Generalized Search Trees of PostgreSQL. In: Kozielski S., Mrozek D., Kasprowski P., Malysiak-Mrozek B., Kostrzewa D. (Eds.). *Beyond Databases, Architectures and Structures. Towards Efficient Solutions for Data Analysis and Knowledge Representation. BDAS 2017. Communications in Computer and Information Science*. 2017;716:224–232. https://doi.org/10.1007/978-3-319-58274-0_19
3. Варакута П.С., Козлов Р.К. Имитационное моделирование пропускной способности пулов соединений к базе данных PostgreSQL. *Трибуна ученого*. 2022;5:48–53.
4. Мухамедина А., Айдаров А.К. Современные средства балансировки перегрузки. *Научные исследования XXI века*. 2021;2:105–109.
5. Gudilin D.S., Zvonarev A.E., Goryachkin B.S., Lychagin D.A. Relational Database Performance Comparison. In: *Proc. 5th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*. March, 16–18, 2023. Moscow. <https://doi.org/10.1109/REEPE57272.2023.10086872>
6. Тупикина М.А. Сравнение систем управления базами данных SQLite, MySQL и PostgreSQL. В сб.: *Студенческая наука для развития информационного общества: сборник материалов VIII Всероссийской научно-технической конференции*. Часть 2; 22–23 мая 2018 г. Ставрополь: Северо-Кавказский федеральный университет; 2018. С. 345–347.
7. Виноградова М.В., Барашкова Е.С., Березин И.С., Ореликов М.Г., Лузин Д.С. Обзор системы полнотекстового поиска в постреляционной базе данных PostgreSQL. *E-SCIO*. 2020;5(44):754–778.
8. Пантелимонов М.В., Буцацкий Р.А., Жуйков Р.А. Кэширование машинного кода в динамическом компиляторе SQL-запросов для СУБД PostgreSQL. *Труды Института системного программирования РАН*. 2020;32(1):205–220. [https://doi.org/10.15514/ISPRAS-2020-32\(1\)-11](https://doi.org/10.15514/ISPRAS-2020-32(1)-11)
9. Портретов В.С. Сравнение PostgreSQL и MySQL. *Молодежная наука в развитии регионов*. 2017;1:136–139.
10. Chauhan C., Kumar D. *PostgreSQL High Performance Cookbook*. 2nd ed. Birmingham: Packt Publishing; 2017. 360 p.
11. Рогов Е.В. *PostgreSQL 15 изнутри*. М.: ДМК Пресс; 2023. 662 с.
12. Новиков Б.А., Горшкова Е.А., Графеева Н.Г. *Основы технологий баз данных*. 2-е изд. М.: ДМК Пресс; 2020. 582 с.
13. Бойченко А.В., Рогожин Д.К., Корнеев Д.Г. Алгоритм динамического масштабирования реляционных баз данных в облачных средах. *Статистика и Экономика*. 2014;6–2:461–465.
14. Афанасьев Г.И., Абулкасимов М.М., Белоногов И.Б. Методика создания Docker-образа PostgreSQL в среде Ubuntu Linux. *Аллея науки*. 2018;2(1–17):913–918.
15. Smolinski M. Impact of storage space configuration on transaction processing performance for relational database in PostgreSQL. In: Kozielski S., Mrozek D., Kasprowski P., Malysiak-Mrozek B., Kostrzewa D. (Eds.). *Beyond Databases, Architectures and Structures. Towards Efficient Solutions for Data Analysis and Knowledge Representation. BDAS 2017. Communications in Computer and Information Science*. 2018;928:157–167. https://doi.org/10.1007/978-3-319-99987-6_12

REFERENCES

1. Sharaev E.V. Using Algorithmic Compositions in PostgreSQL Optimization with Machine Learning Methods. *Nauchnomu Progressu – Tvorchestvo Molodykh*. 2019;3:135–137 (in Russ.).

2. Borodin A., Mirvoda S., Porshnev S., Kulikov I. Optimization of Memory Operations in Generalized Search Trees of PostgreSQL. In: Kozielski S., Mrozek D., Kasprowski P., Małysiak-Mrozek B., Kostrzewa D. (Eds.). *Beyond Databases, Architectures and Structures. Towards Efficient Solutions for Data Analysis and Knowledge Representation. BDAS 2017. Communications in Computer and Information Science*. 2017;716:224–232. https://doi.org/10.1007/978-3-319-58274-0_19
3. Varakuta P.S., Kozlov R.K. Simulation of the capacity of connection pools to the PostgreSQL database. *Tribuna uchenogo = Tribune of the Scientist*. 2022;5:48–53 (in Russ.).
4. Mukhamedina A., Aidarov A.K. Modern load balancing tools. *Nauchnye issledovaniya 21 veka = Scientific Research of the 21st Century*. 2021;2:105–109 (in Russ.).
5. Gudilin D.S., Zvonarev A.E., Goryachkin B.S., Lychagin D.A. Relational Database Performance Comparison. In: *Proc. 5th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*. March 16–18, 2023. Moscow. <https://doi.org/10.1109/REEPE57272.2023.10086872>
6. Tupikina M.A. Comparison of database management systems SQLite, MySQL and PostgreSQL. In: *Student Science for the Development of the Information Society: collection of materials of the 8th All-Russian Scientific and Technical Conference. Part 2*; May 22–23, 2018. Stavropol: North Caucasian Federal University; 2018. P. 345–347 (in Russ.).
7. Vinogradova M.V., Barashkova E.S., Berezin I.S., Orelikov M.G., Luzin D.S. An overview of the full-text search system in PostgreSQL post-relational database. *E-SCIO*. 2020;5(44):754–778 (in Russ.).
8. Pantilimonov M.V., Buchatskiy R.A., Zhuykov R.A. Machine code caching in PostgreSQL query JIT-compiler. *Trudy Instituta sistemnogo programirovaniya RAN = Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS)*. 2020;32(1):205–220 (in Russ.).
9. Portretov V.S. Comparison of PostgreSQL and MySQL. *Molodezhnaya Nauka v Razvitii Regionov*. 2017;1:136–139 (in Russ.).
10. Chauhan C., Kumar D. *PostgreSQL High Performance Cookbook*. 2nd ed. Birmingham: Packt Publishing; 2017. 360 p.
11. Rogov E.V. *PostgreSQL 15 iznutri (PostgreSQL 15 from the Inside)*. Moscow: DMK Press; 2023. 662 p. (in Russ.).
12. Novikov B.A., Gorshkova E.A., Grafeeva N.G. *Osnovy tekhnologii baz dannykh (Bases of Technologies of Databases)*. 2nd ed. Moscow: DMK Press; 2020. 582 p. (in Russ.).
13. Boichenko A.V., Rogojin D.K., Korneev D.G. Algorithm for dynamic scaling relational database in clouds. *Statistika i Ekonomika = Statistics and Economics*. 2014;6–2:461–465 (in Russ.).
14. Afanas'ev G.I., Abulkasimov M.M., Belonogov I.B. How to create a PostgreSQL Docker image on Ubuntu Linux. *Alleya nauki = Alley of Science*. 2018;2(1–17):913–918 (in Russ.).
15. Smolinski M. Impact of storage space configuration on transaction processing performance for relational database in PostgreSQL. In: Kozielski S., Mrozek D., Kasprowski P., Małysiak-Mrozek B., Kostrzewa D. (Eds.). *Beyond Databases, Architectures and Structures. Towards Efficient Solutions for Data Analysis and Knowledge Representation. BDAS 2017. Communications in Computer and Information Science*. 2018;928:157–167. https://doi.org/10.1007/978-3-319-99987-6_12

Об авторах

Боронников Антон Сергеевич, аспирант, старший преподаватель, кафедра вычислительной техники, Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: boronnikov-anton@mail.ru. SPIN-код РИНЦ 8232-6328, <https://orcid.org/0009-0008-4911-6609>

Цынгалёв Павел Сергеевич, студент, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: pstsingalev@mail.ru. <https://orcid.org/0009-0007-6354-1364>

Ильин Виктор Георгиевич, студент, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: vgilyin@yahoo.com. <https://orcid.org/0009-0001-0304-3052>

Деменкова Татьяна Александровна, к.т.н., доцент, кафедра вычислительной техники, Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: demenkova@mirea.ru. Scopus Author ID 57192958412, ResearcherID AAB-3937-2020, SPIN-код РИНЦ 3424-7489, <https://orcid.org/0000-0003-3519-6683>

About the authors

Anton S. Boronnikov, Postgraduate Student, Senior Lecture, Computer Engineering Department, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: boronnikov-anton@mail.ru. RSCI SPIN-code 8232-6328, <https://orcid.org/0009-0008-4911-6609>

Pavel S. Tsingalev, Student, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: pstsingalev@mail.ru. <https://orcid.org/0009-0007-6354-1364>

Victor G. Ilyin, Student, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: vgilyin@yahoo.com. <https://orcid.org/0009-0001-0304-3052>

Tatiana A. Demenkova, Cand. Sci. (Eng.), Associate Professor, Computer Engineering Department, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: demenkova@mirea.ru. Scopus Author ID 57192958412, ResearcherID AAB-3937-2020, RSCI SPIN-code 3424-7489, <https://orcid.org/0000-0003-3519-6683>

УДК 004.056
<https://doi.org/10.32362/2500-316X-2024-12-3-25-36>
EDN LNWLOK



НАУЧНАЯ СТАТЬЯ

Критерии и показатели оценивания качества проведения расследования инцидента информационной безопасности при целевой кибератаке

С.И. Смирнов[@],
М.А. Еремеев,
Ш.Г. Магомедов,
Д.А. Изергин

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: smirnov_si@mirea.ru

Резюме

Цели. В настоящее время при нарастающем числе целевых атак задача расследования инцидента информационной безопасности (ИБ) приобретает важное значение. Компьютерные криминалисты, в зависимости от имеющихся средств защиты, применяют программные и программно-аппаратные средства форензики, проводят анализ цифровых артефактов различных операционных систем и сетевого трафика с построением хронологии событий (таймлайна) инцидента. На сегодняшний день отсутствует какой-либо формальный подход к оцениванию эффективности действий специалистов при проведении расследования инцидента ИБ в рамках целевой кибератаки. Целью работы является формирование частных показателей оперативности, результативности и ресурсоемкости в рамках критерия пригодности при расследовании инцидента ИБ.

Методы. Используются методы теории эффективности целенаправленных процессов, методы экспертных оценок и теории множеств.

Результаты. Проведен анализ актуальных работ в области расследования компьютерных инцидентов. Представлены терминология и основные руководящие документы спецификации проведения расследования инцидента ИБ. Определены примеры цифровых артефактов в виде классификации. Обоснована целесообразность формирования критериев и показателей оценки качества проведения расследования инцидента ИБ. Выбраны критерий пригодности и следующие показатели оценивания качества проведения расследования: показатель результативности (полноты) выявления цифровых артефактов компьютерным криминалистом на основе проведенных мероприятий, показатель ресурсоемкости и показатель оперативности расследования инцидента ИБ.

Выводы. Полученные результаты могут быть использованы не только руководителями подразделений, но и рядовыми специалистами по ИБ для объективного анализа имеющихся программных и человеческих ресурсов, времени, затраченного на эти мероприятия, и выявленных цифровых артефактов в рамках расследования киберинцидента.

Ключевые слова: инцидент информационной безопасности, целевая кибератака, расследование киберинцидента, матрица MITRE ATT&CK, цифровой артефакт, оценка угроз безопасности информации, целенаправленный процесс, критерий и показатели оценки качества

• Поступила: 02.05.2023 • Доработана: 22.11.2023 • Принята к опубликованию: 05.04.2024

Для цитирования: Смирнов С.И., Еремеев М.А., Магомедов Ш.Г., Изергин Д.А. Критерии и показатели оценивания качества проведения расследования инцидента информационной безопасности при целевой кибератаке. *Russ. Technol. J.* 2024;12(3):25–36. <https://doi.org/10.32362/2500-316X-2024-12-3-25-36>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Criteria and indicators for assessing the quality of the investigation of an information security incident as part of a targeted cyberattack

Stanislav I. Smirnov[®],
Mikhail A. Ereemeev,
Shamil G. Magomedov,
Dmitry A. Izergin

MIREA – Russian Technological University, Moscow, 119454 Russia

[®] Corresponding author, e-mail: smirnov_si@mirea.ru

Abstract

Objectives. The currently increasing number of targeted cyberattacks raises the importance of investigating information security incidents. Depending on the available means of protection, computer forensic experts use software and hardware tools for analyzing digital artifacts of various operating systems and network traffic to create an event chronology (timeline) of the incident. However, to date, there is no formal approach for assessing the effectiveness of expert activities when investigating an information security incident within the framework of a targeted cyberattack. The present study aims to develop partial indicators of promptness, effectiveness, and resource intensity as part of the suitability criterion for investigating an information security incident.

Methods. Methods informed by purposeful process efficiency and set theory are used along with expert evaluation approaches.

Results. An analysis of works in the field of investigation of computer incidents is presented. The terminology and main guiding documents on specifics of conducting information security incident investigations are described along with examples of digital artifacts defined in the form of classification. The expediency of forming criteria and indicators for assessing the quality of an information security incident investigation is substantiated. The suitability criterion and subsequent indicators for assessing the quality of the investigation are selected: the effectiveness (completeness) indicator for detecting digital artifacts by a computer criminologist is based on the conducted activities, resource intensity indicator, and promptness indicator for investigating an information security incident.

Conclusions. The obtained results can be used not only by heads of departments but also by rank-and-file information security professionals for objective analysis of the available software and human resources, the time spent on these activities, and the identified digital artifacts as part of a cyber incident investigation.

Keywords: information security incident, targeted cyberattack, cyber incident investigation, MITRE ATT&CK matrix, digital artifact, information security threat assessment, targeted process, quality assessment criteria and indicators

• Submitted: 02.05.2023 • Revised: 22.11.2023 • Accepted: 05.04.2024

For citation: Smirnov S.I., Ereemeev M.A., Magomedov Sh.G., Izergin D.A. Criteria and indicators for assessing the quality of the investigation of an information security incident as part of a targeted cyberattack. *Russ. Technol. J.* 2024;12(3):25–36. <https://doi.org/10.32362/2500-316X-2024-12-3-25-36>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Геополитические события в мире за последние годы в большой степени повлияли на возросшую активность хакерских группировок. Согласно итоговому отчету за 2022 г. компании Positive Technologies¹ количество инцидентов информационной безопасности (ИБ) за прошедший год увеличилось на 20,8%, среди них объем успешных целенаправленных атак с использованием программ-вымогателей (шифровальщиков) составил 51%. Практика реагирования на современные инциденты ИБ показывает, что в последнее десятилетие количество техник и инструментов, используемых злоумышленниками при целевых атаках на информационно-телекоммуникационные сети организаций, увеличивается ежедневно [1].

В свою очередь этот факт позволяет сделать вывод о важности организация комплексного подхода при расследовании инцидента ИБ. Комплексный подход заключается в агрегации выявленных цифровых артефактов из различных систем защиты (например, SIEM-системы², системы обнаружения/предотвращения вторжений IDS/IPS³, системы предотвращения утечки информации DLP⁴, межсетевые экраны нового поколения NGFW⁵ и др.) с построением хронологии кибератаки. В случае их отсутствия в организации специалистам намного сложнее проводить поиск криминалистически значимых данных, и поэтому увеличивается временной интервал проведения расследования. Важную роль в этом процессе играет наличие «живых» систем, в которых компьютерные криминалисты смогут снять дампы оперативной памяти.

Таким образом, можно говорить о том, что одним из перспективных научных направлений в рамках проведения расследования инцидента ИБ является создание теоретической основы с целью расчета эффективности действий компьютерных криминалистов в зависимости от имеющихся программных средств форензики и систем защиты в организации. Целью данной работы является формирование

частных показателей оперативности, результативности, ресурсоемкости в рамках критерия пригодности при расследовании инцидента ИБ от целевых атак.

Статья продолжает исследования авторов [2, 3], направленные на развитие теоретического базиса при оценивании эффективности целенаправленных процессов.

ТЕРМИНОЛОГИЯ И ОСНОВНЫЕ РУКОВОДЯЩИЕ ДОКУМЕНТЫ СПЕЦИФИКИ ПРОВЕДЕНИЯ РАССЛЕДОВАНИЯ ИНЦИДЕНТА ИБ

В настоящее время существует ряд нормативных документов и инструкций коммерческих организаций, описывающих действия специалистов при расследовании и реагировании на компьютерные атаки. Вместе с тем, в государственных стандартах не определены критерии и показатели оценки качества проведения расследования киберинцидентов. Например, в руководящем документе ГОСТ Р 59709-2022⁶ представлены только термины и определения, а также их взаимосвязи в рамках данных процессов. В документе ГОСТ Р 59712-2022⁷ приведено только организационное описание действий подразделений управления киберинцидентами.

В методическом документе Федеральной службы по техническому и экспортному контролю «Методика оценки угроз безопасности информации»⁸

⁶ ГОСТ Р 59709-2022. Национальный стандарт Российской Федерации. *Защита информации. Обнаружение, предупреждение и ликвидация последствий компьютерных атак и реагирование на компьютерные инциденты. Термины и определения*. М.: Росстандарт; 2022. [GOST R 59709-2022. National Standard of the Russian Federation. *Information protection. Detection, prevention and liquidation of the consequences of computer attacks and response to computer incidents. Terms and Definitions*. Moscow: Rosstandart; 2022 (in Russ.).]

⁷ ГОСТ Р 59712-2022. Национальный стандарт Российской Федерации. *Руководство по планированию и подготовке к реагированию на инциденты ISO/IEC 27035-2*. М.: Росстандарт; 2022. [GOST R 59712-2022. National Standard of the Russian Federation. *Guide to Planning and Prepare for Incident Response ISO/IEC 27035-2*. Moscow: Rosstandart; 2022 (in Russ.).]

⁸ Методический документ «Методика оценки угроз безопасности информации» (утв. ФСТЭК России 05.02.2021 г.). <https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-5-fevralya-2021-g>. Дата обращения 06.04.2023. [The methodological document of the Federal Service for Technical and Export Control “Methodology for Assessing Information Security Threats”. <https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-5-fevralya-2021-g> (in Russ.). Accessed April 06, 2023.]

¹ <https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2022/> (in Russ.). Дата обращения 30.03.2023. / Accessed March 30, 2023.

² SIEM – security information and event management.

³ IDS – intrusion detection system, IPS – intrusion prevention system.

⁴ DLP – data leak/loss prevention.

⁵ NGFW – next-generation firewall.

перечислены десять основных тактик и соответствующие им типовые техники, используемые для построения сценариев реализации угроз безопасности информации. При разработке данного перечня за основу взята матрица MITRE ATT&CK⁹. Согласно ГОСТ Р ИСО/МЭК ТО 18044-2007¹⁰ при обнаружении первых признаков инцидента ИБ перед компьютерными криминалистами появляется задача определения их причин. В связи с этим производится сбор цифровых артефактов, основными источниками которых являются: копии жестких дисков, дампы оперативной памяти, журналы событий безопасности, а также трафик сетевых устройств.

Ниже приведена базовая терминология процесса расследования инцидента ИБ, которая будет необходима при дальнейшем исследовании.

Компьютерная атака – целенаправленное несанкционированное сетевое компьютерное воздействие (или их последовательность) на информационный ресурс, осуществляемое нарушителем с применением программных и (или) программно-аппаратных средств и информационных технологий в целях реализации попыток нарушения и (или) прекращения функционирования информационного ресурса или реализации угрозы безопасности информации, обрабатываемой таким ресурсом¹¹.

Целевая кибератака – это непрерывный процесс несанкционированной активности в инфраструктуре атакуемой информационной системы (ИС), удаленно управляемый вручную в реальном времени¹².

Инцидент ИБ – это появление одного или нескольких нежелательных событий ИБ, которые могут вызвать сбой или нарушение функционирования ИС¹³.

Согласно ГОСТ Р ИСО/МЭК 13335-1-2006¹⁴ примерами киберинцидентов могут являться: утрата оборудования (устройств), ошибки пользователей системы, несоблюдение политики ИБ или рекомендаций, нарушение физических мер защиты, сбой программного обеспечения (ПО), отказы технических средств, системные сбои или перегрузки, нарушение правил доступа.

Расследование инцидента ИБ – это совокупность действий специалистов ИБ, направленных на выявление вектора целевой кибератаки, с целью минимизации ущерба и разработки рекомендаций для предотвращения инцидента ИБ в будущем [2].

Реагирование на компьютерный инцидент – это процесс (процедура, функция) автоматической (автоматизированной) обработки компьютерного инцидента¹⁵.

Цифровой артефакт – это потенциальное доказательство, обнаруженное на целевом устройстве (например, на персональном компьютере, мобильном устройстве, сетевом устройстве), которое может использоваться в судебной практике [3].

Криминалистически значимые данные – это компьютерная информация, используемая для обоснования выводов криминалистических исследований и позволяющая решить задачи, поставленные перед криминалистическим исследованием¹⁶.

Классификация цифровых артефактов, которые могут быть использованы специалистами при расследовании инцидента ИБ, представлена на рисунке.

В классификации авторами определены значимые доказательства распространенных операционных систем (ОС) и сетевого трафика.

⁹ <https://attack.mitre.org/>. Дата обращения 06.04.2023. / Accessed April 06, 2023.

¹⁰ ГОСТ Р ИСО/МЭК ТО 18044-2007. Национальный стандарт Российской Федерации. *Информационные технологии. Методы и средства обеспечения безопасности. Менеджмент инцидентов ИБ*. М.: Стандартинформ; 2007. [GOST R ISO/IEC TO 18044-2007. National Standard of the Russian Federation. *Information technologies. Methods and means of ensuring security. Information security incident management*. Moscow: Standartinform; 2007 (in Russ.).]

¹¹ ГОСТ Р 59709-2022. Национальный стандарт Российской Федерации. *Защита информации. Обнаружение, предупреждение и ликвидация последствий компьютерных атак и реагирование на компьютерные инциденты. Термины и определения*. М.: Росстандарт; 2022. [GOST R 59709-2022. National Standard of the Russian Federation. *Information protection. Detection, prevention and liquidation of the consequences of computer attacks and response to computer incidents. Terms and Definitions*. Moscow: Rosstandart; 2022 (in Russ.).]

¹² <https://www.kaspersky.ru/blog/targeted-attack-anatomy/4388/> (in Russ.). Дата обращения 14.03.2023. / Accessed March 14, 2023.

¹³ https://normative_reference_dictionary.academic.ru/23474/ (in Russ.). Дата обращения 15.03.2023. / Accessed March 15, 2023.

¹⁴ ГОСТ Р ИСО/МЭК 13335-1-2006. Национальный стандарт Российской Федерации. *Информационная технология. Методы и средства обеспечения безопасности*. Часть 1. Концепция и модели менеджмента безопасности информационных и телекоммуникационных технологий. М.: Стандартинформ; 2006. [GOST R ISO/IEC 13335-1-2006. National Standard of the Russian Federation. *Information technology. Methods and means of ensuring security*. Part 1. Concept and models of security management of information and telecommunication technologies. Moscow: Standartinform; 2006 (in Russ.).]

¹⁵ ГОСТ Р 59709-2022. Национальный стандарт Российской Федерации. *Защита информации. Обнаружение, предупреждение и ликвидация последствий компьютерных атак и реагирование на компьютерные инциденты. Термины и определения*. М.: Росстандарт; 2022. [GOST R 59709-2022. National Standard of the Russian Federation. *Information protection. Detection, prevention and liquidation of the consequences of computer attacks and response to computer incidents. Terms and Definitions*. Moscow: Rosstandart; 2022 (in Russ.).]

¹⁶ https://www.group-ib.com/wp-content/uploads/media/2016/02/Group-IB_dbo_instruction.pdf (in Russ.). Дата обращения 13.03.2023. / Accessed March 13, 2023.

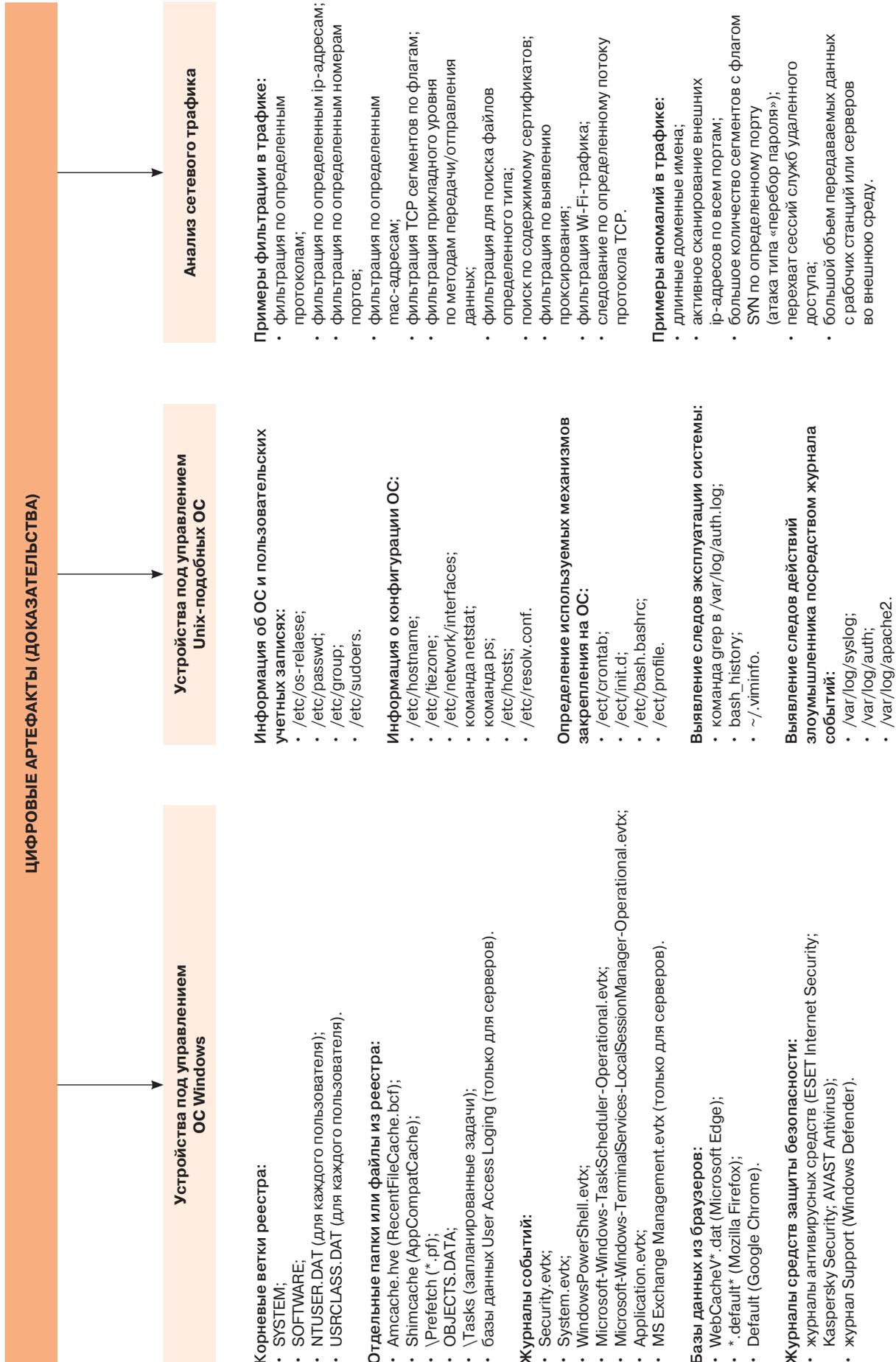


Рисунок. Классификация цифровых артефактов. TCP – Transmission Control Protocol

Стоит отметить, что формирование критериев и показателей оценки качества проведения расследования киберинцидента может повысить эффективность данного процесса, и, следовательно, процесса реагирования на компьютерную атаку с целью минимизации финансового и репутационного ущерба для организации. Таким образом, чем качественнее будет проведено специалистами расследование киберинцидента, тем оперативнее они смогут на него отреагировать и остановить данный вредоносный процесс.

ОБЗОР АКТУАЛЬНЫХ РАБОТ В ОБЛАСТИ РАССЛЕДОВАНИЯ ИНЦИДЕНТОВ ИБ

Практическим вопросам расследования инцидентов ИБ посвящены работы С.И. Макаренко [4], М.А. Еремеева [5], П.Д. Зегжды, Д.П. Зегжды [6], А.Г. Ломако [7], В.А. Овчарова [8], С.А. Петренко [9], И.Б. Саенко [10], И.В. Котенко [11], И.А. Прибылова [12], В.С. Авраменко [13], Д.С. Левшуна [14].

В настоящее время существуют отечественные инструкции коммерческих организаций, описывающие действия специалистов по реагированию на киберинциденты. Эти инструкции имеют ряд недостатков.

Сотрудники компании F.A.C.S.T (ранее Group-IB в России) разработали инструкцию¹⁷ по реагированию на инциденты, связанные с системами дистанционного банковского обслуживания. Недостатком данного документа является возможность его использования только в ИС кредитно-финансовых организаций.

В руководстве¹⁸ сотрудниками Лаборатории Касперского описаны действия специалистов по реагированию на инциденты ИБ. Данный документ не является универсальной инструкцией. В нем описано только применение основных инструментов для сбора данных, анализа потенциальных угроз и их удаления. Данное руководство требует внесения изменений на основе знаний о современных техниках и способах, применяемых злоумышленниками.

В диссертационной работе¹⁹ автором введен обобщенный показатель вредоносной аутентификационной

активности злоумышленника, основанный на показателе полноты выявления аутентификационных действий, совершаемых злоумышленником при «горизонтальном перемещении» в домене, и показателе оперативности расследования киберинцидента. Данный показатель можно рассматривать только в случае кибератаки на домен организации.

Анализ указанных работ в области расследования киберинцидентов показал следующее. Существующие документы, принятые на государственном и коммерческом уровне, описывают только примерные действия специалистов и не определяют временные интервалы, позволяющие оценивать качество проведения расследования киберинцидента.

ОПИСАНИЕ КРИТЕРИЕВ И ПОКАЗАТЕЛЕЙ ОЦЕНИВАНИЯ КАЧЕСТВА ПРОВЕДЕНИЯ РАССЛЕДОВАНИЯ ИНЦИДЕНТА ИБ В РАМКАХ ЦЕЛЕВОЙ КИБЕРАТАКИ

В теории эффективности целенаправленных процессов определено следующее понятие: *эффективность* – это комплексное операционное свойство целенаправленного процесса функционирования, характеризующее его приспособленность к выполнению стоящей перед системой задачи [15].

Понятие эффективности напрямую связано с понятием качества. *Качество* – это свойство или совокупность свойств объекта, обуславливающих его пригодность для использования по назначению. Каждое из свойств объекта может быть описано количественно с помощью некоторой переменной, значение которой характеризует меру (интенсивность) его качества относительно этого свойства. Эту меру называют *показателем свойства* или *единичным, частным показателем качества* объекта [15].

При оценивании качества любого объекта, описываемого n -мерным векторным показателем, реализуется совокупность критериев, каждый из которых в общем случае может принадлежать одному из трех классов:

- класс $\{G\}$ критериев пригодности;
- класс $\{O\}$ оптимальности;
- класс $\{S\}$ превосходства [15].

Критерии могут быть представлены как в векторном, так и в скалярном виде.

Критерий оценивания эффективности – совокупность условий, определяющих цели операции (расследование инцидента ИБ) и в соответствии с ним пригодность, оптимальность или превосходство исследуемой операции [15].

В данной работе для исследования авторами выбран *критерий пригодности G*.

С целью оценивания качества проведения расследования инцидента ИБ в рамках целевой

¹⁷ https://www.group-ib.com/wp-content/uploads/media/2016/02/Group-IB_dbo_instruction.pdf (in Russ.). Дата обращения 13.03.2023. / Accessed March 13, 2023.

¹⁸ https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/07172131/Incident_Response_Guide_rus.pdf (in Russ.). Дата обращения 15.03.2023. / Accessed March 15, 2023.

¹⁹ Смирнов С.И. *Методика проведения расследования киберинцидента на основе автоматизированного анализа событий безопасности домена*: дис. ... канд. техн. наук. СПб., 2022. 124 с. <https://www.elibrary.ru/item.asp?id=54428705>. Accessed March 25, 2023. [Smirnov S.I. *A methodology for conducting a cyber incident investigation based on an automated analysis of domain security events*: Cand. Sci. Thesis. St. Petersburg, 2022. 124 p. (in Russ.). <https://www.elibrary.ru/item.asp?id=54428705>. Accessed March 25, 2023.]

кибератаки (в т.ч. и продолжающейся) авторами предложены следующие показатели:

- показатель результативности (полноты) выявления цифровых артефактов компьютерным криминалистом на основе проведенных мероприятий – r ;
- показатель ресурсоемкости (наличие программных продуктов форензики и затраты людских ресурсов) – p ;
- показатель оперативности расследования инцидента ИБ – t .

ОПИСАНИЕ ПОКАЗАТЕЛЯ РЕЗУЛЬТАТИВНОСТИ

Показатель результативности определен степенью полноты выявления цифровых артефактов r напрямую зависит от девяти основных мероприятий, предложенных авторами, реализуемых при проведении расследования:

- 1) сбор и анализ журналов событий контроллера домена и рабочих станций или серверов;
- 2) анализ процессов из дампа оперативной памяти рабочих станций или серверов;
- 3) анализ истории браузеров рабочих станций;
- 4) анализ почтовых сообщений рабочих станций;
- 5) анализ журналов антивирусных систем, имеющих в организации;
- 6) анализ копий жестких дисков с рабочих станций или серверов;
- 7) анализ трафика с имеющихся сетевых устройств (маршрутизаторы, межсетевые экраны) и с системы обнаружения/предотвращения вторжений (при ее наличии);
- 8) анализ событий SIEM-системы (при ее наличии);
- 9) построение хронологии (таймлайна) целевой атаки.

Основные мероприятия определены на основе практических рекомендаций сотрудников Управления «К» МВД РФ²⁰. Используя собранные цифровые доказательства, специалист должен сформировать отчет о проводимых мероприятиях, в котором будут даны рекомендации по оперативному устранению нештатной ситуации.

²⁰ Управление по организации борьбы с противоправным использованием информационно-коммуникационных технологий Министерства внутренних дел Российской Федерации. <https://xn--b1aew.xn--p1ai/mvd/structure1/Upravlenija/%D1%83%D0%B1%D0%BA/%D0%BF%D0%BE%D0%BB%D0%BE%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5> Дата обращения 30.03.2023 г. [Department of Information Technologies, Communications and Information Protection of the Ministry of Internal Affairs of the Russian Federation. <https://xn--b1aew.xn--p1ai/mvd/structure1/Upravlenija/%D1%83%D0%B1%D0%BA/%D0%BF%D0%BE%D0%BB%D0%BE%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5> (in Russ.). Accessed March 30, 2023.]

На основе описанных выше мероприятий составлена таблица, в которой представлены условные обозначения *показателя результативности (полноты) r в s мероприятиях* и, соответственно, требуемых значений с его описанием (табл. 1).

Таблица 1. Описание требуемого уровня результативности при проведении мероприятий с целью расследования инцидента ИБ

№ мероприятия	Условное обозначение показателя результативности мероприятия s	Описание типовых результатов при проведении каждого мероприятия
1	r_1	Выявление нестандартных способов аутентификации злоумышленника (например, техники Pass-the-Hash, Kerberoasting), методов перебора пароля, аутентификация пользователей домена в нерабочее время
2	r_2	Выявление вредоносных процессов из дампа оперативной памяти «зараженных» машин или серверов
3	r_3	Выявление веб-ресурса, посредством которого произошло заражение рабочей станции
4	r_4	Выявление «фишингового» письма, посредством которого произошло заражение рабочей станции
5	r_5	Выявление в журналах антивирусных систем предупреждений или вредоносного ПО
6	r_6	Выявление в копиях жестких дисков вредоносных файлов или ПО
7	r_7	Выявление вредоносных файлов или ПО, наличие сетевых подключений к командному серверу злоумышленника из трафика
8	r_8	Выявление предупреждений SIEM-системы о кибератаке посредством правил корреляции, нормализации событий безопасности
9	r_9	Построение хронологии целевой атаки с указанием вектора проникновения злоумышленника

Таким образом, показатель результативности (полноты) выявления цифровых артефактов r равен сумме показателей успешных проведенных мероприятий из множества $\{r_s\}$, которые описываются в бинарном виде (выполнено / не выполнено). Формула для расчета показателя результативности имеет вид:

$$r = r_1 + r_2 + (r_3 \vee r_4) + r_5 + r_6 + r_7 + r_8 + r_9. \quad (1)$$

Критерий пригодности по показателю результативности (полноты) выявления цифровых артефактов можно представить в виде неравенства $r > r_{\text{req}}$. Значение требуемого показателя результативности ($r_{\text{req}} = 7$) определено на основе метода экспертных оценок, исключая следующие мероприятия: № 3 или № 4 в зависимости от вектора целевой кибератаки (применено логическое сложение).

ОПИСАНИЕ ПОКАЗАТЕЛЯ РЕСУРСООЕМКОСТИ

Показатель ресурсоемкости p определен наличием программных средств форензики при расследовании киберинцидента и людских ресурсов. Авторами предложено провести разделение программных средств на три подмножества. Данные ОС являются самыми распространенными в организациях при построении информационно-коммуникационных сетей, при этом необходимо проведение анализа сетевого трафика с сетевого оборудования для полноты картины об инциденте ИБ:

- для анализа Windows-подобных систем – $\{p_{\text{win}}\}$;
- для анализа Unix-подобных систем – $\{p_{\text{unix}}\}$;
- для анализа сетевого трафика – $\{p_{\text{traf}}\}$.

Предложенные подмножества рассчитываются на основе имеющихся в арсенале программных средств для расследования инцидента ИБ (имеются / не имеются). Ниже представлены примеры данных программ.

Компьютерные криминалисты для анализа *Windows-подобных систем* должны иметь в своем арсенале необходимый набор программных средств таких как *UserAssist*²¹, *ESEDatabaseView*²², *wmi-parser*²³, *RegRipper*²⁴, утилиты *NirSoft*²⁵ (например, *winprefetchview*, *fulleventlogview*), набор

²¹ https://www.nirsoft.net/utills/userassist_view.html. Дата обращения 30.04.2023. / Accessed April 30, 2023.

²² https://www.nirsoft.net/utills/ese_database_view.html. Дата обращения 30.04.2023. / Accessed April 30, 2023.

²³ <https://github.com/woanware/wmi-parser/releases>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

²⁴ <https://github.com/keydet89/RegRipper3.0>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

²⁵ <https://nirsoft.net/>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

инструментов Эрика Циммермана²⁶ (например, *AncacheParser*, *Registry Explorer*, *MFTECmd*, *AppCompatCacheParser*, *PECmd*), набор программ *Sysinternals Suite* от компании Microsoft²⁷ (например, *PSLoglist*, *Process Monitor*, *Process Explorer*, *Autoruns*, *Autologon* и др.).

Для исследования *Unix-подобных систем* стоит использовать встроенные утилиты такие как *dc3dd*, *ddrescue*, *Autopsy*, *LiME*, *Bulk Extractor*, *Dumpzilla* и др.

Для анализа *трафика* стоит применять *wireshark*, *NetworkMiner*²⁸, *tcpdump*²⁹, *Kismet*³⁰, *SolarWinds Network Bandwidth Analyzer*³¹, *Xplico*³² и др.

Утилиты *FTK Imager*³³, *volatility*³⁴, *artifactcollector*³⁵, *osquery*³⁶, *ir-rescue*³⁷ являются кроссплатформенными и подходят для исследования популярных ОС.

При расследовании инцидента ИБ руководителям подразделений необходимо понимать количество специалистов, которые смогут оперативно провести данное мероприятие. Максимальное значение количества компьютерных криминалистов авторами определено в исследовании как штатная численность отдела.

Затраты людских ресурсов при расследовании инцидента ИБ условно разделены на три подмножества:

- до 2 специалистов – $\{p_{\text{two}}\}$;
- группа специалистов (3–5 чел.) – $\{p_{\text{group}}\}$;
- штатный отдел (10–12 чел.) – $\{p_{\text{dep}}\}$.

²⁶ <https://github.com/EricZimmerman/>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

²⁷ <https://learn.microsoft.com/ru-ru/sysinternals/downloads/sysinternals-suite/> (in Russ.). Дата обращения 30.04.2023. / Accessed April 30, 2023.

²⁸ <https://networkminer.softonic.ru/> (in Russ.). Дата обращения 30.04.2023. / Accessed April 30, 2023.

²⁹ <https://www.microoolap.ru/products/tcpdump-for-windows/> (in Russ.). Дата обращения 30.04.2023. / Accessed April 30, 2023.

³⁰ <https://ru.freedomdownloadmanager.org/Windows-PC/Kismet-FREE.html> (in Russ.). Дата обращения 30.04.2023. / Accessed April 30, 2023.

³¹ <https://softradar.com/solarwinds-network-bandwidth-analyzer-pack/>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

³² <https://www.xplico.org/download>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

³³ <https://accessdata-ftk-imager.software.informer.com/3.1/>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

³⁴ <https://github.com/volatilityfoundation/volatility3>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

³⁵ <https://github.com/forensicanalysis/artifactcollector>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

³⁶ <https://github.com/osquery/osquery>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

³⁷ <https://github.com/diogo-fernan/ir-rescue>. Дата обращения 30.04.2023. / Accessed April 30, 2023.

Математически соотношение данных множеств представлено посредством дизъюнкции. Авторами работы введен коэффициент k , который показывает прямую зависимость от числа задействованных в расследовании компьютерных криминалистов: $k_{two} = 0.8$, $k_{group} = 1$, $k_{dep} = 1.2$. Данные коэффициенты определены на основе метода экспертных оценок исходя из практики расследования киберинцидентов современности. Для успешного решения поставленной перед компьютерными криминалистами задачи необходимо иметь группу специалистов, поэтому данному подмножеству присвоен коэффициент, равный единице.

Таким образом, показатель ресурсоемкости p равен произведению показателей ресурсоемкости трех разновидностей ПО, которые описываются в бинарном виде (в наличии / отсутствуют) на сумму показателей по логическому сложению наличия людских ресурсов с применением коэффициента k .

Математически расчет показателя ресурсоемкости представлен формулой:

$$p = (\{p_{win}\} + \{p_{unix}\} + \{p_{traf}\}) \times (k_{two} \{p_{two}\} \vee k_{group} \{p_{group}\} \vee k_{dep} \{p_{dep}\}). \quad (2)$$

Критерий пригодности по показателю ресурсоемкости (наличие программных средств и людских ресурсов) можно представить в виде неравенства $p > p_{req}$. Значение требуемого показателя ресурсоемкости ($p_{req} = 3$) определено на основе метода экспертных оценок, исключая программные средства для Unix-подобных систем, т.к. при целевых кибератаках злоумышленники в первую очередь нацелены на домен организации с целью шифрования всех данных. Также в него входит значение требуемого показателя ресурсоемкости наличия людских ресурсов.

ОПИСАНИЕ ПОКАЗАТЕЛЯ ОПЕРАТИВНОСТИ

Показатель оперативности расследования инцидента ИБ t рассчитывается на основе временных интервалов мероприятий ($s = 9$), проводимых компьютерным криминалистом. Он определяет временные аспекты проведения расследования.

Показатель оперативности расследования инцидента ИБ t рассчитывается по формуле:

$$t = \sum_{i=1}^s t_i, \quad (3)$$

где t_i – время, затраченное на одно мероприятие.

Требуемые временные интервалы 9 мероприятий при расследовании инцидента приведены в табл. 2. Они рассчитаны на основе метода экспертных оценок. Общее требуемое время расследования t_{req} не должно превышать двух часов.

Таблица 2. Требуемые временные интервалы девяти мероприятий при расследовании инцидента ИБ

№ мероприятия	Условное обозначение требуемого показателя оперативности мероприятия s	Требуемое значение временного интервала мероприятия при расследовании инцидента ИБ, час
1	t_{1_req}	0.25
2	t_{2_req}	0.15
3	t_{3_req}	0.25
4	t_{4_req}	0.25
5	t_{5_req}	0.1
6	t_{6_req}	0.3
7	t_{7_req}	0.2
8	t_{8_req}	0.15
9	t_{9_req}	0.35
10	t_{req}	2

Критерий пригодности по показателю оперативности проведения расследования инцидента ИБ можно представить в виде неравенства $t < t_{req}$. Значение требуемого показателя результативности ($t_{req} = 2$).

Таким образом, критерий пригодности G математически представлен следующим образом на основе трех описанных ранее показателей:

$$G : \left(\bigcap_{i=1}^n \left(\left(r_j^i > \{r_j^{req}\} \right), \left(p_j^i > \{p_j^{req}\} \right), \left(t_j^i > \{t_j^{req}\} \right) \right) \right) \cong U, [j = 1(1)m], \quad (4)$$

где U – достоверное событие (истинное высказывание); \cap – символ булева пересечения событий, $j = 1(1)m$ – упорядоченная совокупность переменных, при которых линейная целевая функция достигает экстремального значения и при этом выполняются (или удовлетворяются) все ограничения в форме равенств или неравенств.

ЗАКЛЮЧЕНИЕ

В статье на основе критерия пригодности описаны показатели оценки качества проведения расследования. Авторами предпринята попытка подвести научную основу под процесс расследования инцидента ИБ от целевой атаки, а именно – на качественном и количественном уровнях сформировать показатели и критерии оценки качества его проведения на основе теории эффективности целенаправленных процессов. Научные результаты, представленные в данной работе, не являются окончательными, а носят дискуссионный характер. Авторы надеются, что

статья станет своеобразной «отправной точкой» для руководителей и специалистов в сфере компьютерной криминалистики при формировании научных подходов в их служебной деятельности.

В дальнейших работах планируется развитие тематического аппарата в данной области исследования и увеличение числа показателей/критериев оценивания процесса расследования инцидента ИБ от целевых атак современности.

Вклад авторов

С.И. Смирнов – описание критериев и показателей оценивания качества проведения расследования ИБ.

М.А. Еремеев – идея исследования, развитие цели и задач, формулировка выводов.

Ш.Г. Магомедов – изучение цифровых доказательств и программных средств при расследовании инцидента ИБ.

Д.А. Изергин – обоснование целесообразности формирования критериев и показателей оценки качества проведения расследования инцидента ИБ.

Authors' contributions

S.I. Smirnov – description of criteria and indicators for assessing the quality of an information security investigation.

M.A. Ereemeev – the idea of research, the development of aims and objectives, the formulation of conclusions.

Sh.G. Magomedov – study of digital evidence and software tools in the investigation of an information security incident.

D.A. Izergin – justification of the expediency of forming criteria and indicators for assessing the quality of the investigation of an information security incident.

СПИСОК ЛИТЕРАТУРЫ

1. Смирнов С.И., Еремеев М.А., Горбачев И.Е., Нефедов В.С., Изергин Д.А. Анализ техник и инструментов, используемых злоумышленником при горизонтальном перемещении в корпоративной сети. *Защита информации. Инсайд*. 2021;1(97):58–61. <https://www.elibrary.ru/pltlpq>
2. Смирнов С.И. Методика расследования киберинцидента, основанная на интеллектуальном анализе событий безопасности домена. *Защита информации. Инсайд*. 2022;4(106):60–69. <https://www.elibrary.ru/mefhpc>
3. Смирнов С.И., Киселев А.Н., Азерский В.Д., Карельский Д.В., Кумуржи Г.М. Комплексная методика проведения расследования инцидента информационной безопасности. *Защита информации. Инсайд*. 2023;2(110):14–26. <https://www.elibrary.ru/fdhgzq>
4. Макаренко С.И. Критерии и показатели оценки качества тестирования на проникновение. *Вопросы кибербезопасности*. 2021;3(43):43–57. <https://www.elibrary.ru/udlknn>
5. Smirnov S.I., Ereemeev M.A., Pribylov I.A. Approach to Recognition of Malicious Behavior Based on Autoregression Model upon Investigation into Cyberincident. *Aut. Control Comp. Sci.* 2021;55(8):1099–1103. <http://doi.org/10.3103/S0146411621080290>, <https://www.elibrary.ru/ubwpa1>
6. Зегжда Д.П., Лаврова Д.С., Павленко Е.Ю. Управление динамической инфраструктурой сложных систем в условиях целенаправленных кибератак. *Известия РАН. Теория и системы управления*. 2020;4(3):50–63. <https://doi.org/10.31857/S0002338820020134>
7. Калинин В.Н., Ломако А.Г., Овчаров В.А., Петренко С.А. Расследование ИБ-инцидентов с использованием профилирования поведения динамических сетевых объектов. *Защита информации. Инсайд*. 2018;3(81):58–67. <https://www.elibrary.ru/xqlamp>
8. Овчаров В.А., Романов П.А. Расследование компьютерных инцидентов на основе идентификации дискретных событий информационной безопасности и обратного анализа по конечным исходам. *Труды Военно-космической академии имени А.Ф. Можайского*. 2015;648:84–89. <https://www.elibrary.ru/uzmkoх>
9. Ломако А.Г., Овчаров В.А., Петренко С.А. Метод расследования инцидентов безопасности на основе профилей поведения сетевых объектов. В сб.: *Дистанционные образовательные технологии: Материалы III Всероссийской научно-практической конференции*, 17–22 сентября 2018 г. Ялта: ООО «Издательство Типография «Ариал»; 2018. С. 366–373. <https://www.elibrary.ru/uzzdah>
10. Саенко И.Б., Лаута О.С., Карпов М.А., Крибель А.М. Модель угроз ресурсам ИТКС как ключевому активу критически важного объекта инфраструктуры. *Электросвязь*. 2021;1:36–44. <https://doi.org/10.34832/ELSV.2021.14.1.004>
11. Быстров И.С., Котенко И.В. Анализ моделей поведения пользователей для задачи обнаружения кибер-инсайдеров. В сб.: *Актуальные проблемы инфотелекоммуникаций в науке и образовании: сборник научных статей*: в 4 т. Т. 1. СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича; 2021. С. 139–143. <https://www.elibrary.ru/sqzvmа>
12. Ereemeev M.A., Smirnov S.I., Pribylov I.A. Detection of malicious actions of an attacker based on event logs when investigating an ongoing cyber incident. В сб.: *Инновационные аспекты развития науки и техники: Сборник статей VII Международной научно-практической конференции*. Саратов: НОО «Цифровая наука»; 2021. С. 22–28. <https://www.elibrary.ru/ygoyfz>

13. Авраменко В.С., Маликов А.В. Нейросетевая модель диагностирования компьютерных инцидентов в инфокоммуникационных системах специального назначения. В сб.: *Проблемы технического обеспечения войск в современных условиях: Труды IV Межвузовской научно-практической конференции*. Т. 1. СПб.; 2019. С. 41–45. <https://www.elibrary.ru/flomvh>
14. Левшун Д.С. Построение модели атакующего для современной киберфизической системы. В сб.: *Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2020). IX Международная научно-техническая и научно-методическая конференция: сборник научных статей*. Т. 1. СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича; 2020. С. 679–682. <https://www.elibrary.ru/krafgr>
15. Петухов Г.Б., Якунин В.И. *Методологические основы внешнего проектирования целенаправленных процессов и целеустремленных систем*. М.: АСТ; 2006. 504 с.

REFERENCES

1. Smirnov S.I., Ereemeev M.A., Gorbachev I.E., Nefedov V.S., Izergin D.A. Analysis of techniques and tools used by an attacker when moving horizontally in the corporate network. *Zashchita Informatsii. Insaidd*. 2021;1(97):58–61 (in Russ.). <https://www.elibrary.ru/pltlpq>
2. Smirnov S.I. Cyber incident investigation methodology based on intelligent analysis of domain security events. *Zashchita Informatsii. Insaidd*. 2022;4(106):60–69 (in Russ.). <https://www.elibrary.ru/mefhpc>
3. Smirnov S.I., Kiselev A.N., Azerskii V.D., Karel'skii D.V., Kumurzhi G.M. Comprehensive methodology for conducting an information security incident investigation. *Zashchita Informatsii. Insaidd*. 2023;2(110):14–26 (in Russ.). <https://www.elibrary.ru/fdhgzq>
4. Makarenko S.I. Criteria and parameters for estimating quality of penetration testing. *Voprosy kiberbezopasnosti = Cybersecurity Issues J*. 2021;3(43):43–57 (in Russ.). <https://www.elibrary.ru/udlkn>
5. Smirnov S.I., Ereemeev M.A., Pribylov I.A. Approach to Recognition of Malicious Behavior Based on Autoregression Model upon Investigation into Cyberincident. *Aut. Control Comp. Sci*. 2021;55(8):1099–1103. <http://doi.org/10.3103/S0146411621080290>, <https://www.elibrary.ru/ubwpa>
6. Zegzhda D.P., Lavrova D.S., Pavlenko E.Y. Management of a Dynamic Infrastructure of Complex Systems Under Conditions of Directed Cyber Attacks. *J. Comput. Syst. Sci. Int*. 2020;59(3):358–370. <https://doi.org/10.1134/S1064230720020124> [Original Russian Text: Zegzhda D.P., Lavrova D.S., Pavlenko E.Y. Management of a Dynamic Infrastructure of Complex Systems Under Conditions of Directed Cyber Attacks. *Izvestiya Rossiiskoi akademii nauk. Teoriya i sistemy upravleniya*. 2020;3:50–63 (in Russ.). <https://doi.org/10.31857/S0002338820020134>]
7. Kalinin V.N., Lomako A.G., Ovcharov V.A., Petrenko S.A. Investigation of information security incidents using the behavior profiling of dynamic network objects. *Zashchita Informatsii. Insaidd*. 2018;3(81):58–67 (in Russ.). <https://www.elibrary.ru/xqlamp>
8. Ovcharov V.A., Romanov P.A. Investigation of computer incidents based on the identification of discrete IS events and reverse analysis by final outcomes. *Trudy Voenno-kosmicheskoi akademii imeni A.F. Mozhaiskogo = Proceedings of the Mozhaisky Military Aerospace Academy*. 2015;648:84–89 (in Russ.). <https://www.elibrary.ru/uzmkox>
9. Lomako A.G., Ovcharov V.A., Petrenko S.A. Method for investigating security incidents based on behavior profiles of network objects. In: *Distance Educational Technologies: Materials of the Third All-Russian Scientific and Practical Conference*, September 17–22, 2018. Yalta: Arial; 2018. P. 366–373 (in Russ.). <https://www.elibrary.ru/uzzdah>
10. Saenko I.B., Lauta O.S., Karpov M.A., Kribel A.M. Model of threats to information and telecommunication network resources as a key asset of critical infrastructure. *Elektrosvyaz*. 2021;1:36–44 (in Russ.). <https://doi.org/10.34832/ELSV.2021.14.1.004>
11. Bystrov I.S., Kotenko I.V. Analysis of user behavior models for the task of detecting cyber insiders. In: *Actual Problems of Infotelecommunications in Science and Education: collection of scientific articles*: in 4 v. V. 1. St. Petersburg: Bonch-Bruevich St. Petersburg State University of Telecommunications; 2021. P. 139–143 (in Russ.). <https://www.elibrary.ru/sqzvm>
12. Ereemeev M.A., Smirnov S.I., Pribylov I.A. Detection of malicious actions of an attacker based on event logs when investigating an ongoing cyber incident. In: *Innovative Aspects of the Development of Science and Technologies: Collection of articles of the 7th International Scientific and Practical Conference*. Saratov: Tsifrovaya nauka; 2021. P. 22–28 (in Russ.). <https://www.elibrary.ru/ygoyfz>
13. Avramenko V.S., Malikov A.V. Neural network model for diagnosing computer incidents in special purpose infocommunication systems. In: *Problems of Technical Support of Troops in Modern Conditions: Proceedings of the Forth Interuniversity Scientific and Practical Conference*. St. Petersburg; 2019. P. 41–45 (in Russ.). <https://www.elibrary.ru/flomvh>
14. Levshun D.S. Building an attacker model for a modern cyberphysical system. In: *Actual Problems of Infotelecommunications in Science and Education (APINO 2020). The 9th International Scientific-Technical and Scientific-Methodological Conference: collection of scientific articles*. V. 1. St. Petersburg: Bonch-Bruevich St. Petersburg State University of Telecommunications; 2020. P. 679–682 (in Russ.). <https://www.elibrary.ru/krafgr>
15. Petukhov G.B., Yakunin V.I. *Metodologicheskie osnovy vneshnego proektirovaniya tselenapravlennykh protsessov i tselestremennykh system (Methodological Foundations of External Design of Purposeful Processes and Purposeful Systems)*. Moscow: AST; 2006. 504 p. (in Russ.).

Об авторах

Смирнов Станислав Игоревич, к.т.н., доцент, кафедра интеллектуальных систем информационной безопасности, Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: smirnov_si@mirea.ru. Scopus Author ID 57475289100, ResearcherID HZM-3994-2023, SPIN-код РИНЦ 1472-6572, <https://orcid.org/0000-0003-4387-0850>

Еремеев Михаил Алексеевич, д.т.н., профессор, кафедра информационно-аналитических систем кибербезопасности, Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: eremeev_m@mirea.ru. Scopus Author ID 57188205500, SPIN-код РИНЦ 3609-5733, <https://orcid.org/0000-0002-5511-4000>

Магомедов Шамиль Гасангусейнович, к.т.н., доцент, заведующий кафедрой интеллектуальных систем информационной безопасности, Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: msgg@list.ru. Scopus Author ID 57204759220, ResearcherID M-5782-2016, SPIN-код РИНЦ 5029-8310, <https://orcid.org/0000-0001-8560-1937>

Изергин Дмитрий Андреевич, к.т.н., доцент, кафедра цифровых технологий обработки данных, Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: izergin@mirea.ru. Scopus Author ID 57224822181, SPIN-код РИНЦ 2318-9152, <https://orcid.org/0000-0002-3174-4550>

About the authors

Stanislav I. Smirnov, Cand. Sci. (Eng.), Assistant Professor, Department of Intelligent Information Security Systems, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: smirnov_si@mirea.ru. Scopus Author ID 57475289100, ResearcherID HZM-3994-2023, RSCI SPIN-code 1472-6572, <https://orcid.org/0000-0003-4387-0850>

Mikhail A. Ereemeev, Dr. Sci. (Eng.), Professor, Department of Information and Analytical Cybersecurity Systems, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: eremeev_m@mirea.ru. Scopus Author ID 57188205500, RSCI SPIN-code 3609-5733, <https://orcid.org/0000-0002-5511-4000>

Shamil G. Magomedov, Cand. Sci. (Eng.), Associate Professor, Head of the Department of Intelligent Information Security Systems, Institute of Cyber Security and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). msgg@list.ru. Scopus Author ID 57204759220, ResearcherID M-5782-2016, RSCI SPIN-code 5029-8310, <https://orcid.org/0000-0001-8560-1937>

Dmitry A. Izergin, Cand. Sci. (Eng.), Assistant Professor, Department of Digital Data Processing Technologies, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: izergin@mirea.ru. Scopus Author ID 57224822181, RSCI SPIN-code 2318-9152, <https://orcid.org/0000-0002-3174-4550>

УДК 004.2

<https://doi.org/10.32362/2500-316X-2024-12-3-37-45>

EDN PXKDKR



НАУЧНАЯ СТАТЬЯ

Методика проектирования специализированных вычислительных систем на основе совместной оптимизации аппаратного и программного обеспечения

И.Е. Тарасов[@],
П.Н. Советов,
Д.В. Люлява,
Д.И. Мирзоян

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[@] Автор для переписки, e-mail: tarasov_j@mirea.ru

Резюме

Цели. Следующим этапом повышения производительности вычислительных систем после завершения этапов роста за счет масштабирования транзисторов (закон Деннарда) и за счет увеличения количества процессорных ядер общего назначения (ограничиваемого законом Амдала) является переход к разработке специализированных вычислительных подсистем для работы в ограниченном подклассе задач. Создание таких систем требует как выбора соответствующих массово востребованных задач, так и применения методик проектирования, обеспечивающих достижение высоких технико-экономических показателей разрабатываемых специализированных сверхбольших интегральных схем. Цель работы – разработка методики проектирования специализированных вычислительных систем на основе совместной оптимизации аппаратного и программного обеспечения применительно к выбранному подклассу задач.

Методы. Используются методы проектирования цифровых систем.

Результаты. Рассмотрены подходы к анализу вычислительных задач путем построения графа вычислений, абстрагированного от вычислительной платформы, однако ограниченного набором архитектурных решений. Предложена методика проектирования, использующая маршрут, основанный на применении синтезатора представления уровня регистровых передач (RTL-представления) вычислительного устройства, ограниченного отдельными вычислительными архитектурами, для которых производятся синтез и оптимизация схемы на основе высокоуровневого входного описания алгоритма. Среди архитектур вычислительных узлов рассмотрены синхронный конвейер и процессорное ядро с древовидным арифметико-логическим устройством. Повышение эффективности вычислительной системы осуществляется путем балансировки конвейера на основе оценок технологического базиса, а для процессора – путем оптимизации набора операций на основе анализа графа абстрактного синтаксического дерева с его оптимальным покрытием подграфами, соответствующим структуре арифметико-логического устройства.

Выводы. Рассмотренные подходы к разработке позволяют ускорить процесс проектирования специализированных вычислительных систем с массово-параллельной архитектурой, основанных на конвейерных вычислительных узлах.

Ключевые слова: процессор, RTL, синтез, транслятор

• Поступила: 18.10.2023 • Доработана: 04.12.2023 • Принята к опубликованию: 22.03.2024

Для цитирования: Тарасов И.Е., Советов П.Н., Люлява Д.В., Мирзоян Д.И. Методика проектирования специализированных вычислительных систем на основе совместной оптимизации аппаратного и программного обеспечения. *Russ. Technol. J.* 2024;12(3):37–45. <https://doi.org/10.32362/2500-316X-2024-12-3-37-45>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Method for designing specialized computing systems based on hardware and software co-optimization

Ilya E. Tarasov[@],
Peter N. Sovietov,
Daniil V. Lulyava,
Dmitry I. Mirzoyan

MIREA – Russian Technological University, Moscow, 119454 Russia

[@] Corresponding author, e-mail: tarasov_j@mirea.ru

Abstract

Objectives. Following the completion of development stages due to transistor scaling (Dennard's law) and an increased number of general-purpose processor cores (limited by Amdahl's law), further improvements in the performance of computing systems naturally proceeds to the stage of developing specialized computing subsystems for performing specific tasks within a limited computational subclass. The development of such systems requires both the selection of the relevant high-demand tasks and the application of design techniques for achieving desired indicators within the developed specializations at very large scales of integration. The purpose of the present work is to develop a methodology for designing specialized computing systems based on the joint optimization of hardware and software in relation to a selected subclass of problems.

Methods. The research is based on various methods for designing digital systems.

Results. Approaches to the analysis of computational problems involving the construction of a computational graph abstracted from the computing platform, but limited by a set of architectural solutions, are considered. The proposed design methodology based on a register transfer level (RTL) representation synthesizer of a computing device is limited to individual computing architectures for which the relevant circuit is synthesized and optimized based on a high-level input description of the algorithm. Among computing node architectures, a synchronous pipeline and a processor core with a tree-like arithmetic-logical unit are considered. The efficiency of a computing system can be increased by balancing the pipeline based on estimates of the technological basis, and for the processor—based on optimizing the set of operations, which is performed based on the analysis of the abstract syntax tree graph with its optimal coverage by subgraphs corresponding to the structure of the arithmetic logic unit.

Conclusions. The considered development approaches are suitable for accelerating the process of designing specialized computing systems with a massively parallel architecture based on pipeline or processor computing nodes.

Keywords: processor, RTL, synthesis, translator

• Submitted: 18.10.2023 • Revised: 04.12.2023 • Accepted: 22.03.2024

For citation: Tarasov I.E., Sovietov P.N., Lulyava D.V., Mirzoyan D.I. Method for designing specialized computing systems based on hardware and software co-optimization. *Russ. Technol. J.* 2024;12(3):37–45. <https://doi.org/10.32362/2500-316X-2024-12-3-37-45>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Ситуация в области проектирования элементной базы для высокопроизводительных вычислений определяется рядом тенденций, соответствующих как объективным техническим ограничениям, так и необходимости интенсификации процессов импортозамещения и обеспечения технологического суверенитета. В связи с этим, анализируя желаемые технические характеристики, необходимо учитывать также возможности их достижения с учетом ограничений возможности производства и необходимости уменьшения технических и экономических рисков.

Анализ архитектурных тенденций в области вычислительных средств представлен Паттерсоном и Хеннесси в работе [1]. Авторы обращают внимание на ряд крупных этапов в области развития процессорных архитектур, начиная с 1970-х гг. Первым из отмеченных архитектурных переходов стала смена концепции CISC (complex instruction set computer) на RISC (reduced instruction set computer). Уменьшение сложности комбинационной логики, ставшее следствием такого перехода, позволило повысить тактовую частоту процессорных устройств.

Дальнейшее повышение тактовой частоты оказалось ограниченным после прекращения действия закона Деннарда, объясняющего увеличение производительности процессоров масштабированием размеров транзистора при переходе к следующему поколению технологических процессов. Реакцией на этот эффект стал переход к многоядерным процессорам, состоявшийся в массовом сегменте персональных компьютеров в середине 2000-х гг.

В свою очередь, повышение производительности за счет увеличения количества процессорных ядер ограничено законом Амдала, который определяет потенциальное повышение производительности многопроцессорного комплекса через долю вычислений, которые могут быть выполнены параллельно. Связанной проблемой является так называемая «стена интерфейсов» [2], учитывающая тот факт, что с уменьшением нормы технологического процесса производительность растет квадратично, а пропускная способность интерфейсов памяти

и периферийных устройств – линейно. Поэтому построение многоядерных систем влечет за собой проблему организации межпроцессорного обмена данными, которая не может быть эффективно реализована из-за опережающего роста объема получаемых данных по сравнению с возможностью их передачи по существующим каналам связи, проектируемым по сопоставимым технологическим нормам.

В связи с изложенными проблемами в [1] предлагается переход к проблемно-ориентированным архитектурам (DSA, domain-specific architecture) по аналогии с проблемно-ориентированными языками программирования (DSL, domain-specific language). При этом специализация процессора для выполнения определенных классов вычислений, по сути, означает снижение его эффективности в других классах, что требует выбора для специализации таких целевых вычислительных задач, которые соответствовали бы актуальным техническим потребностям, допускали бы широкое применение из соображений снижения удельной стоимости проектирования и подготовки производства, а кроме того, соответствовали бы технически реализуемым подходам к проектированию цифровых устройств.

АНАЛИЗ ВЫЧИСЛИТЕЛЬНЫХ ЗАДАЧ ДЛЯ РЕАЛИЗАЦИИ В СОСТАВЕ СПЕЦИАЛИЗИРОВАННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Важными направлениями для применения высокопроизводительных вычислительных комплексов являются системы обработки видео, виртуальная и дополненная реальность, робототехника, промышленная автоматика, цифровая радиосвязь, измерительная техника и ряд других [3, 4]. Среди реализуемых направлений обработки сигналов можно выделить цифровую фильтрацию [5], спектральный анализ [6], алгоритмы машинного обучения [7], в т.ч. на базе специализированных нейропроцессоров [8] или реконфигурируемых ускорителей на базе программируемых логических интегральных схем (ПЛИС).

К вычислительным задачам, которые могут быть ускорены специализированными вычислительными системами, можно отнести следующие подклассы:

Таблица. Интенсивность использования операций, характерных для ряда задач, требующих применения высокопроизводительных вычислительных систем

Вид задачи	Сдвиги, сложение, поразрядные операции	Умножение	Операции с плавающей точкой	Трансцендентные функции	Операции с памятью
Хеш-функции	Массово	Нет	Нет	Нет	Возможно
Реализация нейросетей	Нет	Массово	Возможно	Возможно	Возможно
Цифровая обработка сигналов (фильтрация)	Массово	Массово	Возможно	Возможно	Возможно
Системы дифференциальных уравнений	Нет	Возможно	Массово	Возможно	Массово
Обработка трехмерной графики	Нет	Часто	Массово	Часто	Массово

1. Решение систем дифференциальных уравнений численными методами.
2. Операции с трехмерными изображениями.
3. Цифровая обработка сигналов на основе массового применения операций «умножение с накоплением (multiply and accumulate)».
4. Вычисление хеш-функций в задачах защиты информации.
5. Реализация нейросетей в части вычисления значений функций нейронов (neural net inference), не включая задачи обучения нейросети.

Виды операций, характерных для указанных задач, приведены в таблице. При этом столбцы таблицы размещены так, что сложность реализации соответствующих видов операций возрастает слева направо. Операции с памятью обозначены как имеющие наибольшую сложность вследствие того, что увеличение пропускной способности подсистемы памяти сопряжено с существенными затруднениями, хотя отдельные операции с памятью сами по себе высокой сложностью не обладают.

В таблице использованы следующие оценки, характеризующие интенсивность использования тех или иных видов операций. Оценка «нет» соответствует ситуации, когда операция не используется в алгоритмах и не требует поддержки. Оценка «возможно» характеризует ситуацию, когда такие действия имеют место, но в силу редкого использования не оказывают заметного влияния на эффективность вычислительного устройства. Для таких операций возможно применение неоптимизированных решений или готовых компонентов с функциональной избыточностью. Оценка «массово» соответствует операциям, которые являются основой алгоритмов и в наибольшей степени определяют эффективность вычислительного устройства, предназначенного для их реализации.

По предварительным оценкам видно, что реализация хеш-функций и операций цифровой обработки

сигналов позволяет в большей степени продемонстрировать преимущества конвейерных архитектур, поскольку предусматривает потоковую обработку данных без интенсивного обмена с внешней памятью.

АРХИТЕКТУРЫ ВЫЧИСЛИТЕЛЬНЫХ УЗЛОВ

Архитектурные подходы к реализации отдельных вычислительных узлов в современной цифровой электронике достаточно разнообразны. Для их эффективного практического использования следует ограничиться набором возможных решений, допускающих применение методов автоматизированного проектирования на уровне математических и программных моделей с тем, чтобы дальнейшие преобразования в представление схемы уровня регистровых передач (RTL-представление¹) не вносили существенных изменений в характеристики такой системы. Можно отметить, например, что применение средств высокоуровневого описания класса HLL (high-level language) предполагает автоматизированное построение управляющих схем (flow control), рассчитанных на широкий класс реализуемых архитектурных подходов. Это ведет к излишнему усложнению синтезируемых управляющих схем.

Для практически реализуемой методики проектирования рассмотрены следующие архитектурные подходы к построению вычислительных узлов:

1. Процессорный узел.
2. Синхронный конвейер.
3. Модификация синхронного конвейера с возможностью повторного использования отдельных стадий.

Структурные схемы основных вычислительных узлов показаны на рис. 1.

¹ RTL – register transfer level.

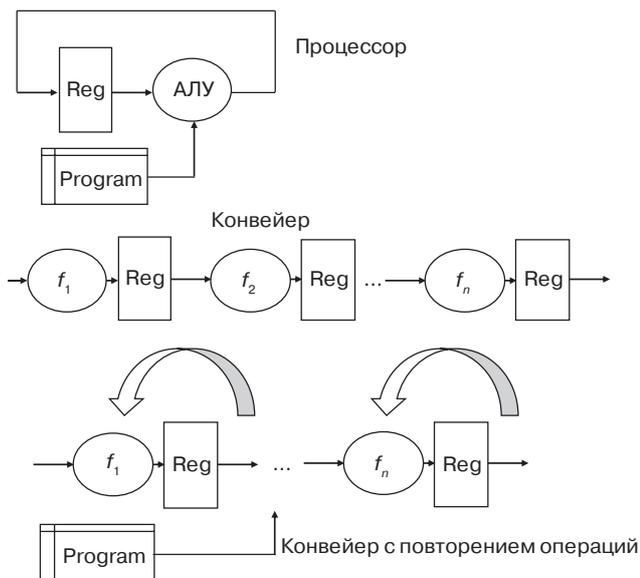


Рис. 1. Схемы основных вычислительных узлов (архитектурные шаблоны). Reg – регистр, Program – программа, f_1, f_2, \dots, f_n – функциональные устройства

Приведенные варианты узлов рассматриваются в качестве архитектурных шаблонов для реализации выбранных подклассов вычислений. При этом для процессора требуемый набор операций реализуется в составе арифметико-логического устройства (АЛУ), а для конвейера – в последовательных стадиях. Для использования конвейера необходимо, чтобы порядок действий для реализации алгоритма оставался неизменным, иначе корректировка потребует усложнения управляющих схем. Повторение операций может быть использовано в таких алгоритмах, как вычисление хеш-функций и реализация фильтров с конечной импульсной характеристикой, при условии, что частота получения входных данных существенно меньше, чем тактовая частота конвейера. Выполнение этого условия позволит использовать одну и ту же стадию конвейера многократно до прихода нового входного значения.

Выбор конкретного архитектурного шаблона определяет операции по синтезу RTL-представления и его оптимизации. В данной статье рассмотрено применение методики для конвейерных вычислительных структур. Добавление архитектурных шаблонов процессора и конвейера с повторением операций предусмотрено для последующих этапов проекта.

Маршрут проектирования согласно предлагаемой методике проектирования вычислительных модулей специализированной вычислительной системы представлен на рис. 2.

В представленном маршруте видно, что входными данными являются исходные тексты реализуемого



Рис. 2. Маршрут проектирования модулей специализированной вычислительной системы

алгоритма и конструктивные ограничения, представленные в виде предельных характеристик требуемого решения. Разработанный П.Н. Советовым специализированный синтезатор [9] на основе архитектурных шаблонов генерирует RTL-представление модуля, используя для предварительной оценки его характеристик данные о технологической библиотеке. Получаемое RTL-представление в дальнейшем используется в маршрутах проектирования ПЛИС или сверхбольших интегральных схем (СБИС), где соответствующие системы автоматизированного проектирования (САПР) позволяют оценить характеристики модуля после синтеза или после выполнения размещения и трассировки (что дает более точную оценку характеристик по сравнению с оценкой после синтеза).

Например, величины задержек распространения сигналов становятся основанием для повторного синтеза RTL-представления с дополнительной конвейеризацией выявленных критических цепей. Кроме этого, синтезатор обеспечивает дополнительные сведения о взаимосвязях синтезируемых узлов, которые позволяют генерировать проектные ограничения для задания координат отдельных узлов синтезированной схемы (стадий конвейера). Подобные возможности частично обеспечивают такие инструменты проектирования, как *Vitis HLS*², однако, в отличие от разработанного маршрута, в *HLS* характеристики технологической платформы задаются в виде библиотек и не подлежат уточнению в процессе оптимизации проекта.

² <https://www.xilinx.com/products/design-tools/vitis/vitis-hls.html>. Дата обращения 10.10.2023. / Accessed October 10, 2023.

ОПТИМИЗАЦИЯ ВЫЧИСЛИТЕЛЬНОГО УЗЛА НА ОСНОВЕ ПРОГРАММНОЙ МОДЕЛИ ВЫЧИСЛЕНИЙ

Разработка компиляторов для новых процессорных архитектур является важной составляющей в обеспечении инструментальных средств проектирования систем на их базе [10]. Трансляторы предметно-ориентированных языков (DSL) в RTL-представление представляют собой перспективный подход к быстрому проектированию аппаратных ускорителей некоторого узкого класса архитектур [11]. Кратко изложим этапы проектирования инструментальной системы, состоящей из встроенного DSL на основе подмножества языка Python и транслятора для синтеза аппаратных ускорителей на основе конвейеризации линейного участка программы.

На вход транслятора подается программа пользователя, представляющая собой поведенческое описание синтезируемого аппаратного ускорителя. Эта программа с помощью модуля `ast` из стандартной библиотеки Python автоматически преобразуется в форму дерева абстрактного синтаксиса (`abstract syntax tree`, AST). Проверка и распространение по AST информации о типах, используемых во входной программе, осуществляются на основе механизма аннотации типов Python.

Кроме того, пользователь предоставляет таблицу задержек и правил комбинирования операций для выбранного типа микросхемы, а также один из выбранных шаблонов управления конвейером на языке Verilog. Результатом работы транслятора является код синтезированного конвейеризованного ускорителя на языке Verilog.

Все функции во входной программе встраиваются в главную функцию, а циклы полностью разворачиваются. Далее программа преобразуется в ациклический граф зависимостей по данным (DDG, `data dependency graph`), при этом на основе нумерации значений осуществляются свертка и продвижение констант, удаление совпадающих выражений и удаление мертвого кода. Для достижения потенциально большего параллелизма вычислений используется сбалансирование высоты деревьев выражений.

Перед непосредственным синтезом конвейера выполняются вспомогательные этапы: этап частичного покрытия DDG с помощью MISO-подграфов (`multiple input single output`, множество входов и один выход) и этап вычисления максимальных задержек между парами узлов в DDG.

В целевом типе микросхемы могут использоваться ресурсы, позволяющие комбинировать, т.е. совмещать во времени выполнение нескольких операций, например, с помощью таблицы истинности (LUT, `look-up table`). В трансляторе используется

частичное покрытие DDG комбинированными операциями с использованием варианта алгоритма MAXMISO для синтеза команд. Этот алгоритм позволяет перечислить в DDG непересекающиеся подграфы, имеющие количество входов, не более заданного, и один выход.

Результатом синтеза конвейера является DDG с добавленными узлами – конвейеризирующими регистрами. Синтез конвейера реализуется преимущественно с использованием сторонних решателей программирования в ограничениях и линейного программирования, что упрощает реализацию генератора кода. Синтез конвейера осуществляется одним из следующих способов: с минимизацией глубины конвейера, с минимизацией общего размера конвейеризирующих регистров или же с минимизацией глубины конвейера, за которой следует минимизация общего размера конвейеризирующих регистров. При синтезе конвейера учитывается представление программы в виде DDG и используется информация об узлах – конечных потребителях, что упрощает алгоритмы синтеза и позволяет сократить число формируемых ограничений. Аналогичный подход был применен для создания компилятора для специализированного контроллера на базе программируемой пользователем вентильной матрицы (FPGA, `field-programmable gate array`) [12]. Можно отметить наличие подобных подходов к трансляции высокоуровневого представления программ [13, 14], которые, тем не менее, не предполагают использование обратной связи от САПР, выполняющей трассировку проекта и определение реально достигнутых временных задержек. При этом уделяется внимание конвейерным архитектурам для FPGA [15].

МЕТОДИКА ОЦЕНКИ ТОПОЛОГИЧЕСКОЙ РЕАЛИЗАЦИИ ВЫЧИСЛИТЕЛЬНОГО УЗЛА

Синтезированное RTL-представление, рассмотренное выше, в качестве входной информации использует оценочные сведения о задержках, вносимых отдельными операциями. В этой связи для выбранного технологического базиса требуется:

1. Определить задержки отдельных элементов, реализующих вычисления, поддерживаемые синтезатором.
2. Выявить возможность использования аддитивной модели задержек или определить способ определения суммарной задержки комбинационного узла с учетом взаимодействия отдельных элементов.

Проверка топологической реализации примера конвейерного вычислителя была проведена в САПР ПЛИС *AMD/Xilinx Vivado*³. Конвейерный

³ <https://docs.xilinx.com/r/en-US/ug910-vivado-getting-started>. Дата обращения 10.10.2023. / Accessed October 10, 2023.

вычислитель реализует алгоритм вращения векторов CORDIC⁴, являющийся основой для вычисления трансцендентных функций. Этот выбор связан с тем, что IP-ядро CORDIC входит в состав библиотечных компонентов САПР *AMD/Xilinx Vivado*, и его характеристики могут быть сопоставлены с полученными результатами. Вычисление шагов алгоритма CORDIC совмещено с последовательным вычислением результата умножения с накоплением. Таким образом, обеспечено превышение функциональных возможностей относительно IP-ядра CORDIC. Количественным критерием оценки качества предварительного моделирования задержек является гистограмма запасов временных задержек (slack histogram)⁵. При статическом временном анализе величиной запаса является разность между значением периода тактового сигнала и максимальной задержкой распространения сигнала между синхронными узлами схемы. В зависимости от сложности выражений и взаимного расположения узлов задержка будет индивидуальной для каждой цепи, что позволяет построить гистограмму, показывающую количество цепей, имеющих соответствующие запасы по времени до прихода следующего фронта тактового сигнала. Такая гистограмма называется гистограммой запасов и формируется в САПР *Vivado* по запросу оператора на основе проведенного в САПР статического временного анализа.

Исходя из соображений балансировки стадий конвейера в идеальном случае, можно предположить, что запасы будут сгруппированы около минимальных значений, что будет говорить об отсутствии цепей, имеющих слишком малую задержку, а, следовательно, неэффективно использующих аппаратные ресурсы. Пример гистограммы запасов временных задержек приведен на рис. 3.

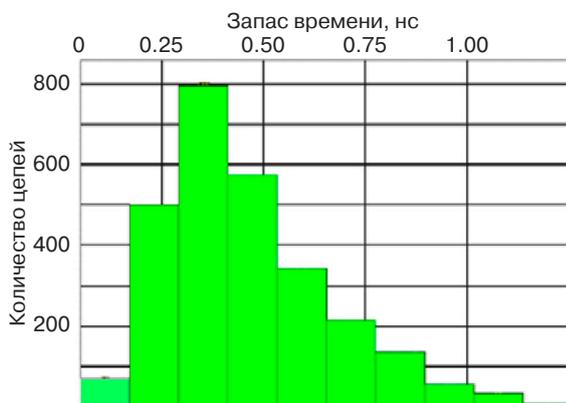


Рис. 3. Гистограмма запасов временных задержек для примера конвейера

⁴ Coordinate rotational digital computer.

⁵ <https://docs.xilinx.com/r/en-US/ug906-vivado-design-analysis/Timing-Analysis>. Дата обращения 10.10.2023. / Accessed October 10, 2023.

Группировка цепей на гистограмме показывает, что балансировка стадий конвейера в целом выполнена корректно, поскольку основная часть цепей находится в левой части гистограммы, которая соответствует небольшим величинам запаса по времени.

ПРИМЕРЫ ПРАКТИЧЕСКОЙ АПРОБАЦИИ МЕТОДИКИ

Практическая апробация методики проведена на базе ряда вычислительных узлов конвейерного типа. В качестве примера реализован конфигурируемый конвейер с совмещением вычислений результата умножения с накоплением и поворота вектора на базе алгоритма CORDIC. Синтезированный конвейер позволяет в режиме коммутации функциональных узлов вычислить пару значений синуса, косинуса или умножать независимые 32-разрядные операнды.

При оценке достижимой тактовой частоты цифровых схем на базе FPGA используется понятие logic levels (логические уровни), которое означает количество последовательно соединенных узлов в цепи максимальной длины. Эта цепь является ограничивающим фактором, где достижимая тактовая частота предварительно оценивается как системная тактовая частота, деленная на показатель logic levels. При системной тактовой частоте порядка 700–750 МГц для современных ПЛИС с архитектурой FPGA достижение logic levels, равного 1, представляет собой достаточно сложную техническую задачу. Тем не менее, проведенная балансировка конвейера позволила получить период тактового сигнала 1.6–1.7 нс, что соответствует тактовой частоте 600–625 МГц, для платформы *AMD/Xilinx Kria*, выполненной по технологическим нормам 16 нм FinFET⁶.

ЗАКЛЮЧЕНИЕ

Рассмотренные в статье подходы позволяют проводить оптимизацию конвейерных вычислителей, предназначенных для работы в составе СБИС. Полученные положительные результаты позволяют расширить методику для проектирования процессорных узлов и конвейеров с повторением операций, архитектурные шаблоны которых были рассмотрены во вводной части статьи. Совместный анализ проекта на нескольких уровнях (программной модели, схемотехнического и топологического представлений) позволяет проводить оптимизацию вычислителя в соответствии с выбираемыми критериями качества, в т.ч. управляемым образом повышать тактовую частоту для высокопроизводительных вычислительных систем за счет балансировки задержек функциональных узлов конвейера.

⁶ <https://www.xilinx.com/products/som/kria/k26c-commercial.html>. Дата обращения 10.10.2023. / Accessed October 10, 2023.

БЛАГОДАРНОСТИ

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № FSFZ-2022-0004 «Архитектуры специализированных вычислительных комплексов, методики, алгоритмы и инструменты проектирования цифровых вычислительных устройств»).

Вклад авторов. Все авторы в равной степени внесли свой вклад в исследовательскую работу.

ACKNOWLEDGMENTS

The work was performed within the framework of the State assignment of the Ministry of Science and Higher Education of the Russian Federation (theme No. FSFZ-2022-0004 “Architectures of specialized computing complexes, methods, algorithms, and tools for designing digital computing devices”).

Authors' contribution. All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ

1. Hennessy J.L., Patterson D.A. A new golden age for computer architecture: Domain-specific hardware/software co-design, enhanced security, open instruction sets, and agile chip development. In: *Proceedings of the 2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture (ISCA)*. IEEE; 2018. <https://doi.org/10.1109/ISCA.2018.00011>
2. Hennessy J.L., Patterson D.A. *Computer Architecture: A Quantitative Approach*. 6th ed. The Morgan Kaufmann Series in Computer Architecture and Design. Morgan Kaufmann; 2017. 936 p.
3. Сесин И.Ю., Болбаков Р.Г. Сравнительный анализ методов оптимизации программного обеспечения для борьбы с предикацией ветвлений на графических процессорах. *Russian Technological Journal*. 2021;9(6):7–15. <https://doi.org/10.32362/2500-316X-2021-9-6-7-15>
4. Слепцов В.В., Афонин В.Л., Аблаева А.Е., Динь Б. Разработка информационно-измерительной и управляющей системы квадрокоптера. *Russian Technological Journal*. 2021;9(6):26–36. <https://doi.org/10.32362/2500-316X-2021-9-6-26-36>
5. Смирнов А.В. Оптимизация характеристик цифровых фильтров одновременно в частотной и временной областях. *Russian Technological Journal*. 2020;8(6):63–77. <https://doi.org/10.32362/2500-316X-2020-8-6-63-77>
6. Умняшкин С.В. *Основы теории цифровой обработки сигналов*. 6-е изд. М.: Litres; 2022. 551 с. ISBN 978-5-4576-1810-7
7. Abadi M., Barham P., Chen J., et al. TensorFlow: A system for Large-Scale Machine Learning. In: *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*. USENIX Association; 2016. P. 265–283.
8. Nurvitadhi E., Sheffield D., Sim J., et al. Accelerating Binarized Neural Networks: Comparison of FPGA, CPU, GPU, and ASIC. In: *2016 International Conference on Field-Programmable Technology (FPT)*. IEEE; 2016. P. 77–84. <https://doi.org/10.1109/FPT.2016.7929192>
9. Советов П.Н. Синтез линейных программ для стековой машины. *Высокопроизводительные вычислительные системы и технологии*. 2019;3(1):17–22.
10. Ахо А.В., Лам М.С., Сети Р., Ульман Д.Д. *Компиляторы: принципы, технологии и инструментарий*: пер. с англ. М.: Вильямс; 2018. ISBN 978-5-8459-1332-8
11. Пратт Т., Зелкович М. *Языки программирования: разработка и реализация*: пер. с англ. СПб.: Питер; 2002. 688 с.
12. Тарасов И.Е., Потехин Д.С., Хренов М.А., Советов П.Н. Автоматизация проектирования многопроцессорной системы на базе ПЛИС для управления во встраиваемых приложениях. *Экономика и менеджмент систем управления*. 2017;25(3–1):179–185.
13. Huang S., Wu K., Jeong H., Wang C., Chen D., Hwu W.M. PyLog: An Algorithm-Centric Python-Based FPGA Programming and Synthesis Flow. *IEEE Trans. Comput.* 2021;70(12):2015–2028. <https://doi.org/10.1109/TC.2021.3123465>
14. Jiang S., Pan P., Ou Y., Batten C. PyMTL3: A Python Framework for Open-Source Hardware Modeling, Generation, Simulation, and Verification. *IEEE Micro*. 2020;40(4):58–66. <https://doi.org/10.1109/MM.2020.2997638>
15. Oishi R., Kadomoto J., Irie H., Sakai S. FPGA-based Garbling Accelerator with Parallel Pipeline Processing. *IEICE Transactions on Information and Systems*. 2023;E106-D(12):1988–1996. <https://doi.org/10.1587/transinf.2023PAP0002>

REFERENCES

1. Hennessy J.L., Patterson D.A. A new golden age for computer architecture: Domain-specific hardware/software co-design, enhanced security, open instruction sets, and agile chip development. In: *Proceedings of the 2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture (ISCA)*. IEEE; 2018. <https://doi.org/10.1109/ISCA.2018.00011>
2. Hennessy J.L., Patterson D.A. *Computer Architecture: A Quantitative Approach*. 6th ed. The Morgan Kaufmann Series in Computer Architecture and Design. Morgan Kaufmann; 2017. 936 p.
3. Sesin I.Yu., Bolbakov R.G. Comparative analysis of software optimization methods in context of branch predication on GPUs. *Russ. Technol. J*. 2021;9(6):7–15 (in Russ.). <https://doi.org/10.32362/2500-316X-2021-9-6-7-15>

- Sleptsov V.V., Afonin V.L., Ablaeva A.E., Dinh B. Development of an information measuring and control system for a quadcopter. *Russ. Technol. J.* 2021;9(6):26–36 (in Russ.). <https://doi.org/10.32362/2500-316X-2021-9-6-26-36>
- Smirnov A.V. Optimization of digital filters performances simultaneously in frequency and time domains. *Russ. Technol. J.* 2020;8(6):63–77 (in Russ.). <https://doi.org/10.32362/2500-316X-2020-8-6-63-77>
- Umnyashkin S.V. *Osnovy teorii tsifrovoi obrabotki signalov (Fundamentals of the Theory of Digital Signal Processing)*. 3rd ed. Moscow: Litres; 2022. 551 p. (in Russ.). ISBN 978-5-4576-1810-7
- Abadi M., Barham P., Chen J., et al. TensorFlow: A system for Large-Scale Machine Learning. In: *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*. USENIX Association; 2016. P. 265–283.
- Nurvitadhi E., Sheffield D., Sim J., et al. Accelerating Binarized Neural Networks: Comparison of FPGA, CPU, GPU, and ASIC. In: *2016 International Conference on Field-Programmable Technology (FPT)*. IEEE; 2016. P. 77–84. <https://doi.org/10.1109/FPT.2016.7929192>
- Sovetov P.N. Synthesis of linear programs for a stack machine. *Vysokoproizvoditel'nye vychislitel'nye sistemy i tekhnologii = High-Performance Computing Systems and Technologies*. 2019;3(1):17–22 (in Russ.).
- Aho A.V., Lam M.S., Sethi R., Ullman J.D. *Kompilyatory: printsipy, tekhnologii i instrumentarii (Compilers: Principles, Techniques, & Tools)*: transl. from Engl. Moscow: Vil'yams; 2018. 1184 p. ISBN 978-5-8459-1932-8 (in Russ.). [Aho A.V., Lam M.S., Sethi R., Ullman J.D. *Compilers: Principles, Techniques, & Tools*. Pearson Addison Wesley; 2007. 1035 p.]
- Pratt T.W., Zelkowitz M.V. *Yazyki programmirovaniya: razrabotka i realizatsiya (Programming Languages. Design and Implementation)*: transl. from Engl. St. Petersburg: Piter; 2002. 688 p. (in Russ.). [Pratt T.W., Zelkowitz M.V. *Programming Languages. Design and Implementation*. Prentice Hall; 2001. 649 p.]
- Tarasov I.E., Potekhin D.S., Khrenov M.A., Sovetov P.N. Computer-aided design of multicore system for embedded applications. *Ekonomika i Menedzhment Sistem Upravleniya*. 2017;25(3–1):179–185 (in Russ.).
- Huang S., Wu K., Jeong H., Wang C., Chen D., Hwu W.M. PyLog: An Algorithm-Centric Python-Based FPGA Programming and Synthesis Flow. *IEEE Trans. Comput.* 2021;70(12):2015–2028. <https://doi.org/10.1109/TC.2021.3123465>
- Jiang S., Pan P., Ou Y., Batten C. PyMTL3: A Python Framework for Open-Source Hardware Modeling, Generation, Simulation, and Verification. *IEEE Micro*. 2020;40(4):58–66. <https://doi.org/10.1109/MM.2020.2997638>
- Oishi R., Kadomoto J., Irie H., Sakai S. FPGA-based Garbling Accelerator with Parallel Pipeline Processing. *IEICE Transactions on Information and Systems*. 2023;E106-D(12):1988–1996. <https://doi.org/10.1587/transinf.2023PAP0002>

Об авторах

Тарасов Илья Евгеньевич, д.т.н., доцент, заведующий лабораторией специализированных вычислительных систем, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: tarasov_i@mirea.ru. Scopus Author ID 57213354150, SPIN-код РИНЦ 4628-7514, <http://orcid.org/0000-0001-6456-4794>

Советов Петр Николаевич, к.т.н., старший научный сотрудник, лаборатория специализированных вычислительных систем, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: sovetov@mirea.ru. Scopus Author ID 57221375427, SPIN-код РИНЦ 9999-1460. <http://orcid.org/0000-0002-1039-2429>

Люлява Даниил Вячеславович, младший научный сотрудник, лаборатория специализированных вычислительных систем, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: lyulyava@mirea.ru. Scopus Author ID 58811698000, SPIN-код РИНЦ 1882-0989, <http://orcid.org/0009-0009-9623-7777>

Мирзоян Дмитрий Ильич, старший научный сотрудник, лаборатория специализированных вычислительных систем, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: mirzoyan@mirea.ru. Scopus Author ID 57432027000, ResearcherID JJE-7844-2023, SPIN-код РИНЦ 8135-9802, <http://orcid.org/0009-0002-4703-8340>

About the authors

Ilya E. Tarasov, Dr. Sci. (Eng.), Associated Professor, Head of the Laboratory of Specialized Computing Systems, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: tarasov_i@mirea.ru. Scopus Author ID 57213354150, RSCI SPIN-code 4628-7514, <http://orcid.org/0000-0001-6456-4794>

Peter N. Sovetov, Cand. Sci. (Eng.), Senior Researcher, Laboratory of Specialized Computing Systems, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: sovetov@mirea.ru. Scopus Author ID 57221375427, RSCI SPIN-code 9999-1460. <http://orcid.org/0000-0002-1039-2429>

Daniil V. Lulyava, Junior Researcher, Laboratory of Specialized Computing Systems, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: lyulyava@mirea.ru. Scopus Author ID 58811698000, RSCI SPIN-code 1882-0989, <http://orcid.org/0009-0009-9623-7777>

Dmitry I. Mirzoyan, Senior Researcher, Laboratory of Specialized Computing Systems, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: mirzoyan@mirea.ru. Scopus Author ID 57432027000, ResearcherID JJE-7844-2023, RSCI SPIN-code 8135-9802, <http://orcid.org/0009-0002-4703-8340>

Современные радиотехнические и телекоммуникационные системы
Modern radio engineering and telecommunication systems

УДК 621.396, 621.371, 621.372
<https://doi.org/10.32362/2500-316X-2024-12-3-46-54>
EDN QREZGM



НАУЧНАЯ СТАТЬЯ

Влияние шумов и помех на вероятность битовых ошибок в системах 5G, использующих банк фильтров с несколькими несущими со смещенной квадратурной амплитудной модуляцией

А. Андрун[®],
О.В. Тихонова

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[®] Автор для переписки, e-mail: abed.androon@yandex.com

Резюме

Цели. Целью работы являются оценка помехоустойчивости метода модуляции сигналов в сетях 5G с использованием банка фильтров с несколькими несущими со смещенной квадратурной амплитудной модуляцией (FBMC-OQAM) и анализ вероятности битовых ошибок.

Методы. В работе применяются методы теории вероятностей и математической статистики, а также методы компьютерного моделирования.

Результаты. Представлен анализ вероятности битовых ошибок метода модуляции сигналов в сетях 5G с использованием банка фильтров с несколькими несущими со смещенной квадратурной амплитудной модуляцией в условиях шумов и исследована устойчивость метода к различным типам помех, таких как внутрисотовые и межсотовые помехи, межлучевые помехи в канале 5G, а также аддитивный белый гауссов шум. Представлены графические и численные данные, полученные компьютерным моделированием, показывающие улучшение вероятности битовых ошибок в сетях 5G, использующих FBMC-OQAM. Проведен сравнительный анализ вероятности ошибки в системе FBMC-OQAM при различных типах шумов и помех, подчеркивающий влияние этих факторов на качество передачи информации.

Выводы. Метод FBMC-OQAM характеризуется малым влиянием на вероятность ошибки системы передачи данных в сетях 5G таких типов помех, как внутрисотовые и межсотовые помехи, межлучевые помехи и нелинейные искажения. В статье подчеркивается необходимость дальнейшей оптимизации и разработки алгоритмов для улучшения вероятности ошибки в системе FBMC-OQAM в реальных условиях сетей 5G. Результаты исследования могут быть использованы при разработке сетей 5G.

Ключевые слова: сеть 5G, метод FBMC-OQAM, белый гауссов шум, вероятность битовой ошибки, отношение сигнал/шум, межсотовые помехи, нелинейные искажения

• Поступила: 04.05.2023 • Доработана: 18.11.2023 • Принята к опубликованию: 12.03.2024

Для цитирования: Андрун А., Тихонова О.В. Влияние шумов и помех на вероятность битовых ошибок в системах 5G, использующих банк фильтров с несколькими несущими со смещенной квадратурной амплитудной модуляцией. *Russ. Technol. J.* 2024;12(3):46–54. <https://doi.org/10.32362/2500-316X-2024-12-3-46-54>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Impacts of noise and interference on the bit error rate of the FBMC-OQAM modulation scheme in 5G systems

Abed Androon[@],
Olga V. Tikhonova

MIREA – Russian Technological University, Moscow, 119454 Russia

[@] Corresponding author, e-mail: abed.androon@yandex.com

Abstract

Objectives. The work sets out to evaluate the noise immunity of the signal modulation method in 5G networks using a filter bank multicarrier with offset quadrature amplitude modulation (FBMC-OQAM) and to analyze the bit error rate (BER).

Methods. In the work, probability theory and mathematical statistics methods are applied according to computer modeling approaches.

Results. An analysis of BER for the signal modulation method in 5G networks, which uses a bank of filters with multiple carriers with offset quadrature amplitude modulation under noise conditions, is presented. The resistance of the method to intra-cell, inter-cell, and inter-beam types of interference in the 5G channel, as well as additive white Gaussian noise, is investigated. The graphical and numerical data obtained through computer modeling demonstrates improved BER in 5G networks using FBMC-OQAM. The presented comparative analysis of error probability in the FBMC-OQAM system under various types of noise and interference emphasizes the impact of these factors on the quality of information transmission.

Conclusions. The FBMC-OQAM method is characterized by the low impact on the error probability of the data transmission system in 5G networks of various types of interference including intra-cell and inter-cell interference, inter-beam interference, and nonlinear distortions. However, it will be necessary to further optimize the method and develop algorithms for enhancing error probability in the FBMC-OQAM system under real conditions in 5G networks. The research results can be used in the development of 5G networks.

Keywords: 5G network, FBMC-OQAM method, white gaussian noise, BER, SNR, inter-cell interference, nonlinear distortion

• Submitted: 04.05.2023 • Revised: 18.11.2023 • Accepted: 12.03.2024

For citation: Androon A., Tikhonova O.V. Impacts of noise and interference on the bit error rate of the FBMC-OQAM modulation scheme in 5G systems. *Russ. Technol. J.* 2024;12(3):46–54. <https://doi.org/10.32362/2500-316X-2024-12-3-46-54>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Ожидается, что пятое поколение систем беспроводной связи, широко известное как 5G, произведет революцию в способах взаимодействия людей. Обещая более высокую скорость передачи данных, меньшую задержку и большую пропускную способность, 5G обладает потенциалом для преобразования широкого спектра отраслей, включая здравоохранение, транспорт и производство. Однако успех 5G зависит от его способности надежно работать в различных средах и при различных типах помех и шумов. Одной из ключевых технологий, используемых в 5G для решения этих проблем, является ортогональное частотное мультиплексирование с использованием банка фильтров (filter-bank multicarrier with offset quadrature amplitude modulation, FBMC-OQAM) [1]. OQAM (смещенная квадратурная амплитудная модуляция) – метод модуляции, при котором синфазная и квадратурная составляющие сигнала расположены в шахматном порядке, т.е. смещены относительно друг друга. Это смещение позволяет более эффективно использовать доступную полосу частот, уменьшает межсимвольные помехи и повышает устойчивость к частотно-селективным замираниям. FBMC-OQAM имеет несколько преимуществ перед традиционным методом ортогонального частотного разделения каналов, широко используемым в 4G и более ранних беспроводных системах связи. Например, метод FBMC-OQAM имеет более высокую спектральную эффективность и более устойчив к помехам и шумам. Тем не менее, вероятность ошибки в системе, использующей FBMC-OQAM, в условиях различных типов помех и шумов исследована недостаточно. Для проектирования сетей 5G такие исследования должны быть проведены.

Целью данной статьи является анализ вероятности битовой ошибки (ВБО) в системах 5G, использующих метод FBMC-OQAM, при различных типах помех и шумов. Результаты этого анализа дают информацию о вероятности ошибки в системе FBMC-OQAM в реальных средах связи и помогают определить облик будущих систем 5G.

1. ТИПЫ ШУМОВ И ПОМЕХ В СОТОВОЙ СИСТЕМЕ 5G

Системы «5G Новое радио»¹ должны обеспечить улучшенную спектральную эффективность, более высокую пропускную способность, увеличение скорости передачи данных и надежности связи по сравнению с предыдущими поколениями сотовых сетей [1]. Ресурсы спектра, используемого

в системах 5G и определенные в протоколе 3GPP (3rd Generation Partnership Project), разделены на два диапазона частот (ДЧ): ДЧ1 < 6 ГГц и ДЧ2 > 24 ГГц (миллиметровый диапазон) [2]. Диапазон частот ДЧ1 почти полностью занят и имеет ограниченные ресурсы, которые можно использовать для сети 5G, тогда как ДЧ2 представляет собой часть спектра, где большая часть частот пока не задействована и может быть легко использована для будущих сотовых сетей [3].

Ниже под помехами понимается нежелательное воздействие на передаваемый сигнал. Помехи обычно приводят к изменению или искажению сигнала, они могут возникать от различных источников, включая соседние сигналы, электромагнитные поля, перекрытия сигналов и препятствия на пути распространения сигнала [4]. Под шумовыми помехами понимается широкополосное воздействие на сигнал, действующее в течение длительного времени. В системах 5G шум (шумовые помехи) может возникать от различных источников, включая электронные компоненты и тепловые эффекты. Помехи ухудшают качество сигнала, снижая пропускную способность и вызывая ошибки при передаче данных [5]. Они оказываются основной проблемой при освоении нового и использовании существующего участков спектра современными системами 5G [6].

В беспроводной сотовой сети с малыми ячейками многоуровневые помехи определяются специфическими особенностями каждого узла малой мощности: он непрерывно генерирует и принимает нежелательные сигналы от различных близлежащих источников [7]. Наиболее распространенными помехами, связанными с радиосетями, являются собственные помехи, помехи по смежному каналу, внутри- и межсотовые помехи. Тем не менее, мобильная сеть не ограничивается только этими помехами. Каждая сеть подвержена воздействию помех, возникающих в зависимости от конкретного сценария ее развертывания.

1.1. Внутрисотовые и межсотовые помехи

Межсотовые помехи являются одной из существенных причин ухудшения работы сети. Когда пользователи двух соседних ячеек пытаются одновременно использовать одну и ту же полосу частот, то возникают помехи [8]. Более того, межсотовые помехи сильно влияют на качество связи пользователей на краях ячейки, потому что пользователь получает сигнал от макробазовой станции своей ячейки и от соседней ячейки из-за повторного использования частоты (рис. 1). Искажения, вызываемые дополнительным оборудованием в пределах одной и той же ячейки, называются внутрисотовыми помехами.

¹ <https://www.3gpp.org/technologies/5g-system-overview>. Дата обращения 20.01.2023. / Accessed January 20, 2023.

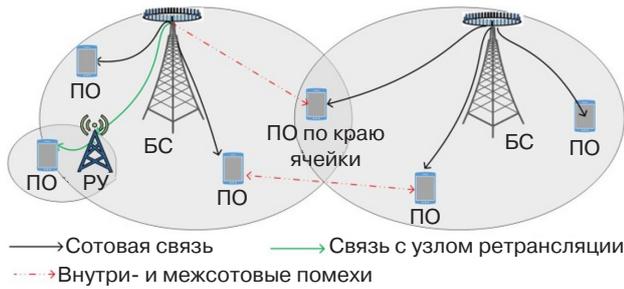


Рис. 1. Внутрисотовые и межсотовые помехи. БС – базовая станция, ПО – пользовательское оборудование, РУ – ретрансляционный узел

Межканальные помехи возникают, когда сигналы двух и более отдельных каналов мешают друг другу, как показано на рис. 2 (на горизонтальной оси отложена частота f (Гц), на вертикальной оси – мощность сигнала (дБ); f_1 – центральная частота канала 1; f_2 – центральная частота канала 2), при наличии множества устройств беспроводной связи, работающих на близком расстоянии. Из-за этого передатчик мощного сигнала создает помехи приемнику слабого сигнала.

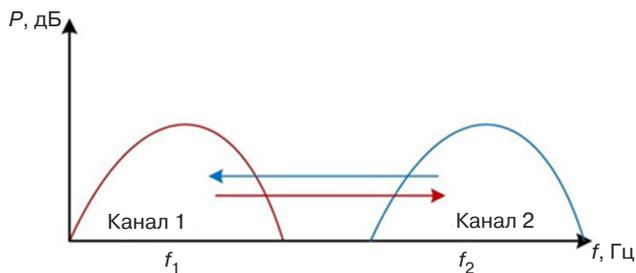


Рис. 2. Межканальные помехи

1.2. Межлучевые помехи

Формирование луча – это новая технология, используемая в современной сотовой связи, которая определяет лучший маршрут и обеспечивает оптимальную пропускную способность для конкретного пользователя в определенном направлении. Такой подход необходим для компенсации потерь на затухание при передаче сигнала, особенно при связи на миллиметровых волнах. Базовая станция (БС) генерирует многократные узкие лучи радиочастотных сигналов во всех направлениях зоны покрытия. Антенны БС и/или мобильного терминала настраиваются таким образом, чтобы сосредоточить передаваемый сигнал в определенном направлении, образуя так называемый «луч» или направленный сигнал [9]. Однако пространственное разделение нескольких лучей приводит к межлучевой помехе [10]. Она вызвана соседними лучами БС той же или соседней ячейки, как показано на рис. 3.

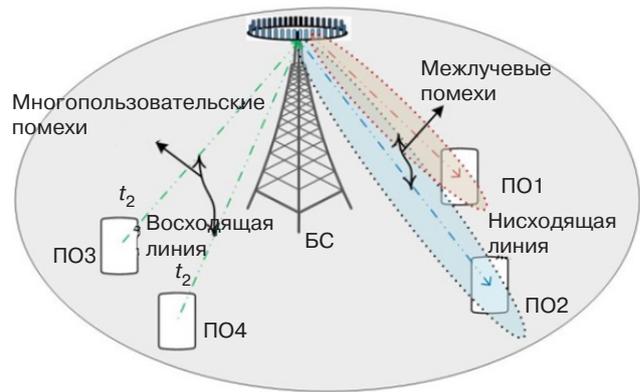


Рис. 3. Межлучевые и многопользовательские помехи. t_2 – путь передачи данных

1.3. Нелинейные искажения в сетях 5G

Нелинейные искажения являются распространенной проблемой в беспроводных коммуникационных системах, в т.ч. в сетях 5G. Эти искажения возникают, когда передаваемый сигнал значительно усиливается, и его мощность становится слишком высокой, в результате чего усилитель переходит в нелинейный режим [11]. При этом сигнал искажается, и декодирование информации в приемнике происходит некорректно.

В системах 5G нелинейные искажения могут возникать в цепях передатчика или приемника из-за множества факторов и влиять на характеристики передаваемого сигнала [12]. В контексте FBMC нелинейные искажения могут вызывать интермодуляционные искажения, когда передаваемые сигналы смешиваются друг с другом и генерируют дополнительные нежелательные частоты, что приводит к возрастанию уровня помех между поднесущими. Следствием этого является увеличение частоты битовых ошибок и снижение эффективности передачи данных.

Выходной сигнал $y(t)$, прошедший через усилитель с нелинейными искажениями, может быть представлен с помощью разложения в ряд Тейлора:

$$y(t) = a_0x(t) + a_1x^2(t) + a_2x^3(t), \quad (1)$$

где $x(t)$ – входной сигнал, a_0 – линейный коэффициент усиления усилителя мощности, a_1 и a_2 – коэффициенты при нелинейных членах разложения.

1.4. Шумовая помеха

Вероятность появления ошибки в канале связи зависит от уровня шума. Аддитивный белый гауссов шум (АБГШ) возникает из многих естественных источников – от движения атомов в проводнике до излучения от земли и космических объектов. Канал АБГШ – самая простая модель, которая хорошо

подходит для беспроводной связи. Этот канал линеен и не зависит от времени [13]. Модулированные сигналы проходят через него без каких-либо потерь амплитуды и фазовых искажений. Выходной сигнал канала определяется как

$$y(t) = x(t) + n(t), \quad (2)$$

где слагаемое $n(t)$ – шум, имеющий гауссово распределение с нулевым средним значением и дисперсией в качестве мощности шума, а $x(t)$ – передаваемый сигнал.

Канал АБГШ позволяет разработчикам оценить влияние различных факторов на вероятность появления ошибки в системе, в частности, оценить влияние схем модуляции, методов канального кодирования и алгоритмов исправления ошибок [14], параметров физического уровня, таких как несущая частота, полоса пропускания и мощность передачи.

2. ВБО И ОТНОШЕНИЕ СИГНАЛ/ШУМ

ВБО – это показатель точности передачи цифровых данных в системе связи. Она определяется как отношение количества принятых битов с ошибкой к общему количеству битов, переданных по каналу связи, выраженное в процентах или десятичной дробью. ВБО – это вероятность того, что бит в передаваемом сигнале будет принят с ошибкой. На величину ВБО могут влиять различные факторы, такие как шумы, помехи, схема модуляции и дальность передачи.

Важным параметром, определяющим качество работы системы передачи информации, является отношение сигнал/шум (signal-to-noise ratio, SNR), обычно выражаемое в децибелах:

$$\text{SNR} = 10 \lg \frac{P_s}{P_n}, \quad (3)$$

где P_s – мощность сигнала, а P_n – мощность шума.

С практической точки зрения высокое значение SNR предпочтительно для любой системы связи, поскольку оно обеспечивает более высокую точность и надежность передачи информации. С другой стороны, низкое значение SNR может вызвать ошибки в передаваемых данных, что может привести к снижению качества системы связи.

3. МОДЕЛИРОВАНИЕ И АНАЛИЗ РЕЗУЛЬТАТОВ

Эта статья основана на ранее опубликованном исследовании авторов и продолжает анализ модели, представленной в работах [15, 16].

Модель состоит из передатчика и приемника FBMC-OQAM сигнала, где имитируются сам

сигнал и возмущающие воздействия: белый гауссов шум, паразитные составляющие соседних каналов, внутрисотовые и межсотовые помехи, межлучевые помехи и нелинейные искажения. В модели имитируется деление сигнала на несколько поднесущих, каждая со своим собственным узкополосным фильтром, что обеспечивает эффективное использование частоты и высокую спектральную эффективность. Используется смещенная квадратурная модуляция, при которой каждая поднесущая разделяется на два параллельных потока – действительную (I – in-phase) и мнимую (Q – quadrature) части, что помогает уменьшить межсимвольные помехи, вызванные перекрытием соседних поднесущих. С использованием этой модели можно получить численные оценки ВБО в зависимости от SNR для различных условий.

В настоящей статье представлены новые результаты, полученные с помощью описанной модели и ее дальнейшего анализа, в частности, изучения влияния различных шумов и помех в сетях 5G.

Система способна обрабатывать I-данные и Q-данные различного размера и с различной длиной быстрого преобразования Фурье (БПФ). Были проведены эксперименты с использованием разных комбинаций параметров, чтобы продемонстрировать адаптивность системы к различным условиям.

Для верификации модели было проведено сравнение экспериментально полученных результатов ВБО с известной теоретически полученной ВБО при идентичных условиях канала АБГШ. На рис. 4 представлена зависимость ВБО от величины SNR для системы FBMC-OQAM, полученная с помощью моделирования и подтверждающая работоспособность модели [17].

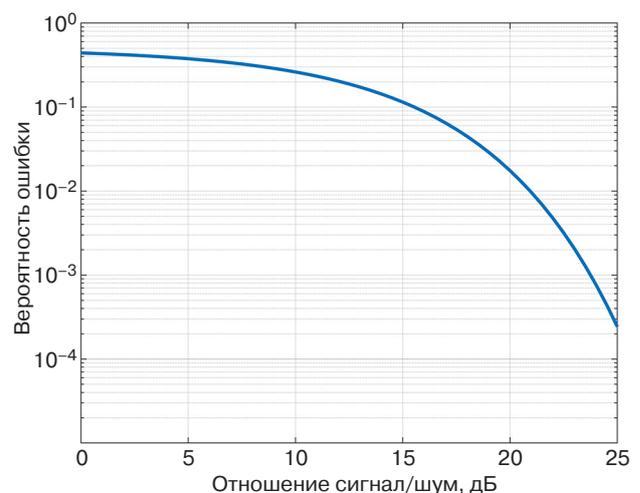


Рис. 4. Зависимость вероятности ошибки от отношения сигнал/шум, полученная с помощью моделирования

Результаты моделирования получены при значениях параметров, которые приведены в таблице.

Таблица. Параметры модели

Количество поднесущих	2048
Размер данных I/Q	$N = 64, 32, 16$
Длина кадров данных	1
Размер БПФ	8192, 4096, 2048, 512
Тип фильтра	Фильтр-прототип, используемый в FBMC-OQAM
Коэффициент перекрытия	$K = 4$
Модель канала	Случайная исходная последовательность
Тип шума	АБГШ
Тип помехи	Помехи от других систем сотовой связи, нелинейные искажения, межсотовые помехи
Диапазон SNR	От -10 дБ до 30 дБ

Исследуется вероятность появления ошибки в системе для 512-точечного, 2048-точечного, 4096-точечного и 8192-точечного БПФ. В системе используется согласованная фильтрация с коэффициентом перекрытия K , представляющим количество символов мультинесущей, которые перекрываются на временном интервале. В этом исследовании анализируется влияние изменения длины символа для 512-точечного, 1024-точечного и 2048-точечного БПФ. Результаты показывают, что изменение длины символов практически не влияет на вероятность появления ошибки в системе FBMC-OQAM.

Основываясь на результатах (рис. 5) для $N = 64$, можно заметить, что с уменьшением длины БПФ SNR, требуемое для достижения ВБО значения, равного 10^{-6} , увеличивается. В частности, когда длина БПФ равна 8192, системе требуется SNR = 6 дБ для достижения желаемого значения ВБО, в то время как для длин БПФ 4096, 2048 и 512 требуемые значения SNR составляют 7.8, 12 и 20 дБ соответственно. Это значит, что большая длина БПФ обеспечивает лучшую величину ВБО при более низких

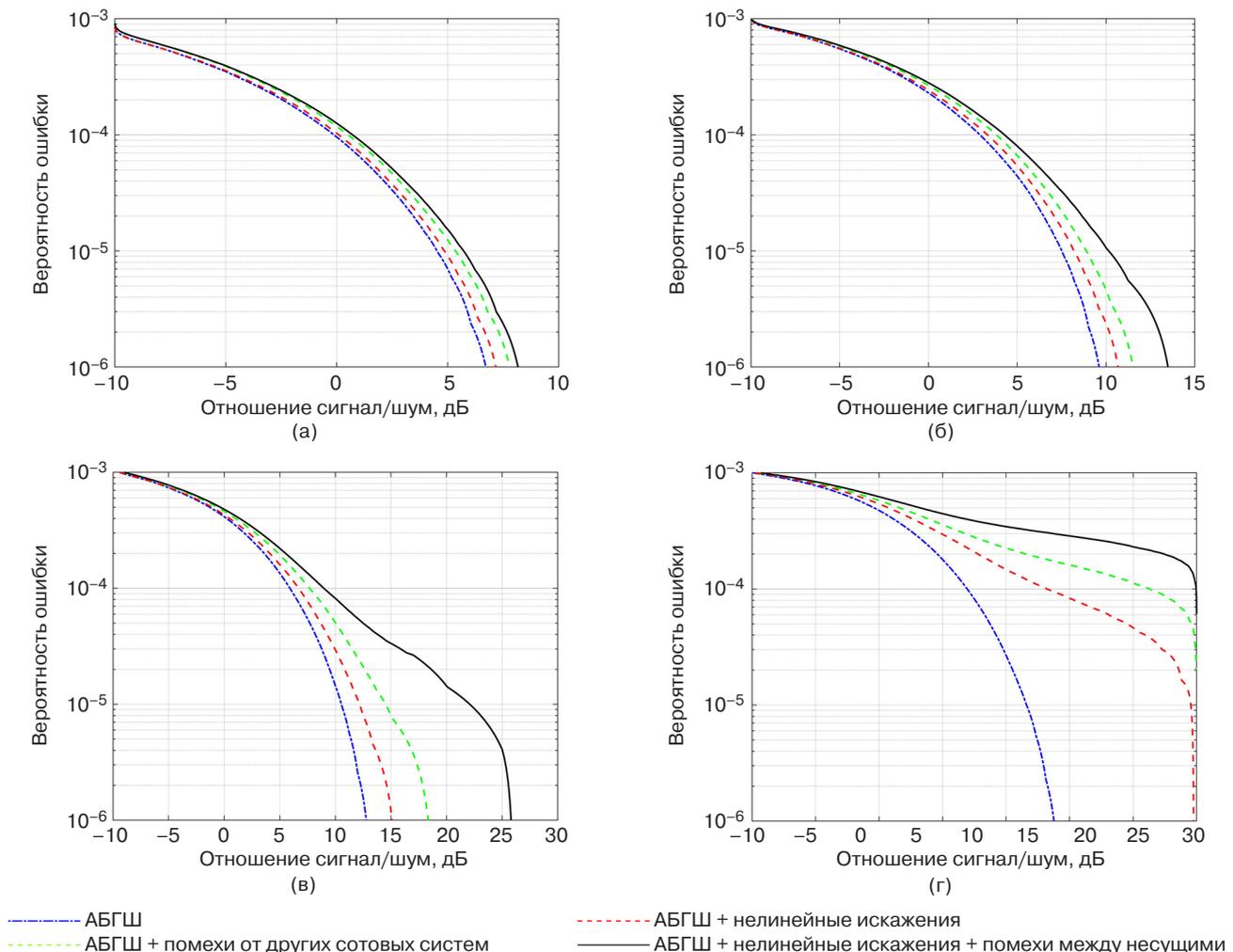


Рис. 5. Результаты моделирования при $N = 64$ и длине БПФ: (а) 8192, (б) 4096, (в) 2048, (г) 512

значениях SNR. Следовательно, увеличение длины БПФ может быть полезной стратегией для снижения вероятности ошибки.

На рис. 6 показаны результаты моделирования для $N = 32$ с различной длиной БПФ. Можно заметить, что с уменьшением длины БПФ увеличивается SNR, необходимое для достижения значения ВБО, равного 10^{-6} . В частности, когда длина БПФ равна 8192, системе требуется SNR = 1.5 дБ для достижения желаемой величины ВБО, в то время как для длин БПФ 4096, 2048 и 512 требуемые значения SNR составляют 5, 8 и 14 дБ соответственно. Следовательно, и в этом случае увеличение длины БПФ может быть полезной стратегией для уменьшения вероятности ошибки.

Из рис. 7 для $N = 16$ можно заметить, что тенденция увеличения требуемого SNR для достижения того же уровня ВБО по мере уменьшения длины БПФ становится еще более заметной. Та же зависимость прослеживается на рис. 5 и 6. Видно,

что ВБО уменьшается по мере увеличения SNR для всех длин БПФ. Кроме того, система с наибольшей длиной БПФ, равной 8192, имеет низкую ВБО, требуя самого высокого SNR для достижения значения ВБО, равного 10^{-6} . С другой стороны, система с наименьшей длиной БПФ, равной 512, обладает наилучшей ВБО, требуя наименьшего SNR для достижения той же величины ВБО. Системы с длиной БПФ 4096 и 2048 имеют схожие вероятности ошибки, при этом для первой требуется несколько более высокое SNR для достижения того же уровня, что и для второй. Следовательно, основываясь на этих результатах, можно сделать вывод, что для этой системы с фиксированным $N = 16$ меньшие длины БПФ приводят к лучшей ВБО при более высоких уровнях SNR, в то время как большие длины БПФ требуют более высокого SNR для достижения той же ВБО. Это подчеркивает важность выбора подходящей длины БПФ в данной системе.

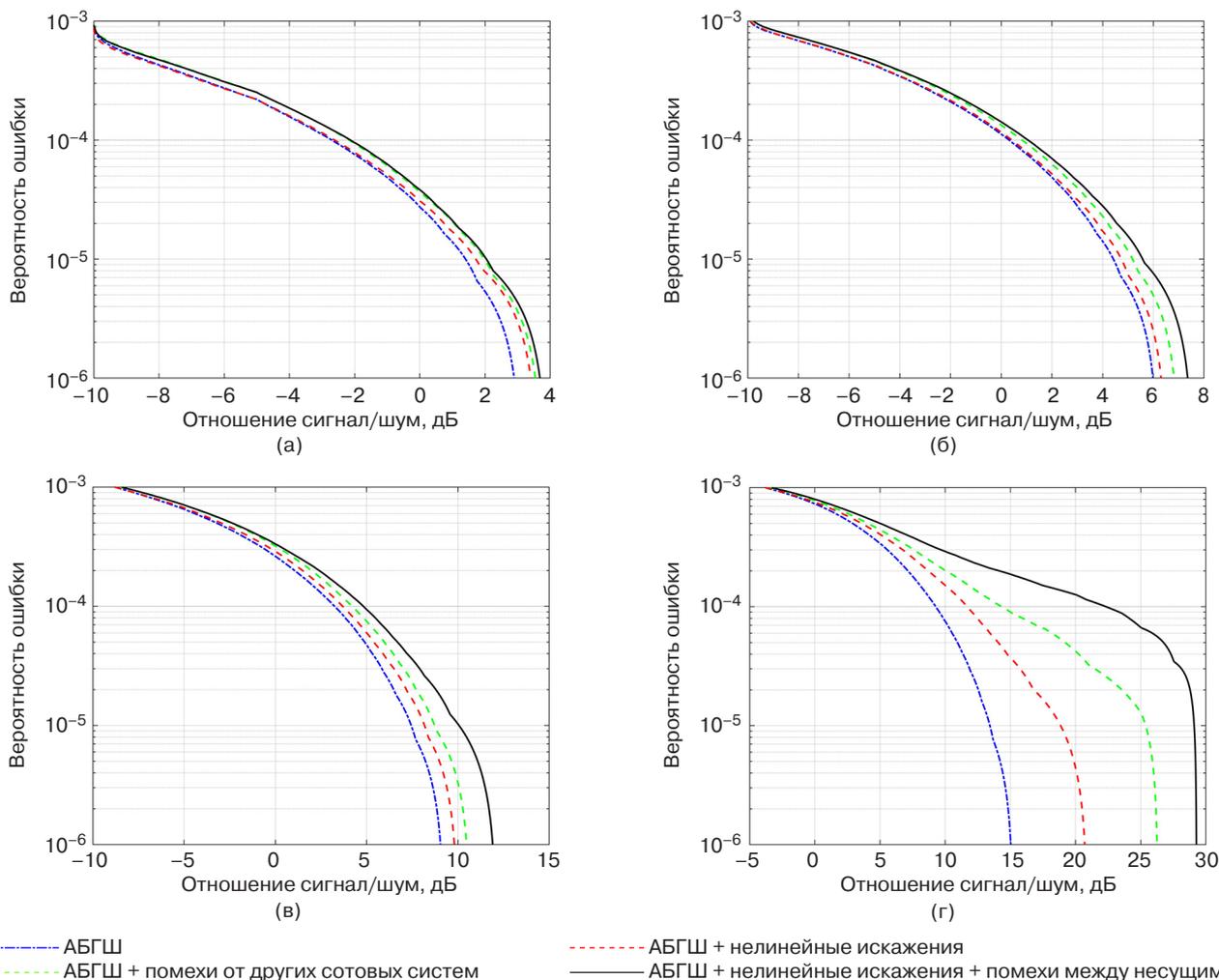


Рис. 6. Результаты моделирования при $N = 32$ и длине БПФ: (а) 8192, (б) 4096, (в) 2048, (г) 512

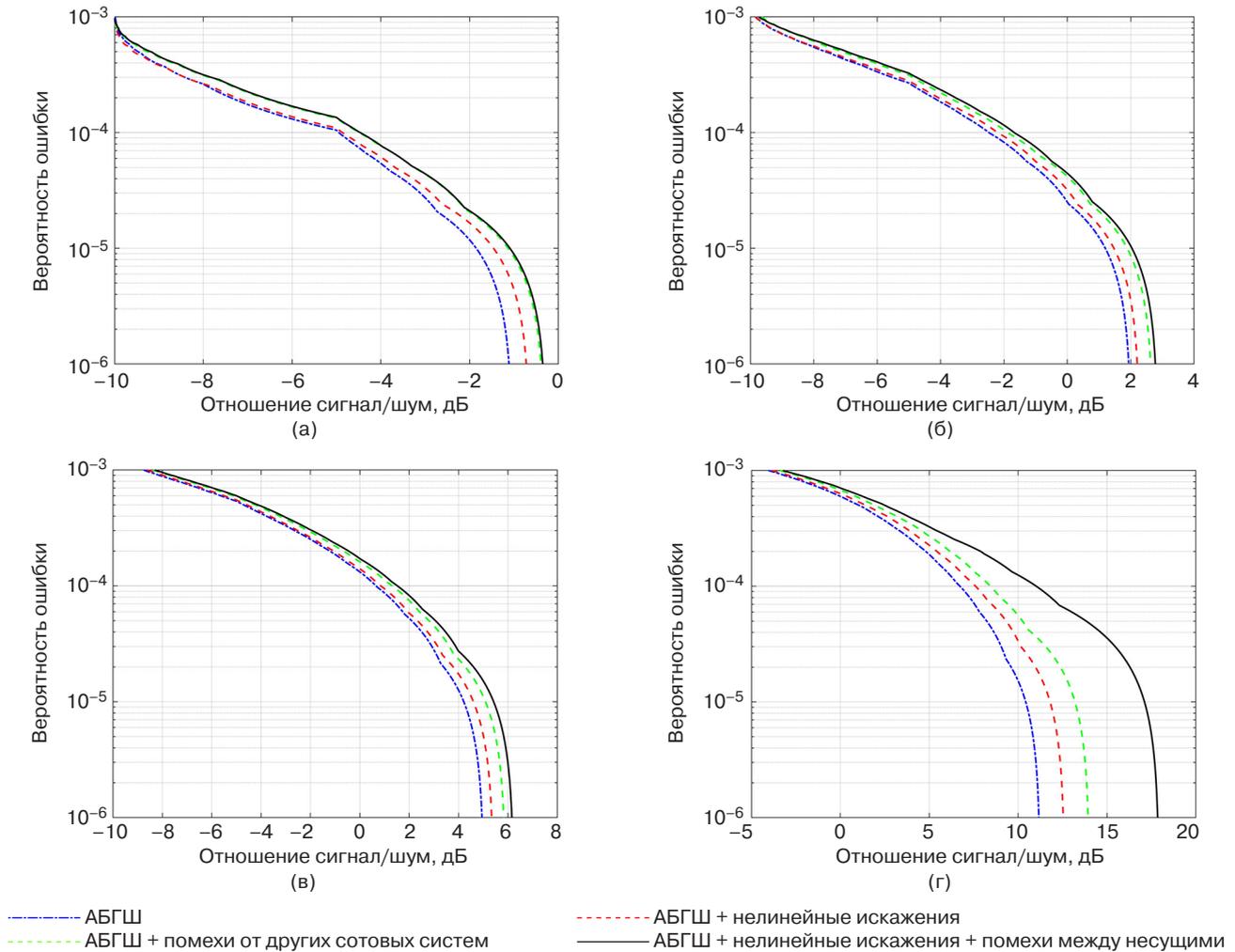


Рис. 7. Результаты моделирования при $N = 16$ и длине БПФ: (а) 8192, (б) 4096, (в) 2048, (г) 512

ЗАКЛЮЧЕНИЕ

Результаты моделирования подтверждают высокую устойчивость метода модуляции FBMC-OQAM в сетях 5G к различным видам помех, что существенно улучшает качество связи. Увеличение длины БПФ эффективно снижает вероятность ошибок передачи

данных, а оптимальный выбор длины БПФ зависит от конкретных условий сетей 5G.

Вклад авторов. Все авторы в равной степени внесли свой вклад в исследовательскую работу.

Authors' contribution. All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Qamar F., Hindia M.N., Abbas T., Dimiyati K.B., Amiri I.S. Investigation of QoS performance evaluation over 5G network for indoor environment at millimeter wave bands. *Int. J. Electron. Telecommun.* 2019;65(1):95–101.
2. Sanfilippo G., Galinina O., Andreev S., Pizzi S., Araniti G. A concise review of 5G new radio capabilities for directional access at mmWave frequencies. In: Galinina O., Andreev S., Balandin S., Koucheryavy Y. (Eds.). *Internet of Things, Smart Spaces, and Next Generation Networks and Systems. NEW2AN ruSMART 2018. Lecture Notes in Computer Science.* 2018. P. 340–354. Correction: https://doi.org/10.1007/978-3-030-01168-0_65
3. Dilli R. Analysis of 5G Wireless Systems in FR1 and FR2 Frequency Bands. In: *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*. 2020. IEEE. P. 767–772. <https://doi.org/10.1109/ICIMIA48430.2020.9074973>

4. Rebato M., Mezzavilla M., Rangan S., Boccardi F., Zorzi M. Understanding noise and interference regimes in 5G millimeter wave cellular networks. In: *European Wireless 2016; 22th European Wireless Conference*. 2016; VDE. P. 84–88. <https://doi.org/10.48550/arXiv.1604.05622>
5. Liu J., Sheng M., Liu L., Li J. Effect of densification on cellular network performance with bounded pathloss model. *IEEE Communications Letters*. 2016;21(2):346–349. <https://doi.org/10.1109/LCOMM.2016.2615298>
6. Liu J., Sheng M., Liu L., Li J. Interference management in ultra-dense networks: Challenges and approaches. *IEEE Network*. 2017;31(6):70–77. <https://doi.org/10.1109/MNET.2017.1700052>
7. Shende N., Gurbuz O., Erkip E. Half-duplex or full-duplex relaying: A capacity analysis under self-interference. In: *2013 47th Annual Conference on Information Sciences and Systems (CISS)*. 2013; IEEE, 6 p. <https://doi.org/10.1109/CISS.2013.6552276>
8. Qamar F., Dimiyati K.B., Hindia M.N., Noordin K.A.B., Al-Samman A.M. A comprehensive review on coordinated multipoint operation for LTE-A. *Computer Networks*. 2017;123:19–37. <https://doi.org/10.1016/j.comnet.2017.05.003>
9. Guo Z., Fei Y. On the Cross Link Interference of 5G with Flexible Duplex and Full Duplex. In: *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. 2020; IEEE. 4 p. <https://doi.org/10.1109/WCNCW48565.2020.9124866>
10. Xue Q., Li B., Zuo X., Yan Z., Yang M. Cell capacity for 5G cellular network with inter-beam interference. In: *2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. 2016; IEEE. 5 p. <https://doi.org/10.1109/ICSPCC.2016.7753608>
11. El Ghzaoui M., Hmamou A., Foshi J., Mestoui J. Compensation of non-linear distortion effects in MIMO-OFDM systems using constant envelope OFDM for 5G applications. *J. Circuits, Syst. Comput.* 2020;29(16):2050257. <https://doi.org/10.1142/S0218126620502576>
12. Mathews A.B., Mathews A.B., Kumar C.A. A Non-Linear Improved CNN Equalizer with Batch Gradient Decent in 5G Wireless Optical Communication. *IETE Journal of Research*. 2023. P. 1–13. <https://doi.org/10.1080/03772063.2022.2163930>
13. Kaur N., Kansal L. Performance Comparison of MIMO Systems over AWGN and Rician Channels with Zero Forcing Receivers. *International Journal of Wireless & Mobile Networks (IJWMN)*. 2013;5(1):73–84. <http://doi.org/10.5121/ijwmn.2013.5106>
14. Mezher M., AlAbbas A.R. BER performance of Reed-Solomon codes with 16 PSK modulation over *in-vivo* radio channel. In: *Proc. International Conference on Electrical, Computer and Energy Technologies (ICECET)*. 2022; IEEE. 6 p. <https://doi.org/10.1109/ICECET55527.2022.9873485>
15. Андрун А., Тихонова О.В. Преобразование цифровых данных для передачи по сетям 5G. В сб.: *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем «Радиоинфоком-2021»: Сборник научных статей V Международной научно-практической конференции*. М.: РТУ МИРЭА; 2021. С. 22–26. [Androon A., Tikhonova O.V. Digital data conversion for transmission over 5G networks. In: *Actual Problems and Prospects for the Development of Radio Engineering and Infocommunication Systems Radioinfocom 2021: Collection of scientific articles of the 5th International Scientific and Practical Conference*. Moscow: RTU MIREA; 2021. P. 22–26 (in Russ.).]
16. Андрун А., Тихонова О.В. Метод обработки сигнала в современных сетях беспроводной связи 5G. *Вопросы электромеханики. Труды ВНИИЭМ*. 2023;192(1):21–26. [Androon A., Tikhonova O.V. Signal processing method in modern wireless networks 5G. *Voprosy elektromekhaniki. Trudy VNIIEEM = Electromechanical Matters. VNIIEEM Studies*. 2023;192(1):21–26 (in Russ.).]
17. Kumar A., Bharti S. Design and performance analysis of OFDM and FBMC modulation techniques. *The Scientific Bulletin of Electrical Engineering Faculty*. 2017;17(2):30–34. <https://doi.org/10.1515/sbeef-2017-0007>

Об авторах

Андрун Абед, аспирант, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: abed.androon@yandex.com. <https://orcid.org/0000-0001-9624-2060>

Тихонова Ольга Вадимовна, д.т.н., старший научный сотрудник, профессор, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: o_tikhonova@inbox.ru. Scopus Author ID 57208923772, SPIN-код РИНЦ 3362-9924, <https://orcid.org/0009-0009-4013-9182>

About the authors

Abed Androon, Postgraduate Student, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: abed.androon@yandex.com. <https://orcid.org/0000-0001-9624-2060>

Olga V. Tikhonova, Dr. Sci. (Eng.), Senior Researcher, Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: o_tikhonova@inbox.ru. Scopus Author ID 57208923772, RSCI SPIN-code 3362-9924, <https://orcid.org/0009-0009-4013-9182>

Микро- и нанoeлектроника. Физика конденсированного состояния
Micro- and nanoelectronics. Condensed matter physics

УДК 537.632

<https://doi.org/10.32362/2500-316X-2024-12-3-55-64>

EDN SWVVUI



НАУЧНАЯ СТАТЬЯ

Моделирование магниторефрактивного эффекта в нанокompозитах $\text{Co-Al}_2\text{O}_3$ в рамках приближения Бруггемана

М.А. Мухутдинова,
А.Н. Юрасов[@]

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: alexey_yurasov@mail.ru

Резюме

Цели. Цель работы – изучить магниторефрактивный эффект (МРЭ) в нанокompозитах, заключающийся в изменении коэффициентов отражения, пропускания и поглощения света образцов с большим магнитосопротивлением (МС) при их намагничивании. Существует ряд материалов, обладающих большой магнитооптической активностью и значительным МС. К таким материалам относятся магнитные нанокompозиты. Они представляют из себя материалы на основе полимерной матрицы, в которую включены неорганические магнитные частицы, волокна или слоистые частицы, с нанометровыми размерами от 1 до 100 нм хотя бы в одном измерении. Главной целью создания таких нанокompозитов является совмещение нескольких компонентов с их особыми свойствами в одном материале. Наличие в таких материалах гигантского, колоссального и туннельного МС, гигантского аномального эффекта Холла представляет практический интерес. Данные материалы применяют для магнитной записи, модуляции света, как приемники теплового излучения, а сам МРЭ является перспективным методом неразрушающего контроля любых наноструктур, например, для измерения МС.

Методы. Для описания оптики и магнитооптики дисперсных сред рассмотрена теория эффективной среды, благодаря которой можно решить задачу определения комплексной диэлектрической проницаемости среды через диэлектрические проницаемости составляющих ее компонент или наоборот. В статье этот подход рассматривался на примере нанокompозита $\text{Co-Al}_2\text{O}_3$ с концентрацией ферромагнитного металла Co , равной 0.4, вблизи порога перколяции. Для изучения рассмотрен именно этот случай, т.к. вблизи порога перколяции кардинально меняются все свойства нанокompозитов.

Результаты. Используя приближение Бруггемана (effective medium approximation, ЕМА) для описания оптических и магнитооптических свойств нанокompозитов на примере $\text{Co-Al}_2\text{O}_3$, авторы получили характеристики МРЭ, а именно: изменение МРЭ на отражение и пропускание света при нормальном падении и при угле падения вблизи угла Брюстера (ниже порога перколяции) или главного угла падения для металлов (выше порога перколяции), что усиливает МРЭ. Преимущество ЕМА заключается в возможности исследовать магнитооптические спектры в диапазоне средних объемных концентраций металлической компоненты.

Выводы. Полученные значения хорошо соответствуют известным экспериментальным данным. Важно отметить, что данный подход позволяет исследовать любые наноструктуры.

Ключевые слова: магниторефрактивный эффект, нанокompозиты, магнитосопротивление, тензор диэлектрической проницаемости, приближение Бруггемана

• Поступила: 16.05.2023 • Доработана: 16.10.2023 • Принята к опубликованию: 08.04.2024

Для цитирования: Мухутдинова М.А., Юрасов А.Н. Моделирование магниторефрактивного эффекта в нанокompозитах Co-Al₂O₃ в рамках приближения Бруггемана. *Russ. Technol. J.* 2024;12(3):55–64. <https://doi.org/10.32362/2500-316X-2024-12-3-55-64>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Modeling of the magnetorefractive effect in Co-Al₂O₃ nanocomposites in the framework of the Bruggeman approximation

Muza A. Mukhutdinova,
Alexey N. Yurasov[@]

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: alexey_yurasov@mail.ru

Abstract

Objectives. To investigate the magnetorefractive effect (MRE) in nanocomposites, which consists in changing the reflection, transmittance and light absorption coefficients of samples with large magnetoresistance (MR) upon their magnetization. Materials offering high magneto-optical activity and significant MR include magnetic nanocomposites. These materials are based on a polymer matrix, which includes inorganic magnetic particles, fibers or layered particles, whose nanometer sizes range from 1 to 100 nm in at least one dimension. The main purpose of creating such nanocomposites is to combine the special properties of several components in one material. The presence in such materials of gigantic, colossal and tunneling MR, as well as the giant anomalous Hall effect, is of practical interest. Uses range from magnetic recording, light modulation, and receivers for thermal radiation, while the MRE itself is a promising method for the non-destructive testing of any nanostructures, e.g., measuring MR.

Methods. The use of effective medium theory to describe the optics and magneto-optics of dispersed media provides a means to determine the complex permittivity of a medium through the permittivity of its constituent components or vice versa. The present work considers the example of a Co-Al₂O₃ nanocomposite with a concentration of ferromagnetic metal Co 0.4 near the percolation threshold. This particular case was considered for study, since all the properties of nanocomposites change dramatically near the percolation threshold.

Results. Using the Bruggeman effective medium approximation (EMA) to describe the optical and magneto-optical properties of nanocomposites on the example of Co-Al₂O₃, the characteristics of MRE are obtained, namely, the change in MRE for reflection and transmission of light at normal incidence and at the angle of incidence near the Brewster angle (below the percolation threshold) or the main angle of incidence for metals (above the percolation threshold), which enhances MRE. The advantage of the EMA is the ability to study magneto-optical spectra in the range of average volume concentrations of the metal component.

Conclusions. The obtained values correspond well to the known experimental data. Moreover, the described approach can be used to study any nanostructures.

Keywords: magnetorefractive effect, nanocomposites, magnetoresistance, dielectric permittivity tensor, Bruggeman approximation

• Submitted: 16.05.2023 • Revised: 16.10.2023 • Accepted: 08.04.2024

For citation: Mukhutdinova M.A., Yurasov A.N. Modeling of the magnetorefractive effect in Co-Al₂O₃ nanocomposites in the framework of the Bruggeman approximation. *Russ. Technol. J.* 2024;12(3):55–64. <https://doi.org/10.32362/2500-316X-2024-12-3-55-64>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Магниторефрактивный эффект (МРЭ) заключается в изменении коэффициентов отражения $R\left(\frac{\Delta R}{R}\right)$, пропускания $T\left(\frac{\Delta T}{T}\right)$ и поглощения A света у образцов с большим магнитосопротивлением (МС) при их намагничивании [1].

Нанокompозиты – перспективные многофазные материалы, получаемые при введении наночастиц в матрицу основного материала с геометрическими размерами частиц, лежащими в диапазоне от 1 до 100 нм [1]. Нанокompозиты обладают значительным МС $\frac{\Delta\rho}{\rho}$, где ρ – электрическое сопротивление. Существенные возможности открываются при применении нанокompозитов в области магнитооптики, которая изучает явления, возникающие в магнитном поле, в результате взаимодействия оптического излучения с веществом. Например, нанокompозиты применяют для измерения МС. Также магнитные нанокompозиты, представляющие из себя неоднородные магнетики, в которых ферромагнитные частицы помещены в металлическую или диэлектрическую матрицу, используются для магнитной записи, модуляции света, как приемники теплового излучения [1–7]. В свою очередь магнитооптические методы являются методами неразрушающего контроля любых наноструктур [1, 2].

Целью работы является исследование МРЭ в нанокompозитах вблизи порога перколяции.

МЕТОДИКА РАСЧЕТОВ

Под влиянием магнитного поля изменяются дисперсионные кривые показателя преломления n и коэффициента поглощения k , приводя к появлению или изменению оптической анизотропии среды. При изменении МС меняется значение коэффициентов отражения и пропускания. Коэффициент отражения (формулы Френеля) при нормальном падении имеет вид:

$$R = \frac{(1-n)^2 + k^2}{(1+n)^2 + k^2}, \quad (1)$$

где n – коэффициент рефракции, k – коэффициент экстинкции нанокompозита [1].

Формула (1) используется для расчета величины $\frac{\Delta R}{R}$ [2]:

$$\frac{\Delta R}{R} = -(1-R) \cdot \frac{\Delta\rho}{\rho} \cdot k^2 \times \frac{3n^2 - k^2 - 1}{(n^2 + k^2)((1-n)^2 + k^2)} \cdot 100\%. \quad (2)$$

При угле падения, отличном от нуля, коэффициент отражения для полубесконечной среды рассчитывают по формулам:

$$R = |r_{12}|^2, \quad (3)$$

$$r_{12} = \frac{g_1\eta_2^2 - g_2n_1^2}{g_1\eta_2^2 + g_2n_1^2}, \quad (4)$$

где $g_1 = \sqrt{n_1 - n_1(\sin\varphi_0)^2}$, $g_2 = \sqrt{n_2 - n_2(\sin\varphi_0)^2}$, φ_0 – угол падения света на поверхность нанокompозита, $\eta_2 = n_2 - k_2i$ – комплексный показатель преломления нанокompозита, n_1 – коэффициент преломления среды, из которой падает свет [2].

Для коэффициента пропускания при нормальном падении света имеем [2]:

$$T = |t_{12}|^2, \quad (5)$$

$$t_{12} = \frac{g_1^2\eta_2}{g_1\eta_2^2 + g_2n_1^2}, \quad (6)$$

$$\frac{\Delta T}{T} = \frac{1}{2} \cdot \frac{\Delta\rho}{\rho} \cdot Tk^2 \cdot \left(\frac{2n^2 + n}{n^2 + k^2}\right) \cdot 100\%. \quad (7)$$

В связи с тем, что нанокompозит является сложным материалом, для описания его электрических, оптических и магнитных свойств рассматривается эффективный тензор диэлектрической проницаемости (ТДП):

$$\tilde{\epsilon}^{\text{eff}} = \begin{pmatrix} \epsilon_{xx}^{\text{eff}} & i\gamma^{\text{eff}} & 0 \\ -i\gamma^{\text{eff}} & \epsilon_{xx}^{\text{eff}} & 0 \\ 0 & 0 & \epsilon_{xx}^{\text{eff}} \end{pmatrix}, \quad (8)$$

где диагональные компоненты $\epsilon_{xx}^{\text{eff}} = (\epsilon_{xx}^{\text{eff}})' - i(\epsilon_{xx}^{\text{eff}})''$ – оптическая составляющая, а недиагональные

компоненты $\gamma^{\text{eff}} = (\gamma^{\text{eff}})' - i(\gamma^{\text{eff}})''$ – магнитооптическая составляющая ТДП [4–8].

Наиболее интересен случай, когда нанокompозит находится вблизи порога перколяции, в этой области происходит значительное усиление магнитооптических эффектов. В связи с этим все расчеты проводились, когда объемная концентрация $X = 0.4$, что соответствует близости к порогу перколяции.

Воспользуемся приближением эффективной среды Бруггемана (effective medium approximation, ЕМА) при $0.3 < X < 0.7$, которое хорошо работает для описания нанокompозитов при средней концентрации магнитной (металлической) компоненты [7–10].

Чтобы найти ε^{eff} и γ^{eff} , используем следующие уравнения:

$$\begin{aligned}
 & X \frac{(\varepsilon_1 - \varepsilon^{\text{EMA}})}{\varepsilon^{\text{EMA}} + (\varepsilon_1 - \varepsilon^{\text{EMA}})L_{xx}} + \\
 & + (1 - X) \frac{(\varepsilon_0 - \varepsilon^{\text{EMA}})}{\varepsilon^{\text{EMA}} + (\varepsilon_0 - \varepsilon^{\text{EMA}})L_{xx}} = 0, \\
 & X \frac{(\gamma^{\text{EMA}} - \gamma)}{[\varepsilon^{\text{EMA}} + (\varepsilon_1 - \varepsilon^{\text{EMA}})L_{xx}]^2} + \\
 & + (1 - X) \frac{\gamma^{\text{EMA}}}{[\varepsilon^{\text{EMA}} + (\varepsilon_0 - \varepsilon^{\text{EMA}})L_{xx}]^2} = 0,
 \end{aligned} \tag{9}$$

где $\varepsilon^{\text{EMA}} = \varepsilon^{\text{eff}}$, $\gamma^{\text{EMA}} = \gamma^{\text{eff}}$, $L_{xx} = \frac{1-L}{2} = \frac{1-\frac{1}{3}}{2} = \frac{1}{3}$ –

форм-фактор частиц ($L = \frac{1}{3}$ для сферических частиц), $\varepsilon_0 = \varepsilon'_0 - i\varepsilon''_0$ – диэлектрическая проницаемость неметаллической (неферромагнитной) составляющей, $\varepsilon_1 = \varepsilon'_1 - i\varepsilon''_1$ – диэлектрическая проницаемость для ферромагнитной компоненты (диагональная часть соответствующего ТДП), $\gamma = \gamma' - i\gamma''$ – недиагональная компонента ТДП ферромагнитной составляющей [2].

РЕЗУЛЬТАТЫ РАСЧЕТОВ

В настоящей работе рассчитываются абсолютные значения МРЭ на отражении $\frac{\Delta R}{R}$ и пропускании $\frac{\Delta T}{T}$ при углах падения $\varphi_1 = 0^\circ$ и $\varphi_2 = 70^\circ$, а также эффективный ТДП при различных длинах волн инфракрасного (ИК) диапазона (от 0.8 до 10 мкм), падающих на поверхность нанокompозит Co-Al₂O₃. Выбор угла $\varphi_2 = 70^\circ$ обусловлен близостью его к углу Брюстера, если нанокompозит находится в состоянии ниже порога перколяции, или близостью к главному углу падения для металлов, если нанокompозит находится в состоянии выше порога перколяции, что обуславливает усиление оптических эффектов.

Результаты расчетов приведены в табл. 1–6. Здесь E – энергия падающей электромагнитной волны.

Таблица 1. Абсолютные значения коэффициентов отражения $\frac{\Delta R}{R}$ при нормальном падении света $\frac{\Delta \rho}{\rho} = 1\%$

E , эВ	λ , мкм	$n_1(\text{Al}_2\text{O}_3)$	$n_2(\text{Co})$	$k_2(\text{Co})$	$\frac{\Delta R}{R}$, %		
					$\varphi_1 = 0^\circ$		
					$\frac{\Delta \rho}{\rho} = 1\%$	$\frac{\Delta \rho}{\rho} = 5\%$	$\frac{\Delta \rho}{\rho} = 10\%$
1.550	0.8	1.76	1.90	4.95	0.117	0.583	1.167
1.240	1	1.76	2.20	5.50	0.099	0.495	0.989
0.620	2	1.74	5.15	7.00	-0.069	-0.344	-0.687
0.413	3	1.71	4.90	8.45	0.0006	0.003	0.006
0.310	4	1.68	4.70	11.00	0.043	0.214	0.429
0.248	5	1.62	4.70	14.70	0.045	0.225	0.450
0.207	6	1.56	5.00	17.50	0.039	0.195	0.389
0.177	7	1.46	5.40	20.90	0.033	0.163	0.325
0.155	8	1.32	5.80	24.00	0.028	0.140	0.280
0.138	9	1.15	6.56	27.20	0.025	0.123	0.247
0.124	10	0.85	7.10	29.50	0.023	0.114	0.228

Таблица 2. Абсолютные значения коэффициентов отражения $\frac{\Delta R}{R}$ при падении света под углом 70°

E, эВ	λ , мкм	$n_1(\text{Al}_2\text{O}_3)$	$n_2(\text{Co})$	$k_2(\text{Co})$	$\frac{\Delta R}{R}$, %		
					$\varphi_2 = 70^\circ$		
					$\frac{\Delta \rho}{\rho} = 1\%$	$\frac{\Delta \rho}{\rho} = 5\%$	$\frac{\Delta \rho}{\rho} = 10\%$
1.550	0.8	1.76	1.90	4.95	-0.277	-1.383	-2.767
1.240	1	1.76	2.20	5.50	-0.281	-1.406	-2.811
0.620	2	1.74	5.15	7.00	-0.351	-1.756	-3.512
0.413	3	1.71	4.90	8.45	-0.300	-1.500	-3.001
0.310	4	1.68	4.70	11.00	-0.226	-1.131	-2.262
0.248	5	1.62	4.70	14.70	-0.154	-0.771	-1.541
0.207	6	1.56	5.00	17.50	-0.122	-0.612	-1.224
0.177	7	1.46	5.40	20.90	-0.095	-0.473	-0.947
0.155	8	1.32	5.80	24.00	-0.076	-0.379	-0.758
0.138	9	1.15	6.56	27.20	-0.064	-0.322	-0.645
0.124	10	0.85	7.10	29.50	-0.059	-0.294	-0.588

Таблица 3. Абсолютные значения коэффициентов пропускания $\frac{\Delta T}{T}$ при нормальном падении света

E, эВ	λ , мкм	$n_1(\text{Al}_2\text{O}_3)$	$n_2(\text{Co})$	$k_2(\text{Co})$	$\frac{\Delta T}{T}$, %		
					$\varphi_1 = 0^\circ$		
					$\frac{\Delta \rho}{\rho} = 1\%$	$\frac{\Delta \rho}{\rho} = 5\%$	$\frac{\Delta \rho}{\rho} = 10\%$
1.550	0.8	1.76	1.90	4.95	1.378	6.888	13.777
1.240	1	1.76	2.20	5.50	1.513	7.567	15.134
0.620	2	1.74	5.15	7.00	1.763	8.817	17.635
0.413	3	1.71	4.90	8.45	2.023	10.114	20.228
0.310	4	1.68	4.70	11.00	2.211	11.053	22.107
0.248	5	1.62	4.70	14.70	2.243	11.213	22.427
0.207	6	1.56	5.00	17.50	2.269	11.347	22.694
0.177	7	1.46	5.40	20.90	2.298	11.491	22.982
0.155	8	1.32	5.80	24.00	2.331	11.653	23.306
0.138	9	1.15	6.56	27.20	2.442	12.208	24.416
0.124	10	0.85	7.10	29.50	2.518	12.591	25.182

Таблица 4. Абсолютное значение коэффициентов пропускания $\frac{\Delta T}{T}$ при падении света под углом 70°

E, эВ	λ , мкм	$n_{1(Al_2O_3)}$	$n_{2(Co)}$	$k_{2(Co)}$	$\frac{\Delta T}{T}$, %		
					$\varphi_2 = 70^\circ$		
					$\frac{\Delta \rho}{\rho} = 1\%$	$\frac{\Delta \rho}{\rho} = 5\%$	$\frac{\Delta \rho}{\rho} = 10\%$
1.550	0.8	1.76	1.90	4.95	-0.026	-0.128	-0.255
1.240	1	1.76	2.20	5.50	-0.023	-0.114	-0.228
0.620	2	1.74	5.15	7.00	-0.013	-0.064	-0.129
0.413	3	1.71	4.90	8.45	-0.012	-0.058	-0.117
0.310	4	1.68	4.70	11.00	-0.009	-0.047	-0.093
0.248	5	1.62	4.70	14.70	-0.007	-0.033	-0.065
0.207	6	1.56	5.00	17.50	-0.005	-0.025	-0.050
0.177	7	1.46	5.40	20.90	-0.004	-0.019	-0.038
0.155	8	1.32	5.80	24.00	-0.003	-0.015	-0.030
0.138	9	1.15	6.56	27.20	-0.002	-0.012	-0.024
0.124	10	0.85	7.10	29.50	-0.002	-0.010	-0.021

Таблица 5. Спектральные значения ϵ^{EMA} при различных величинах λ , рассчитанные с помощью коэффициентов n и k

E, эВ	λ , мкм	$n_{1(Al_2O_3)}$	$n_{2(Co)}$	$k_{2(Co)}$	ϵ_1	ϵ_0	ϵ^{eff}
0.539	2.30	1.73	5.15	7.27	-26.330 - 74.881i	2.993	5.018 - 13.031i
0.729	1.70	1.74	4.6	6.70	-23.730 - 61.640i	3.028	4.666 - 11.622i
1.000	1.24	1.75	3.2	6.10	-26.970 - 39.040i	3.063	2.975 - 9.599i
1.253	0.99	1.76	2.94	5.50	-21.606 - 32.340i	3.098	3.033 - 8.558i
1.494	0.83	1.76	2.53	4.95	-18.102 - 25.047i	3.098	2.762 - 7.487i
1.797	0.69	1.76	2.31	4.45	-14.466 - 20.559i	3.098	2.712 - 6.634i
2.000	0.62	1.77	2.19	4.11	-12.096 - 18.002i	3.133	2.729 - 6.108i
2.296	0.54	1.77	2.05	3.81	-10.314 - 15.621i	3.133	2.661 - 5.606i
2.480	0.50	1.77	1.88	3.55	-9.068 - 13.348i	3.133	2.524 - 5.158i
2.696	0.46	1.78	1.78	3.30	-7.722 - 11.748i	3.168	2.498 - 4.782i
3.024	0.41	1.78	1.61	3.05	-6.710 - 9.821i	3.168	2.343 - 4.361i
3.263	0.38	1.79	1.53	2.82	-5.612 - 8.629i	3.204	2.325 - 4.025i

Таблица 6. Спектральные значения ϵ^{EMA} и γ^{EMA} при различных величинах λ с табличными значениями ϵ и γ

$E, \text{эВ}$	$\lambda, \text{мкм}$	$X = 0.4$			ϵ^{eff}	γ^{eff}
		ϵ_1	ϵ_0	γ		
0.539	2.30	2.993	-23.835 - 57.178i	1.529 - 3.008i	4.381 - 11.119i	0.179 + 0.017i
0.729	1.70	3.028	-18.091 - 44.063i	0.748 - 2.051i	4.213 - 9.450i	0.134 + 0.004i
1.000	1.24	3.063	-13.307 - 31.657i	0.203 - 1.241i	3.857 - 7.761i	0.098 + 0.008i
1.253	0.99	3.098	-11.358 - 24.914i	-0.009 - 0.895i	3.520 - 6.829i	0.083 - 0.007i
1.494	0.83	3.098	-9.474 - 17.882i	-0.135 - 0.530i	3.018 - 5.762i	0.060 - 0.017i
1.797	0.69	3.098	-8.295 - 14.346i	-0.160 - 0.380i	2.739 - 5.160i	0.051 - 0.024i
2.000	0.62	3.133	-7.613 - 12.339i	-0.162 - 0.299i	2.580 - 4.823i	0.045 - 0.028i
2.296	0.54	3.133	-6.507 - 9.779i	-0.144 - 0.199i	2.349 - 4.301i	0.033 - 0.028i
2.480	0.50	3.133	-5.945 - 8.626i	-0.130 - 0.159i	2.240 - 4.042i	0.029 - 0.029i
2.696	0.46	3.168	-5.217 - 7.378i	-0.111 - 0.118i	2.144 - 3.745i	0.023 - 0.029i
3.024	0.41	3.168	-4.407 - 6.281i	-0.088 - 0.084i	2.059 - 3.415i	0.018 - 0.026i
3.263	0.38	3.204	-3.504 - 5.468i	-0.066 - 0.059i	2.058 - 3.103i	0.013 - 0.023i

Анализируя величины в таблицах, можно сделать вывод, что рассчитанные значения эффективного ТДП хорошо совпадают с литературными данными, например, приведенными в [4].

На основе выполненных расчетов построены графики зависимости МРЭ на отражении и пропускании при нормальном падении света и угле падения 70° (рис. 1–4).

Анализируя эти графики, можно сделать главный вывод: спектральные зависимости $\frac{\Delta R}{R}$ и $\frac{\Delta T}{T}$ в ИК-диапазоне линейно коррелируют со значениями МС $\frac{\Delta \rho}{\rho}$, причем зависимость прямо пропорциональная. Чем больше значение МС, тем сильнее изменяются величины $\frac{\Delta R}{R}$ и $\frac{\Delta T}{T}$.

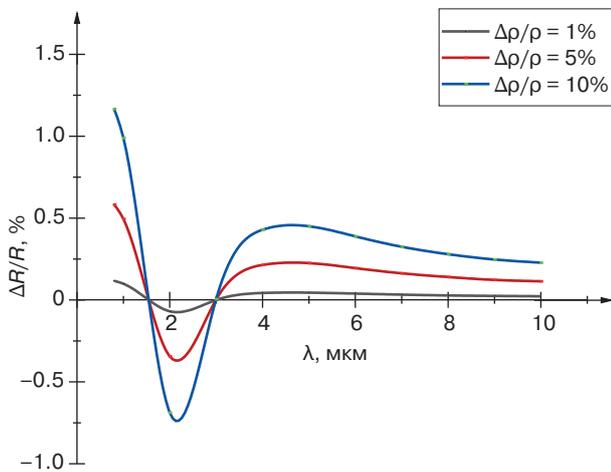


Рис. 1. Спектральная зависимость МРЭ на отражении при нормальном падении света

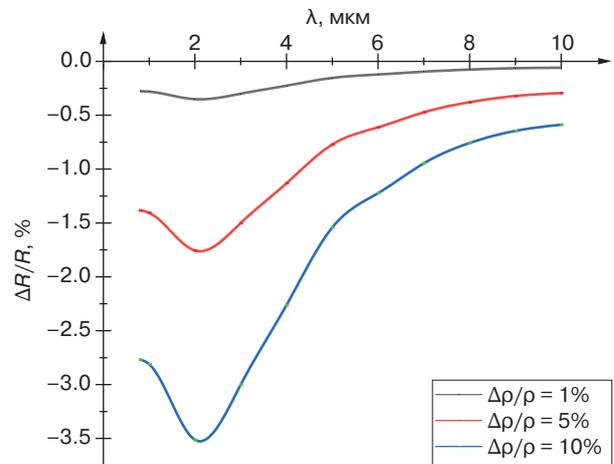


Рис. 2. Спектральная зависимость МРЭ на отражении при $\phi_2 = 70^\circ$

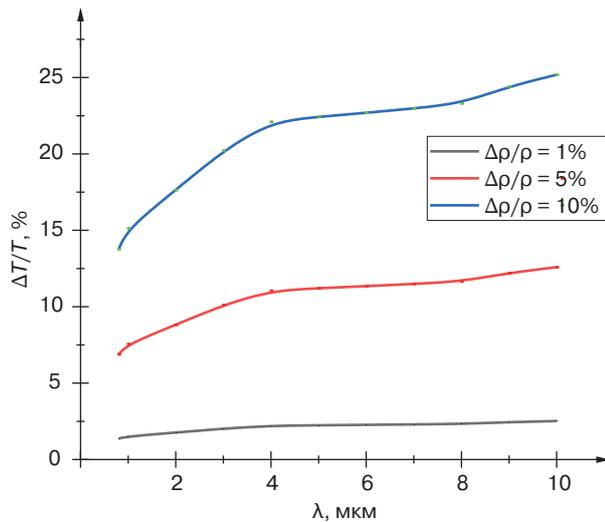


Рис. 3. Спектральная зависимость МРЭ на пропускании при нормальном падении света

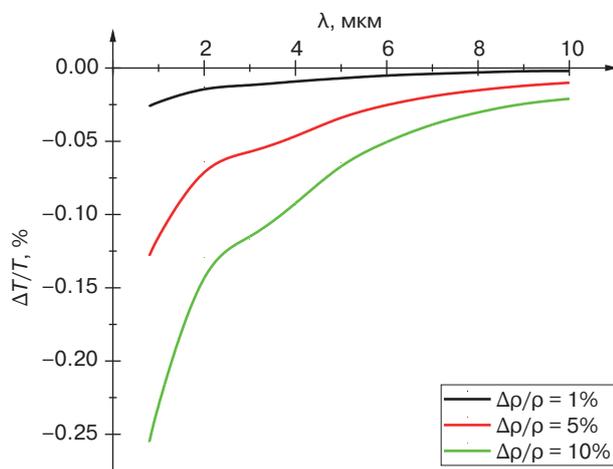


Рис. 4. Спектральная зависимость МРЭ на пропускании при φ₂ = 70°

Из рис. 1 видно, что в ближней ИК-области спектра МРЭ на отражении сильно меняется, что связано с межзонными переходами, которые значительно влияют на оптические характеристики материала в ИК-диапазоне. На рис. 2 можно видеть большую абсолютную величину параметра $\frac{\Delta R}{R}$. Этот факт можно объяснить тем, что в данной области, помимо межзонных переходов, существенную роль играют внутризонные переходы, и угол падения близок к углу Брюстера (главному углу падения для металлов) [3]. На рис. 3 наблюдается плавное увеличение МРЭ на пропускании $\frac{\Delta T}{T}$, а на рис. 4 видно, что

с увеличением длины волны (т.е. с уменьшением E) значения МРЭ на пропускании $\frac{\Delta T}{T}$ сначала резко возрастают, а далее плавно меняются. Изменяя длину волны λ и действительную часть коэффициента преломления n , при данной толщине пленки d нанокompозита можно получить условия интерференции, при которых величина МРЭ существенно возрастает [11–15].

ЗАКЛЮЧЕНИЕ

В работе проведено теоретическое исследование оптических и магнитооптических спектров нанокompозитов на примере Co-Al₂O₃ в рамках теории МРЭ. Расчет спектров проводился с помощью формул (2)–(7). Спектры эффективного ТДП рассчитывались в приближении Бруггемана, т.к. именно этот метод позволяет хорошо описать свойства нанокompозита при средних объемных концентрациях металлической компоненты ($X = 0.4$). Важно отметить, что МРЭ – перспективный инструмент изучения материалов, в первую очередь, как бесконтактный метод измерения МС – неразрушающий метод контроля любых наноструктур, а приведенные в данной статье подходы справедливы для исследования широкого класса наноструктур.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (государственное задание для университетов № ФГФЗ-2023-0005).

ACKNOWLEDGMENTS

The work was supported by the Ministry of Education and Science of the Russian Federation (State task for universities No. FGFZ-2023-0005).

Вклад авторов

М.А. Мухутдинова – обработка литературных источников, компьютерное моделирование, обсуждение результатов, написание статьи.

А.Н. Юрасов – моделирование, обсуждение результатов, написание и редактирование статьи.

Authors' contributions

M.A. Mukhutdinova – literature review, computer simulation, discussion of results, and writing the text of the article.

A.N. Yurasov – computer simulation, discussion of results, writing and editing the article.

СПИСОК ЛИТЕРАТУРЫ

1. Юрасов А.Н. Магниторефрактивный эффект в наноструктурах. *Приборы*. 2022. № 4(262). С. 22–25.
2. Юрасов А.Н. *Магнитооптические эффекты и магниторефрактивный эффект в нанокompозитах*. М.: МИРЭА; 2016. 55 с.
3. Кринчик Г.С., Артемьев В.А. Магнитооптические свойства Ni, Co, и Fe в ультрафиолетовой, видимой и инфракрасной областях спектра. *Журнал экспериментальной и теоретической физики*. 1967;53(6):1901–1912.
4. Niklasson G.A., Granqvist C.G. Optical properties and solar selectivity of coevaporated CoAl₂O₃ composite films. *J. Appl. Phys.* 1984;55(9):3382–3410. <https://doi.org/10.1063/1.333386>
5. Быков И.В., Ганьшина Е.А., Грановский А.Б., Гуцин В.С. Магниторефрактивный эффект в гранулированных пленках с туннельным магнитосопротивлением. *Физика твердого тела*. 2000;42(3):487–491.
6. Гуцин В.С., Ганьшина Е.А., Козлов А.А., Быков И.В. Магниторефрактивный эффект в нанокompозитах. *Вестник Московского университета. Серия 3. Физика. Астрономия*. 2005;1:45–58.
7. Yurasov A., Yashin M., Ganshina E., Granovsky A., Garshin V., Semenova D., Mirzokulov K. Simulation of magneto-optical properties of nanocomposites (CoFeZr)_x(Al₂O₃)_(1-x). *J. Phys.: Conf. Ser.* 2019;1389(1):012113. <http://doi.org/10.1088/1742-6596/1389/1/012113>
8. Апресян Л.А., Власова Т.В., Красовский В.И., Крыштоб В.И., Расмагин С.И. Приближения эффективной среды для описания многокомпонентных композитов. *Журнал технической физики (ЖТФ)*. 2020;90(7):1175–1183. <https://doi.org/10.21883/JTF.2020.07.49453.446-18>
9. Фадеев Е.А., Блинов М.И., Гаршин В.В., Тарасова О.С., Ганьшина Е.А., Прудникова М.В., Прудников В.Н., Ляхдеранта Э., Рыльков В.В., Грановский А.Б. Магнитные свойства нанокompозитов (Co₄₀Fe₄₀B₂₀)_x(SiO₂)_(100-x) вблизи порога перколяции. *Известия РАН. Серия физическая*. 2019;83(7):917–920. <https://doi.org/10.1134/S0367676519070159>
10. Ганьшина Е.А., Припеченков И.М., Перова Н.Н., Каназакова Е.С., Николаев С.Н., Ситников А.С., Грановский А.Б., Рыльков В.В. Магнитооптическая спектроскопия нанокompозитов (CoFeB)_x(LiNbO₃)_(100-x) до порога перколяции: от суперпарамагнетизма и суперферромагнетизма до ферромагнетизма. *Физика металлов и металловедение*. 2023;124(2):134–140.
11. Granovsky A., Sukhorukov Yu., Gan'shina E., Telegin A. Magnetorefractive effect in magnetoresistive materials. In: *Magnetophotonics: From Theory to Applications*. Berlin Heidelberg: Springer. 2013. P. 107–133.
12. Shkurdoda Yu.O., Dekhtyaruk L.V., Basov A.G., Chornous A.M., Shabelnyk Yu.M., Kharchenko A.P., Shabelnyk T.M. The giant magnetoresistance effect in Co/Cu/Co three-layer films. *Journal of Magnetism and Magnetic Materials*. 2019;477:88–91. <https://www.doi.org/10.1016/j.jmmm.2019.01.040>
13. Dekhtyaruk L.V., Kharchenko A.P., Klymenko Yu.O., Shkurdoda Yu.O., Shabelnyk Yu.M., Bezdidko O.V., Chornous A.M. Negative and Positive Effect of Giant Magnetoresistance in The Magnetically Ordered Sandwich. *2020 IEEE 10th International Conference Nanomaterials: Applications & Properties (NAP)*. Sumy, Ukraine. 2020. P. 01NMM13-1-01NMM13-3. <https://www.doi.org/10.1109/NAP51477.2020.9309694>
14. Kelley C.S., Naughton J., Benson E., Bradley R.C., Lazarov V.K., Thompson S.M., Matthew J.A. Investigating the magnetic field-dependent conductivity in magnetite thin films by modelling the magnetorefractive effect. *J. Phys.: Condens. Matter*. 2014;26(3):036002. <https://doi.org/10.1088/0953-8984/26/3/036002>
15. Yurasov A., Yashin M., Ganshina E., et al. Simulation of magneto-optical properties of nanocomposites (CoFeZr)_x(Al₂O₃)_{1-x}. *J. Phys.: Conf. Ser.* 2019;1389:012113. <http://doi.org/10.1088/1742-6596/1389/1/012113>

REFERENCES

1. Yurasov A.N. Magnetorefractive effect in nanostructures. *Pribory = Instruments*. 2022;4(262):22–25 (in Russ.).
2. Yurasov A.N. *Магнитооптические эффекты и магниторефрактивный эффект в нанокompозитах (Magneto-Optical Effects and Magnetorefractive Effect in Nanocomposites)*. Moscow: MIREA; 2016. 55 p. (in Russ.).
3. Krinchik G.S., Artem'ev V.A. Magneto-optical properties of Ni, Co and Fe in the ultraviolet visible and infrared parts of the spectrum. *Journal of Experimental and Theoretical Physics (JETP)*. 1968;26(6):1080–1085. [Original Russian Text: Krinchik G.S., Artem'ev V.A. Magneto-optical properties of Ni, Co and Fe in the ultraviolet visible and infrared parts of the spectrum. *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki*. 1967;53(6):1901–1912 (in Russ.).]
4. Niklasson G.A., Granqvist C.G. Optical properties and solar selectivity of coevaporated CoAl₂O₃ composite films. *J. Appl. Phys.* 1984;55(9):3382–3410. <https://doi.org/10.1063/1.333386>
5. Bykov I.V., Gan'shina E.A., Granovskii A.B., et al. Magnetorefractive effect in granular films with tunneling magnetoresistance. *Phys. Solid State*. 2000;42(3):498–502. <https://doi.org/10.1134/1.1131238> [Original Russian Text: Bykov I.V., Gan'shina E.A., Granovskii A.B., Gushchin B.C. Magnetorefractive effect in granular films with tunneling magnetoresistance. *Fizika tverdogo tela*. 2000;42(3):487–491 (in Russ.).]
6. Gushchin B.C., Gan'shina E.A., Kozlov A.A., Bykov I.V. Magnetorefractive effect in nanocomposites. *Moscow University Physics Bulletin*. 2005;60(1):57–75. [Original Russian Text: Gushchin B.C., Gan'shina E.A., Kozlov A.A., Bykov I.V. Magnetorefractive effect in nanocomposites. *Vestnik Moskovskogo universiteta. Seriya 3. Fizika. Astronomiya*. 2005;1:45–58 (in Russ.).]

7. Yurasov A., Yashin M., Ganshina E., Granovsky A., Garshin V., Semenova D., Mirzokulov K. Simulation of magneto-optical properties of nanocomposites (CoFeZr)_x(Al₂O₃)_(1-x). *J. Phys.: Conf. Ser.* 2019;1389(1):012113. <http://doi.org/10.1088/1742-6596/1389/1/012113>
8. Apresyan L.A., Vlasova T.V., Krasovskii V.I., et al. Effective Medium Approximations for the Description of Multicomponent Composites. *Tech. Phys.* 2020;65(7):1130–1138. <https://doi.org/10.1134/S106378422007004X>
[Original Russian Text: Apresyan L.A., Vlasova T.V., Krasovskii V.I., Kryshob V.I., Rasmagin S.I. Effective Medium Approximations for the Description of Multicomponent Composites. *Zhurnal tekhnicheskoi fiziki.* 2020;90(7):1175–1183 (in Russ.). <https://doi.org/10.21883/JTF.2020.07.49453.446-18>]
9. Fadeev E.A., Blinov M.I., Garshin V.V. et al. Magnetic Properties of (Co₄₀Fe₄₀B₂₀)_x(SiO₂)_(100-x) Nanocomposites near the Percolation Threshold. *Bull. Russ. Acad. Sci. Phys.* 2019;83(7):835–837. <https://doi.org/10.3103/S1062873819070153>
[Original Russian Text: Fadeev E.A., Blinov M.I., Garshin V.V., Tarasova O.S., Gan'shina E.A., Prudnikova M.V., Prudnikov V.N., Lyakhderanta E., Ryl'kov V.V., Granovskii A.B. Properties of (Co₄₀Fe₄₀B₂₀)_x(SiO₂)_(100-x) Nanocomposites near the Percolation Threshold. *Izvestiya Rossiiskoi akademii nauk. Seriya fizicheskaya.* 2019;83(7):917–920 (in Russ.). <https://doi.org/10.1134/S0367676519070159>]
10. Gan'shina E.A., Pripechenkov I.M., Perova N.N., et al. Magneto-Optical Spectroscopy of Nanocomposites (CoFeB)_x(LiNbO₃)_(100-x) with Concentrations up to the Percolation Threshold: From Superparamagnetism and Superferromagnetism to Ferromagnetism. *Phys. Metals Metallogr.* 2023;124(2):126–132. <https://doi.org/10.1134/s0031918x22601949>
[Original Russian Text: Gan'shina E.A., Pripechenkov I.M., Perova N.N., Kanazakova E.S., Nikolaev S.N., Sitnikov A.S., Granovskii A.B., Ryl'kov V.V. Magneto-optical spectroscopy of nanocomposites (CoFeB)_x(LiNbO₃)_(100-x) with concentrations up to the percolation threshold: from superparamagnetism and superferromagnetism to ferromagnetism. *Fizika metallov i metallovedenie.* 2023;124(2):134–140 (in Russ.).]
11. Granovsky A., Sukhorukov Yu., Gan'shina E., Telegin A. Magnetorefractive effect in magnetoresistive materials. In: *Magnetophotonics: From Theory to Applications.* Berlin Heidelberg: Springer. 2013. P. 107–133.
12. Shkurdoda Yu.O., Dekhtyaruk L.V., Basov A.G., Chornous A.M., Shabelnyk Yu.M., Kharchenko A.P., Shabelnyk T.M. The giant magnetoresistance effect in Co/Cu/Co three-layer films. *Journal of Magnetism and Magnetic Materials.* 2019;477:88–91. <https://www.doi.org/10.1016/j.jmmm.2019.01.040>
13. Dekhtyaruk L.V., Kharchenko A.P., Klymenko Yu.O., Shkurdoda Yu.O., Shabelnyk Yu.M., Bezdidko O.V., Chornous A.M. Negative and Positive Effect of Giant Magnetoresistance in The Magnetically Ordered Sandwich. *2020 IEEE 10th International Conference Nanomaterials: Applications & Properties (NAP).* Sumy, Ukraine. 2020. P. 01NMM13-1-01NMM13-3. <https://www.doi.org/10.1109/NAP51477.2020.9309694>
14. Kelley C.S., Naughton J., Benson E., Bradley R.C., Lazarov V.K., Thompson S.M., Matthew J.A. Investigating the magnetic field-dependent conductivity in magnetite thin films by modelling the magnetorefractive effect. *J. Phys.: Condens. Matter.* 2014;26(3):036002. <https://doi.org/10.1088/0953-8984/26/3/036002>
15. Yurasov A., Yashin M., Ganshina E., et al. Simulation of magneto-optical properties of nanocomposites (CoFeZr)_x(Al₂O₃)_{1-x}. *J. Phys.: Conf. Ser.* 2019;1389:012113. <http://doi.org/10.1088/1742-6596/1389/1/012113>

Об авторах

Мухутдинова Муза Александровна, студент, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: mukhutdinova-03@mail.ru. <https://orcid.org/0009-0001-5013-5208>

Юрасов Алексей Николаевич, д.ф.-м.н., профессор, профессор кафедры наноэлектроники, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: alexey_yurasov@mail.ru. ResearcherID M-3113-2016, Scopus Authors ID 6602974416, SPIN-код РИНЦ 4259-8885, <https://orcid.org/0000-0002-9104-3529>

About the authors

Muza A. Mukhutdinova, Student, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: mukhutdinova-03@mail.ru. <https://orcid.org/0009-0001-5013-5208>

Alexey N. Yurasov, Dr. Sci. (Phys.-Math.), Professor, Department of Nanoelectronics, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: alexey_yurasov@mail.ru. ResearcherID M-3113-2016, Scopus Authors ID 6602974416, RSCI SPIN-code 4259-8885, <https://orcid.org/0000-0002-9104-3529>

Математическое моделирование
Mathematical modeling

УДК 621.396.969.1

<https://doi.org/10.32362/2500-316X-2024-12-3-65-77>

EDN TUXBXL



НАУЧНАЯ СТАТЬЯ

Двумерный алгоритм с последовательной оценкой углов прихода сигналов в системе когерентных распределенных автомобильных радаров с несколькими приемными и передающими антеннами

И.В. Артюхин[®]

Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, 603950 Россия

[®] Автор для переписки, e-mail: artjukhin@rf.unn.ru

Резюме

Цели. Одной из актуальных задач в радиолокации является проблема повышения пространственного разрешения целей при ограниченной апертуре антенной решетки радара и короткой выборке входных отсчетов. Разрабатываемые алгоритмы должны обеспечивать высокое угловое разрешение и иметь малую вычислительную сложность. В настоящее время автомобили для выполнения требований систем безопасности и помощи водителю оснащаются не одним, а несколькими радарными с общей схемой обработки сигналов для улучшения характеристик при обнаружении, позиционировании и распознавании целей по сравнению с одиночным радаром. Цель работы – разработка двумерного алгоритма оценки угловых координат с низкой вычислительной сложностью в системе распределенных когерентных автомобильных радаров для случая короткой выборки входных отсчетов.

Методы. Использованы алгоритм формирования виртуальной антенной решетки, двумерный метод Кейпона. Предложена модификация метода Кейпона на основе последовательной оценки углов прихода сигналов применительно к системе распределенных радаров. Для сравнения эффективности рассматриваемых алгоритмов используется метод Монте-Карло.

Результаты. Представлен алгоритм с последовательной оценкой азимута и угла места для системы распределенных когерентных автомобильных радаров. Приведены результаты сравнительного анализа предложенного алгоритма и классического двумерного метода Кейпона на основе численного моделирования при помощи метода Монте-Карло. Показано, что предложенная схема приводит к улучшению целевых метрик (вероятности правильного определения числа целей, среднеквадратической и систематической ошибок измерения азимута и угла места) по сравнению с одиночным радаром. Последовательный алгоритм обеспечивает выигрыш в использовании вычислительных ресурсов по сравнению с полным двумерным методом Кейпона.

Выводы. Предложенный двумерный метод оценки углов прихода сигналов в азимутальной и угломестной плоскостях может быть применен для распределенной системы бистатических когерентных радаров с несколькими приемными и передающими антеннами (MIMO-радаров). Метод основан на последовательной оценке углов прихода (на первом шаге – в азимутальной плоскости, на втором – в вертикальной). Характеристики системы когерентных радаров с ограниченной конфигурацией антенной решетки, сравнимы с характеристиками высокопроизводительного 4D-радаром со значительно большим числом элементов антенной решетки.

Ключевые слова: распределенные автомобильные радары, когерентная обработка, алгоритмы сверхразрешения, двумерный алгоритм Кейпона

• Поступила: 18.07.2023 • Доработана: 18.03.2024 • Принята к опубликованию: 05.04.2024

Для цитирования: Артюхин И.В. Двумерный алгоритм с последовательной оценкой углов прихода сигналов в системе когерентных распределенных автомобильных радаров с несколькими приемными и передающими антеннами. *Russ. Technol. J.* 2024;12(3):65–77. <https://doi.org/10.32362/2500-316X-2024-12-3-65-77>

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

RESEARCH ARTICLE

High-resolution 2D-DoA sequential algorithm of azimuth and elevation estimation in automotive distributed system of coherent MIMO radars

Igor V. Artyukhin @

Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, 603950 Russia

@ Corresponding author, e-mail: artjukhin@rf.unn.ru

Abstract

Objectives. One of the main tasks of radiolocation involves the problem of increasing spatial resolution of the targets in the case of limited aperture of the radar antenna array and short length of time samples (snapshots). Algorithms must be developed to provide high angular resolution and low computational complexity. In order to conform with the existing Advanced Driver Assistance Systems requirements, modern cars are equipped with more than one radar having a common signal processing scheme to improve performance during target detection, positioning, and recognition as compared to a single radar. The present study aims to develop a two-dimensional Direction-of-Arrival algorithm with low computation complexity as part of distributed coherent automotive radar system for cases involving short time samples (snapshots).

Methods. A virtual antenna array formation algorithm is formulated according to the two-dimensional Capon method. A proposed modification of two-dimensional Capon algorithm is based on sequentially estimating the directions of arrival for the distributed radar system. The Monte Carlo method is used to compare the effectiveness of the considered algorithms.

Results. The 2D-DoA sequential algorithm of azimuth and elevation estimation is proposed. The comparative analysis results for the developed algorithm and classical 2D Capon method based on numerical simulation using Monte Carlo method are presented. The proposed scheme of DoA estimation for coherent signal processing of distributed radars is shown to lead to an improvement of the main considered metrics representing the probability of correctly estimating the number of targets, mean square error, and square error compared to a single radar system. The proposed low-computational algorithm shows the gain in complexity compared to full 2D Capon algorithm.

Conclusions. The proposed two-stage algorithm for estimating the directions of arrival of signals in azimuth and elevation planes can be applied to the distributed system of coherent radars with several receiving and transmitting antennas representing multiple input multiple output (MIMO) radars. The algorithm is based on sequentially estimating the directions of arrival, implying estimation in the azimuthal plane at the first stage and estimation in the vertical plane at the second stage. The performance of a coherent radar system with limited antenna array configuration of separate radar is close in characteristics to a high-performance 4D-radar with a large antenna array system.

Keywords: automotive distributed radars, coherent signal processing, high-resolution algorithms, 2D Capon algorithm

• Submitted: 18.07.2023 • Revised: 18.03.2024 • Accepted: 05.04.2024

For citation: Artyukhin I.V. High-resolution 2D-DOA sequential algorithm of azimuth and elevation estimation in automotive distributed system of coherent MIMO radars. *Russ. Technol. J.* 2024;12(3):65–77. <https://doi.org/10.32362/2500-316X-2024-12-3-65-77>

Financial disclosure: The author has no a financial or property interest in any material or method mentioned.

The author declares no conflicts of interest.

ВВЕДЕНИЕ

Современные разработки в области автомобильной электроники привели к появлению нового поколения радаров – «4D-радары»¹. Такие устройства обеспечивают одновременное измерение трех пространственных координат (дальность, азимут и угол места) и скорости (доплеровское смещение). При этом за счет высокого пространственного разрешения обеспечивается распознавание наложенных друг на друга изображений, таких как пешеход, стоящий у забора или мотоцикл, едущий рядом с грузовиком.

Один из способов улучшения углового разрешения состоит в увеличении количества передающих (Tx) и принимающих (Rx) антенн. Такой подход ограничен сложностями аппаратной реализации и высокой стоимостью конечного продукта. На сегодняшний день количество антенн варьируется от нескольких единиц до нескольких десятков на радаре. Например, радар Continental ARS540 (Xilinx, Германия) оснащен 12 передающими и 16 приемными антеннами². Аналогичную конфигурацию имеет автомобильный радар S80 производства Uhnder, США [1]. В более продвинутом радаре Phoenix компании Arbe (Израиль), реализуется до 48 Tx и 48 Rx антенн³.

Принимая во внимание ограничение на количество физических антенн, разрешение по углам может быть улучшено путем применения дополнительных методов сверхразрешения близкорасположенных источников сигналов (целей). Из классических алгоритмов сверхразрешения можно выделить multiple signal classification (MUSIC), estimation of signal parameters via rotational invariant techniques (ESPRIT), метод Кейпона, метод минимального многочлена и их вариации [2–6]. Однако эти методы, обладая высокой эффективностью, требуют больших вычислительных ресурсов, особенно при использовании

антенных решеток (АР) с достаточно большим числом антенн.

Дополнительно следует отметить, что на практике при обработке сигналов в автомобильных радаров оценка углов прихода (direction of arrival, DoA) обычно осуществляется только по одной последовательности коротких импульсов (так называемый «фрейм»), которая дает одну выборку (один временной отсчет) в каждой приемной антенне [3, 7]. Таким образом, в рассматриваемой задаче число выборок меньше числа элементов АР (так называемая короткая выборка), что приводит к сложности использования отмеченных выше сверхразрешающих методов, связанных с обращением плохо обусловленной корреляционной матрицы такого короткого входного процесса.

Второй способ увеличения углового разрешения основан на реализации когерентного радара с несколькими приемными и передающими каналами – так называемый MIMO (multiple-input-multiple-output) радар. MIMO-радары с N_{Tx} передающими антеннами и N_{Rx} принимающими антеннами формируют виртуальную АР, эквивалентную АР традиционного радара с одной передающей антенной и $N_{Tx} \times N_{Rx}$ принимающими антеннами (MISO – multiple-input-single-output радар). Таким образом, MIMO-радары обеспечивают эффективный путь, позволяющий получить высокое угловое разрешение при меньшем числе реальных антенн [7–9]. Для работы такого радара необходимо, чтобы все принимаемые сигналы были когерентными, а форма сигналов – ортогональной. При построении MIMO-радаров используют различные подходы к ортогонализации, такие как временное или кодовое разделение сигналов [8, 9].

Используя концепцию MIMO, радар Phoenix компании Arbe может создать до 2304 виртуальных антенн, обеспечивая угловое разрешение (ширина луча по уровню 3 дБ), равное 1.25° и 1.5° по азимуту и углу места соответственно. При этом сектор обзора в азимутальной плоскости составляет $\pm 60^\circ$, а по углу места составляет $\pm 15^\circ$. Следуя аналогичному подходу, радар Continental ARS540 может создать 192 виртуальных антенн с разрешением по азимуту в 1.2° , а по углу места – в 2.3° .

Для выполнения существующих требований системы безопасности и помощи водителю (advanced safety systems and advanced driver assistance systems, ADAS) современные автомобили оснащаются

¹ <https://autotech.news/the-future-of-automotive-radar-4d-imaging-radar/>. Дата обращения 07.02.2024. / Accessed February 07, 2024.

² <https://www.xilinx.com/publications/presentations/continental-ARS540-powered-by-xilinx.pdf>. Дата обращения 21.08.2023. / Accessed August 21, 2023.

³ https://arberobotics.com/wp-content/uploads/2021/04/4D_Imaging_Radar_Product_Overview.pdf. Дата обращения 21.08.2023. / Accessed August 21, 2023.

несколькими датчиками (сенсорами), которые могут работать как единое целое и имеют общую (централизованную) обработку сигналов. В качестве преимуществ таких систем можно отметить существенное улучшение их характеристик при решении задач обнаружения, позиционирования, распознавания целей по сравнению с одиночным радаром (традиционным или MIMO).

Обычно выделяют два подхода к интеграции изменений с разных радаров, основанных на когерентной и некогерентной обработке сигналов. Некогерентная схема предполагает объединение сигналов отдельных радаров, работающих в моностатичном режиме [10, 11]. Когерентная схема основана на построении совместной виртуальной AP распределенных радаров, работающих в режиме бистатических измерений, при этом требуется полная или частичная синхронизация сигналов между радаром [12].

Настоящая статья посвящена двумерному методу оценки DoA при когерентной обработке сигналов в системе распределенных автомобильных радаров для случая короткой выборки входного процесса. Данный метод должен обеспечить высокое угловое разрешение за счет когерентной обработки сигналов в распределенных радарах с небольшим числом реальных антенн вместо использования одного высокопроизводительного 4D-радар с большой AP. При этом разрешение по азимуту в 1° и углу места в 2° будем считать необходимыми для достижения характеристиками.

1. ПОСТРОЕНИЕ ВИРТУАЛЬНОЙ AP ДЛЯ РАСПРЕДЕЛЕННОЙ СИСТЕМЫ КОГЕРЕНТНЫХ РАДАРОВ

Рассмотрим распределенную систему, состоящую из двух когерентных MIMO-радаров. Предположим, что данная система имеет N_{Tx} и N_{Rx} реальных антенн. Общее число антенн виртуальной AP будет равно $N_{virt} = N_{Tx} \cdot N_{Rx}$.

Расположение виртуальных антенн определяется как свертка координат реальных приемных и передающих антенн [13, 14]. Координаты m -й виртуальной антенны будут рассчитываться по формуле:

$$\rho^{(m)} = \mathbf{r}_{Tx}^{(p)} + \mathbf{r}_{Rx}^{(q)} - \mathbf{r}_0, \quad (1)$$

где $\mathbf{r}_{Tx}^{(p)}$ – радиус-вектор координат p -го элемента передающей антенны, $\mathbf{r}_{Rx}^{(q)}$ – радиус-вектор координат q -го элемента приемной антенны, \mathbf{r}_0 – радиус-вектор определяющий выбор системы координат, $m = q \cdot p$. Необходимо отметить, что при определенных условиях архитектуры AP радаров несколько виртуальных антенн могут иметь одинаковые координаты в пространстве.

Уравнение (1) может быть переписано в следующем виде:

$$\rho^{(m)} = \mathbf{r}_{Tx}^{(p)} + \mathbf{r}_{Rx}^{(q)} - \mathbf{r}_{Tx}^{(1)} - \mathbf{r}_{Rx}^{(1)}, \quad (2)$$

где $\mathbf{r}_{Tx}^{(1)}$ и $\mathbf{r}_{Rx}^{(1)}$ – радиус-векторы координат первого передающего и первого приемного элементов AP соответственно.

Свойствами виртуальной AP можно управлять, изменяя координаты реальных Rx и Tx антенн. Как правило, выбор положения антенн направлен на получение виртуальной решетки с большей апертурой, что ведет к улучшению углового разрешения близких объектов. При этом концепция виртуальной AP позволяет применять традиционную обработку AP (включая алгоритмы сверхвысокого разрешения) для оценки угловых координат. Основная идея виртуальной AP заключается в совпадении ее выходного сигнала с сигналом, создаваемым некоторой реальной AP. В свободном пространстве такое совпадение однозначно, однако в случае многолучевого распространения необходимо учитывать критерии совпадения для определения оптимальной конфигурации реальной AP. В [14] показано, что в случае наличия отраженного сигнала от Земли либо все передающие, либо все приемные антенны должны иметь одинаковую высоту над поверхностью Земли. Простейшей формой такой конфигурации реальных передающих и приемных антенн является L-образная AP.

На рис. 1 показан пример построения виртуальной AP для распределенной системы, состоящей из двух когерентных радаров с L-образной AP. Передающие антенны обозначены крестом, а приемные антенны – кругом. AP одиночных радаров расположены симметрично относительно вертикальной оси (центра радаров).

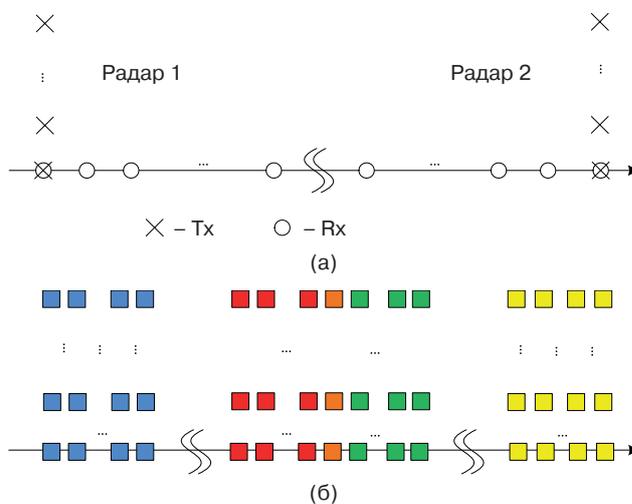


Рис. 1. Общая геометрия распределенной системы радаров: (а) AP, состоящая из двух радаров L-типа; (б) виртуальная AP

Результирующая виртуальная АР состоит из двух виртуальных моностатических решеток (выделенных синим и желтым цветами на рис. 1б, и одной виртуальной бистатической решетки, расположенной посередине. Бистатическая АР состоит из удвоенного числа элементов виртуальной решетки одиночного радара [13, 14].

Красным и зеленым цветом показаны виртуальные антенны, сформированные с помощью бистатического отклика левого и правого радаров соответственно. Средние элементы, выделенные оранжевым цветом в виртуальной АР, содержат перекрывающиеся бистатические измерения от правого и левого радаров. Такое положение антенн в среднем столбце необходимо для учета и компенсации разности фаз, которая возникает между бистатическими сигналами в виртуальной АР при неидеальной синхронизации.

2. ОПИСАНИЕ ИСПОЛЬЗУЕМЫХ АЛГОРИТМОВ

2.1. Алгоритм компенсации фаз

Предлагается алгоритм компенсации фаз сигналов, состоящий из двух этапов. Первый этап используется в случае ненулевого установочного угла α одиночных радаров (рис. 2). Если угол установки одиночных радаров равен нулю, то первый этап пропускается и применяется только второй.

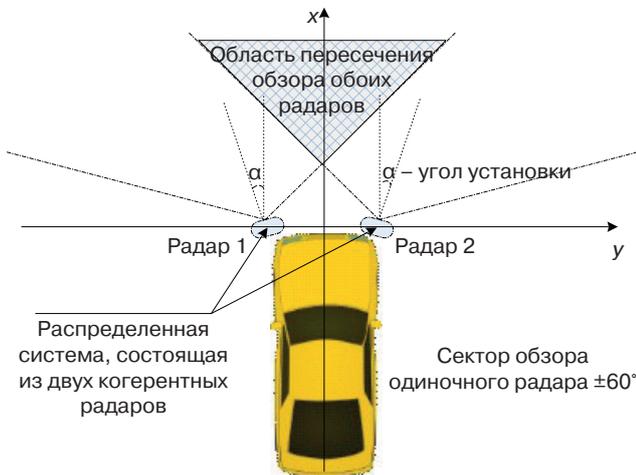


Рис. 2. Общая геометрия распределенной системы радаров. Вид сверху

Первый этап: компенсация дополнительного набега фаз, связанного с ненулевым установочным углом отдельного радара.

Обозначим через матрицы \mathbf{Y}_i ($i = 1, 2$) сигналы виртуальных бистатических решеток (откликов) от левого и правого радаров. И введем p -ю компоненту сканирующего вектора $\mathbf{s}(\alpha)$ АР в азимутальной плоскости для установленного угла α радара как:

$$s_p(\alpha) = \exp(2\pi j d_\lambda (p-1) \sin(\alpha)), \quad p = \overline{1, N}, \quad (3)$$

где N – число антенных элементов в горизонтальной плоскости одиночного радара, j – мнимая единица.

Для компенсации дополнительного набега фаз, связанного с ненулевым установочным углом отдельного радара, сделаем следующее преобразование:

$$\mathbf{Z}_1 = \mathbf{G}_1 \cdot \mathbf{Y}_1, \quad \mathbf{Z}_2 = \mathbf{G}_2 \cdot \mathbf{Y}_2, \quad (4)$$

где

$$\mathbf{G}_1 = \text{diag}\{\mathbf{s}\}, \quad \mathbf{G}_2 = \text{diag}\{\mathbf{s}^*\} \quad (5)$$

и $\text{diag}\{\mathbf{s}\}$ – диагональная матрица, элементы вектора \mathbf{s} лежат на главной диагонали, $*$ – знак комплексного сопряжения.

Введем матрицу \mathbf{Z} размерности $N_{\text{Tx}}/2 \times 2N$, включающую в себя сигналы \mathbf{Z}_1 и \mathbf{Z}_2 :

$$\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2]. \quad (6)$$

Второй этап: скачок в фазах между бистатическими измерениями радаров компенсируется за счет учета разности фаз радаров в среднем элементе построенной виртуальной АР.

Выделим из полного набора \mathbf{Z} отдельную строку, состоящую из $2N$ элементов: $\{z(1), z(2), \dots, z(2N)\}$. Виртуальная АР (выделенная строка) состоит из $2N - 1$ элементов с перекрывающимися измерениями в среднем элементе (рис. 3).

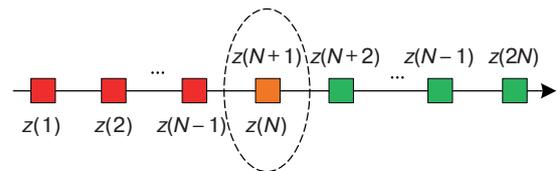


Рис. 3. Иллюстрация построения отсчетов для отдельной строки виртуальной решетки

Найдем разность фаз в среднем элементе $\Delta\Phi = \text{phase}(z(N)) - \text{phase}(z(N+1))$, и скомпенсируем полученную разность для антенных элементов с индексами $N+2, \dots, 2N$:

$$\{z(N+2), z(N+3), \dots, z(2N)\} \cdot \exp(-j \cdot \Delta\Phi). \quad (7)$$

Преобразованные сигналы, используемые для оценки углов прихода, состоят из следующих $2N - 1$ элементов:

$$\{z(1), z(2), \dots, z(N), z(N+2) \cdot \exp(-j\Delta\Phi), \dots, z(N+3) \cdot \exp(-j\Delta\Phi), \dots, z(2N) \cdot \exp(-j\Delta\Phi)\}. \quad (8)$$

Элемент с индексом $N + 1$ из полного набора не используется в конечном наборе (рис. 4).

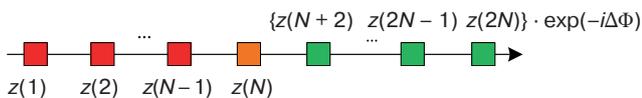


Рис. 4. Иллюстрация построения отсчетов для отдельной строки виртуальной решетки

2.2. Двумерный алгоритм оценки углов прихода

Рассмотрим два метода оценки азимута и угла места целей: классический двумерный (2D) алгоритм Кейпона и предлагаемый (новый) двухшаговый метод последовательной оценки азимута и угла места.

Непараметрический двумерный метод Кейпона выбран в качестве базового. Для него не требуются априорные знания о количестве цели (или их предварительная оценка), что упрощает оценку углов прихода, но при этом метод имеет большую вычислительную сложность, связанную с нахождением обратной матрицы и двумерным поиском экстремумов [2, 13]. Данный алгоритм основан на поиске максимумов двумерной разрешающей функции $\eta_C(\varphi, \theta)$:

$$\eta_C(\varphi, \theta) = \frac{1}{\mathbf{S}(\varphi, \theta)^H \mathbf{M}_{2D}^{-1} \mathbf{S}(\varphi, \theta)}, \quad (9)$$

где \mathbf{M}_{2D} – оценка корреляционной матрицы входного процесса двумерной AP, $\mathbf{S}(\varphi, \theta)$ – соответствующий сканирующий вектор.

Предлагаемый в данной статье метод оценки DoA заключается в последовательной оценке азимута (первый шаг) и угла места (второй шаг). Рассмотрим в деталях предложенную схему оценки углов прихода.

Первый шаг предполагает реализацию одномерного алгоритма Кейпона. Пусть имеется $m \times n$ виртуальная AP (обозначена на рис. 5 слева синими квадратами). Для оценки DoA используется один временной отсчет, который может быть представлен в математическом виде с помощью матрицы \mathbf{A} (см. рис. 5 справа).

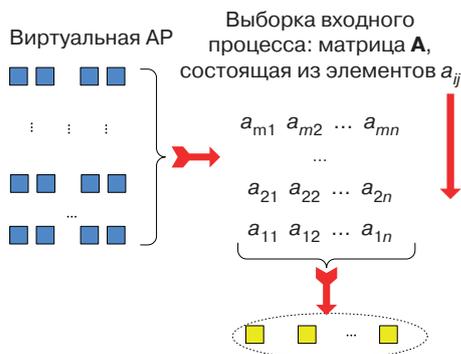


Рис. 5. Пример построения отсчетов для оценки в азимутальной плоскости

Выделим горизонтальную строку длины n (одномерная AP – желтые квадраты, отмеченные эллипсом на рис. 5) из полной виртуальной AP. При этом вся матрица \mathbf{A} будет использоваться для оценки корреляционной матрицы одномерной AP. Отдельные строки матрицы \mathbf{A} можно интерпретировать как «временные отсчеты» для одномерной AP. Сформируем матрицу \mathbf{A} в виде:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}, \quad (10)$$

тогда одномерную разрешающую функцию метода Кейпона можно представить в виде выражения:

$$\eta(\varphi) = \frac{1}{\mathbf{S}(\varphi)^H \mathbf{M}_{1D}^{-1} \mathbf{S}(\varphi)}, \quad (11)$$

где матрица $\mathbf{M}_{1D} = \frac{1}{L} \cdot \mathbf{A} \cdot \mathbf{A}^H$, $L = m$ – число строк в виртуальной AP, $(\bullet)^H$ – эрмитово сопряжение.

В результате решения задачи максимизации (11), мы получим k оценок азимута: $\varphi_1, \dots, \varphi_k$.

На втором шаге реализуется двумерный алгоритм Кейпона для оценки углов места. Однако, в отличие от полного 2D-метода Кейпона, поиск максимумов разрешающей функции $\eta_C(\varphi, \theta)$ осуществляется только при фиксированных значениях азимутов целей $\varphi_1, \dots, \varphi_k$, найденных на первом этапе:

$$\eta_C(\varphi, \theta) \Big|_{\varphi = \varphi_l} = \frac{1}{\mathbf{S}(\varphi, \theta)^H \mathbf{M}_{2D}^{-1} \mathbf{S}(\varphi, \theta)} \Big|_{\varphi = \varphi_l}, \quad (12)$$

φ_l – фиксированы, $l = \overline{1, k}$.

Таким образом, задача поиска максимумов двумерной функции $\eta_C(\varphi, \theta)$ (9) свелась к k одномерным задачам поиска экстремумов (12). Решением (12) являются оценки угла места $\theta_1, \dots, \theta_k$.

Структурная схема предложенного алгоритма представлена на рис. 6.

2.3. Процедура пространственного сглаживания

В практических задачах для оценки углов прихода сигналов используется только одна временная выборка. В этом случае корреляционные матрицы \mathbf{M}_{1D} и \mathbf{M}_{2D} являются вырожденными. В такой ситуации применяется процедура пространственного сглаживания для входных данных [3]. Для этого в основной AP выделяются подрешетки, оптимальный размер

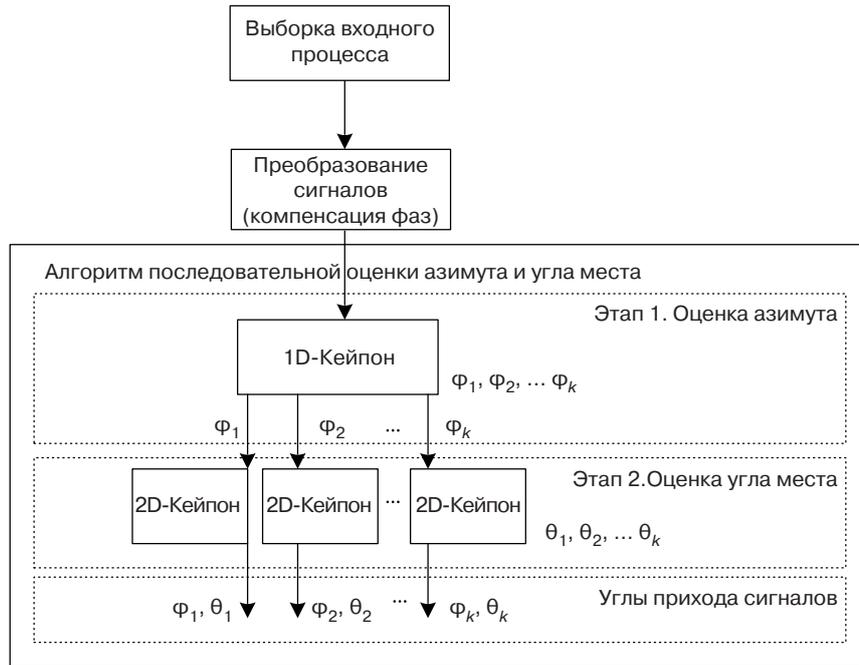


Рис. 6. Последовательный алгоритм оценки DoA

которых составляет $Q \approx 0.7N_{\text{dim}}$ [3, 5] для каждой размерности N_{dim} .

Для пояснения процедуры пространственного сглаживания рассмотрим формирование подрешеток на примере системы радаров L-типа с конфигурацией 6T8R. Тогда бистатистическая виртуальная решетка имеет размер 6×15 , размер подрешетки для пространственного сглаживания будет равен 4×10 при использовании двумерного алгоритма Кейпона (отмечен красным цветом на рис. 7а). В случае одномерного алгоритма Кейпона, размер подрешетки будет равен 1×10 (отмечен зеленым цветом на рис. 7б).

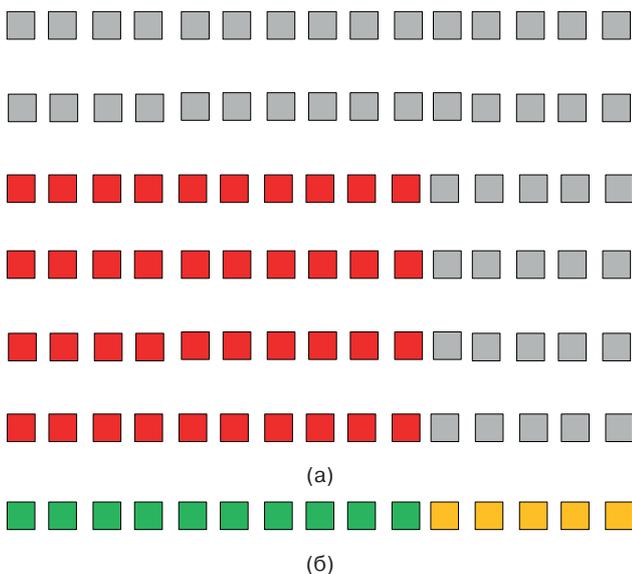


Рис. 7. Конфигурация подрешеток:
(а) подрешетка для 2D-алгоритма Кейпона;
(б) подрешетка для 1D-алгоритма Кейпона

После выбора подрешетки формируются несколько выборок. Каждая выборка является частью принятого сигнала, соответствующего смещенным копиям выбранной подрешетки (так называемое прямое сглаживание). Число выборок можно удвоить, применив процедуру обратного сглаживания [3, 5]. В результате пространственного сглаживания получаем $L_{2D} = 36$ и $L_{1D} = 12 \times 6 = 72$ пространственных выборок для 2D- и 1D-алгоритмов Кейпона соответственно.

2.4. Оценка вычислительных ресурсов

Для оценки требуемых вычислительных ресурсов при аппаратной реализации алгоритмов будем учитывать только операцию комплексного умножения как наиболее затратную процедуру [15, 16]. Вычислительная сложность полного 2D-метода Кейпона представлена в табл. 1, 2. Оценка вычислительных ресурсов для предложенного последовательного алгоритма оценки DoA собрана в табл. 3, 4.

Численные оценки вычислительных ресурсов в табл. 2, 4 представлены для следующего случая:

- AP одиночного радара 6T8R, L-типа;
- размер виртуальной решетки (бистатистические измерения): 6×15 ;
- один временной отсчет;
- применяется процедура пространственного сглаживания.

Параметры «число антенных элементов» и «число временных отсчетов» в табл. 2, 4 – это параметры после процедуры пространственного сглаживания.

Таблица 1. Вычислительные затраты 2D-метода Кейпона

Процедура	Число операций
Оценка корреляционной матрицы	$L_{2D}(N_{2D})^2 + (N_{2D})^2$
Обращение матрицы	$4(N_{2D})^2$
Сканирование по углам разрешающей функции	$(N_{2D})^2 + N_{2D}$
Поиск максимумов	$4J_1 N_{steps}$
Общая сложность	$(N_{2D})^2(L_{2D} + 6) + N_{2D} + 4J_1 N_{steps}$

Таблица 2. Вычислительные затраты 2D-метода Кейпона

Параметр	Значение
Число антенных элементов	$N_{2D} = 40$
Число временных отсчетов	$L_{2D} = 36$
Число целей	$J_1 = 2$
Число шагов (точек) для поиска максимумов	$N_{steps} = N_h N_v$ $N_h = N_v = 100$
Общая сложность	147240

Таблица 3. Вычислительные затраты последовательного алгоритма оценки DoA

Процедура	Число операций
1-й этап (1D-Кейпон)	
Оценка корреляционной матрицы	$L_{1D}(N_{1D})^2 + (N_{1D})^2$
Обращение матрицы	$4(N_{1D})^2$
Сканирование по углам разрешающей функции	$(N_{1D})^2 + N_{1D}$
Поиск максимумов	$4J_1 N_h$
2-й этап (2D-Кейпон)	
Оценка корреляционной матрицы	$4(L_{2D}(N_{2D})^2 + (N_{2D})^2)$
Обращение матрицы	$4(N_{2D})^2$
Сканирование по углам разрешающей функции	$(N_{2D})^2 + N_{2D}$
Поиск максимумов	$2 \cdot 4J_2 N_v$
Общая сложность	$(N_{2D})^2(L_{2D} + 6) + N_{2D} + 2 \cdot 4J_2 N_v + (N_{1D})^2(L_{2D} + 6) + N_{1D} + 4J_1 N_h$

Таблица 4. Вычислительные затраты последовательного алгоритма оценки DoA

Параметр	Значение
Число антенных элементов	$N_{1D} = 10$
Число временных отсчетов	$L_{1D} = 72$
Число целей	$J_2 = 1$
Общая сложность	76650

Вычислительная сложность полного 2D-метода Кейпона составляет 147240 операций (комплексных умножений) и 76650 операций для последовательного алгоритма оценки DoA при заданных параметрах. Таким образом, выигрыш предложенного алгоритма в вычислительных ресурсах составляет 1.9 раз по сравнению с классическим двумерным алгоритмом Кейпона.

3. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Предположим, что распределенная система состоит из двух когерентных радаров миллиметрового диапазона длин волн (77 ГГц), расстояние между сенсорами 1.48 м, угол установки радаров 0° (радары смотрят вперед). Антенная решетка одиночного радара имеет L-образную 6T8R конфигурацию, представленную на рис. 8.

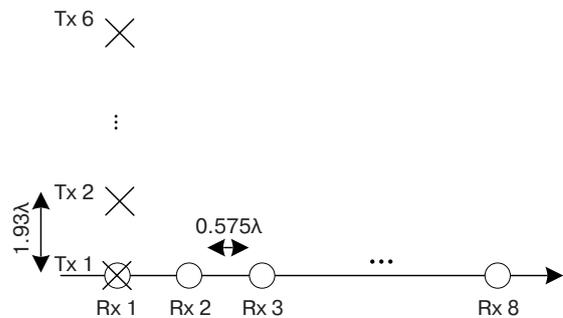


Рис. 8. AP одиночного радара: 6T8R (L-типа)

Период AP равен 0.575λ (горизонтальная плоскость) и 1.93λ (вертикальная плоскость). Параметры одиночного радара обеспечивают сектор обзора $\pm 60^\circ$ в азимутальной плоскости и $\pm 15^\circ$ в вертикальной плоскости. Ширина главного луча по уровню половинной мощности (-3дБ) составляет 11.25° в горизонтальной плоскости и 4.46° – в вертикальной плоскости.

Для распределенной системы радаров с учетом получаемой виртуальной бистатической AP ширина главного луча в вертикальной плоскости не изменится, а в горизонтальной плоскости – уменьшится до $\approx 6^\circ$ по сравнению с одиночным радаром.

Для иллюстрации эффективности распределенной системы когерентных радаров рассмотрим два

сценария пространственного расположения двух близких целей. В первом сценарии, представленном на рис. 9а, две цели расположены симметрично относительно оси x с угловым расстоянием 1° в азимутальной плоскости и имеют одинаковый угол места 0° . Угловое расстояние между целями в 11 раз меньше ширины луча АР одиночного радара. Во втором сценарии (рис. 9б) две цели имеют одинаковый азимут 0° и расположены симметрично относительно оси x в вертикальной плоскости с угловым расстоянием 2° . Общее расположение целей и радаров проиллюстрировано на рис. 9в. Угловое расстояние между целями при этом примерно в 2.2 раза меньше ширины луча АР одиночного радара в вертикальной плоскости.

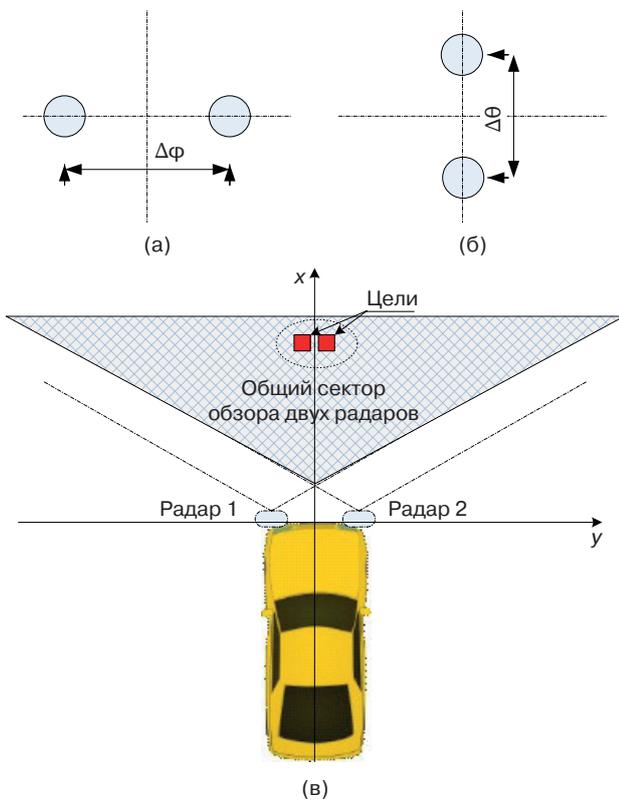


Рис. 9. Сценарии пространственного расположения двух близких целей: (а) сценарий 1; (б) сценарий 2; (в) общая геометрия распределенной системы двух радаров

Полный 2D-метод Кейпона для одиночного радара и распределенной системы используется в качестве базового алгоритма для сравнения с предложенным алгоритмом последовательной оценки азимута и угла места для системы распределенных радаров. Один временной отсчет взят для оценки углов прихода сигналов.

Для сравнения эффективности алгоритмов используется метод Монте-Карло. Для каждого значения отношения сигнал/шум набирался статистический ансамбль, состоящий из 2000 экспериментов. Рассматриваются три основные метрики как функции параметра «отношение сигнал/шум» (ОСШ):

1. Вероятность правильной оценки числа близких целей p :

$$p = \frac{m}{n}, \quad (13)$$

где m – число экспериментов с правильным обнаружением числа целей, n – общее число экспериментов.

2. Среднеквадратичная ошибка (СКО, mean square error, MSE):

$$MSE = \sqrt{\frac{1}{J} \sum_{k=1}^J \sigma_k^2}, \quad (14)$$

$$\sigma_k^2 = \frac{1}{m-1} \sum_{i=1}^m (\hat{\varphi}_{ki} - \bar{\varphi}_k)^2, \quad \bar{\varphi}_k = \frac{1}{m} \sum_{i=1}^m \hat{\varphi}_{ki},$$

где $\hat{\varphi}_{km}$ – оценка DoA для k -й цели в m -м эксперименте.

3. Систематическая ошибка (СО, systematic error, SE):

$$SE = \sqrt{\frac{1}{J} \sum_{k=1}^J (\bar{\varphi}_k - \varphi_k)^2}. \quad (15)$$

Обобщенные результаты численного моделирования для предложенного алгоритма последовательной оценки азимута и угла места представлены в табл. 5. Видно, что разрешение по азимуту и углу места (1° и 2° соответственно) достигается при ОСШ в 36 дБ (сценарий 1) и 20 дБ (сценарий 2) для вероятности правильной оценки числа целей равной 0.5.

Таблица 5. Обобщенные результаты численного моделирования для предложенного алгоритма последовательной оценки азимута и угла места

Сценарий	Разрешение		Точность		ОСШ
	Азимут, $\Delta\varphi$	Угол места, $\Delta\theta$	Азимут	Угол места	
1	1°	–	0.12° (СКО) 0.11° (СО)	0.6° (СКО) 0.04° (СО)	36 дБ
2	–	2°	0.08° (СКО) 0.02° (СО)	0.45° (СКО) 0.04° (СО)	20 дБ

Сравнение метрик как функций параметра ОСШ для рассматриваемых алгоритмов при различных сценариях представлено ниже на рис. 10 и 11. Полный набор исследуемых метрик состоит из пяти графиков. В левой части представлены СКО и СО метрики в зависимости от ОСШ для оценки азимута, в правой – для угла места. Метрики СКО и СО

представлены с учетом дополнительного ограничения на вероятность корректного обнаружения числа близких целей ($p > 0.5$).

Сценарий 1 (рис. 9а): две цели имеют азимуты $\pm 0.5^\circ$ и одинаковый угол места 0° .

Черная пунктирная кривая со звездочками соответствует предложенному последовательному алгоритму для распределенной системы радаров; черная пунктирная кривая с кружками соответствует распределенной системе радаров для полного

2D-алгоритма Кейпона; синяя кривая с треугольниками/красная кривая с крестиками — оценка DoA для одиночного радара (левый/правый), используется полный 2D-алгоритм Кейпона. Как видно из графиков, одиночные радары не смогли достичь целевого уровня метрики правильной оценки количества целей в $p = 0.5$. Таким образом, на графиках отсутствуют результаты для метрик СКО/СО.

Из результатов моделирования для текущего сценария видно, что предложенный алгоритм

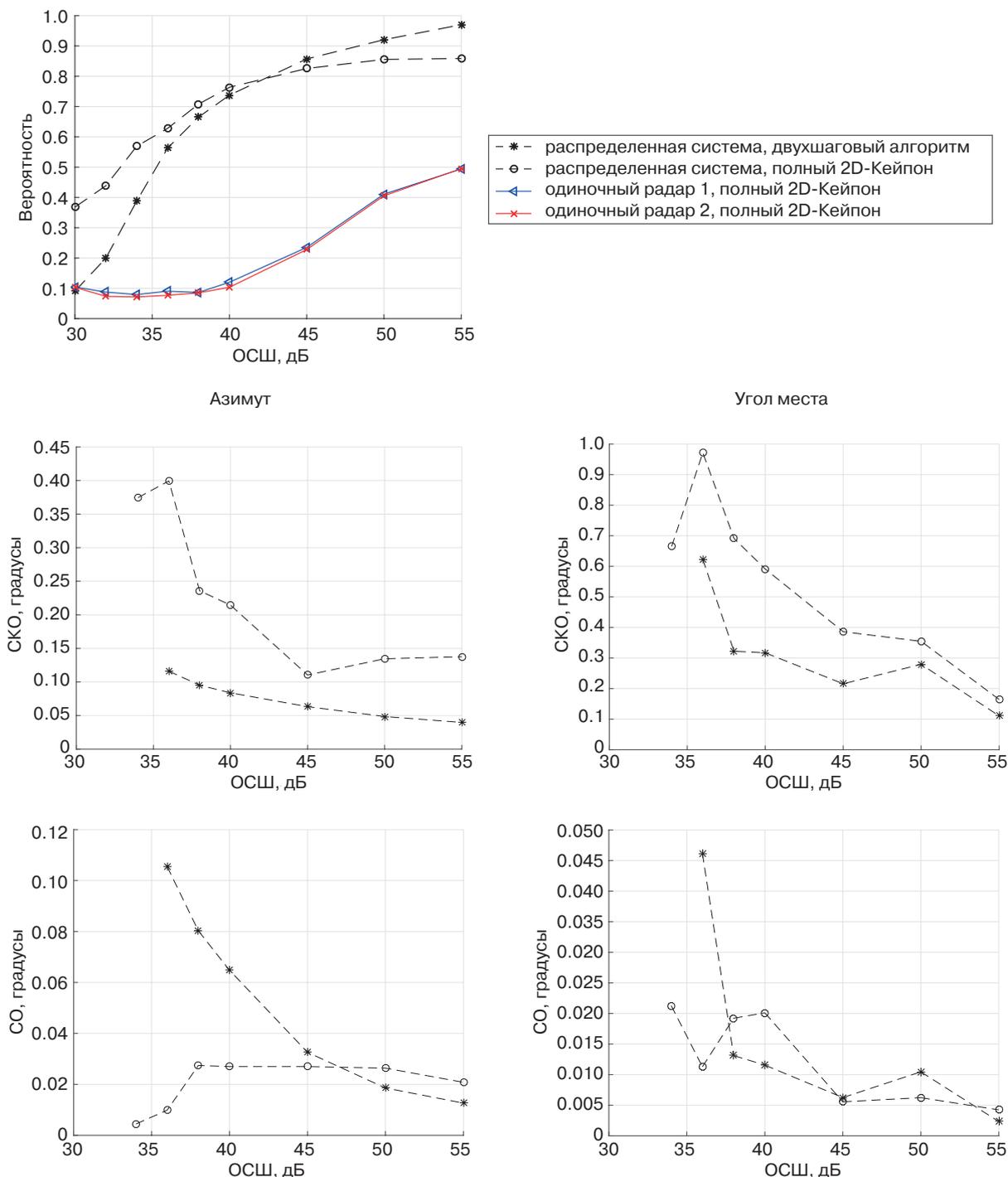


Рис. 10. Метрики как функция ОСШ. Сценарий 1

последовательной оценки углов прихода сигналов выигрывает у полного 2D-алгоритма Кейпона в метриках правильной оценки числа целей (для больших значений ОСШ) и СКО за счет большего числа пространственных выборок при оценке азимута на первом этапе (72 против 36, см. п. 2.4). При этом метрики СО для двух алгоритмов имеют близкие значения.

Сценарий 2 (рис. 9б): две цели имеют одинаковый азимут 0° и углы места $\pm 1^\circ$ (угловое расстояние 2° в вертикальной плоскости).

Из представленных результатов видно, что последовательный алгоритм оценки углов прихода приводит к улучшению метрик СКО и СО. Одиночный радар также способен различать две близких цели в сценарии 2 за счет одинаковых апертур АР отдельного радара и распределенной системы в вертикальной плоскости. При этом одиночный радар с полным 2D-алгоритмом Кейпона проигрывает в метриках СКО и СО распределенной системе радаров с двухшаговым алгоритмом.

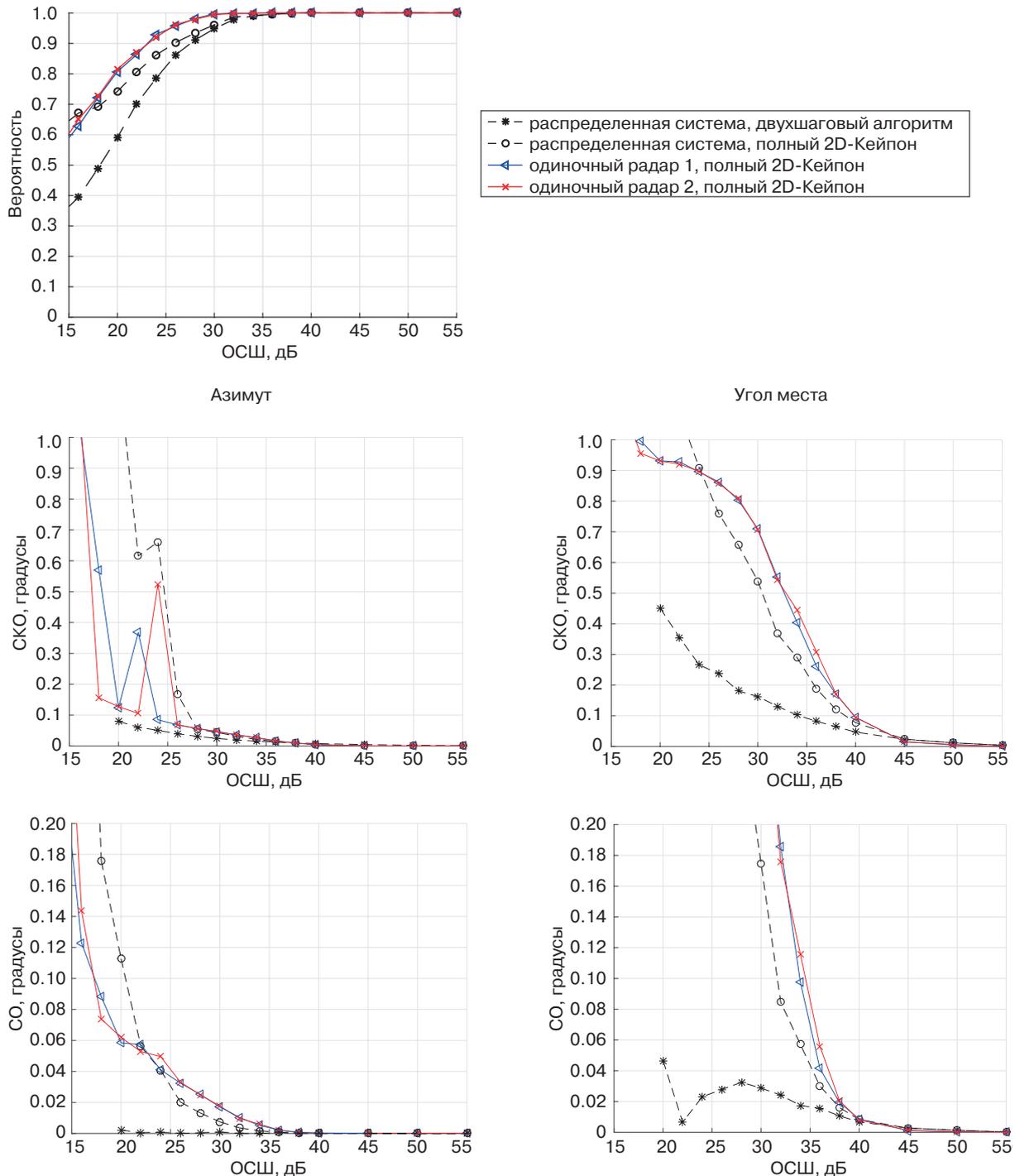


Рис. 11. Метрики как функция ОСШ. Сценарий 2

Выигрыш предложенного алгоритма последовательной оценки углов прихода сигналов по сравнению с полным 2D-алгоритмом Кейпона можно объяснить более точной оценкой азимута на первом этапе метода за счет большего числа пространственных выборок (72 против 36), в то время как классический двумерный алгоритм Кейпона вынужден оценивать одновременно азимут и угол места при меньшей длине пространственной выборки.

ЗАКЛЮЧЕНИЕ

В работе предложен двумерный метод оценки углов прихода сигналов в азимутальной и угломестной плоскостях для распределенной системы бистатических когерентных MIMO-радаров. Метод основан на последовательной оценке углов прихода

(на первом шаге – в азимутальной плоскости, на втором – в вертикальной) и обеспечивает выигрыш (при выбранных параметрах системы) в 1.9 раз в использовании вычислительных ресурсов по сравнению с полным классическим двумерным методом Кейпона.

Проведенное сравнительное численное моделирование на основе метода Монте-Карло показало, что предложенная схема когерентной обработки сигналов распределенных радаров для оценки DoA приводит к улучшению целевых метрик (вероятности корректного определения числа целей, СКО и СО) по сравнению с одиночным радаром. Характеристики системы, состоящей из когерентных радаров с ограниченной конфигурацией AP, сравнимы с характеристиками высокопроизводительного 4D-радар со значительно большим числом элементов AP.

СПИСОК ЛИТЕРАТУРЫ

- Giannini V., Goldenberg M., Eshrahi A., et al. 9.2 A 192-Virtual-Receiver 77/79GHz GMSK Code-Domain MIMO Radar System-on-Chip. In: *2019 IEEE International Solid-State Circuits Conference (ISSCC)*, San Francisco, CA, USA. 2019. P. 164–166. <https://doi.org/10.1109/ISSCC.2019.8662386>
- Godara L.C. *Smart Antennas*. CRC Press; 2004. 472 p.
- Tuncer T.E., Friedlander B. (Eds.). *Classical and Modern Direction-of-Arrival Estimation*. Academic Press, Inc.; 2009. 456 p. <https://doi.org/10.1016/C2009-0-19135-3>
- Ермолаев В.Т., Флакман А.Г., Елохин А.В., Купцов В.В. Метод минимального многочлена для оценки параметров сигналов, принимаемых антенной решеткой. *Акустический журнал*. 2018;64(1):78–85. <https://doi.org/10.7868/S0320791918010057>
- Ермолаев В.Т., Флакман А.Г., Елохин А.В., Шмонин О.А. Угловое сверхразрешение сигналов в антенной решетке с помощью корневого метода минимального многочлена корреляционной матрицы. *Известия вузов. Радиофизика*. 2018;61(3):261–272.
- Родионов А.А., Турчин В.И. Обработка сигналов в антенных решётках на основе модели помехи, включающей корреляционную матрицу неполного ранга. *Известия вузов. Радиофизика*. 2017;60(1):60–71.
- Patole S.M., Torlak M., Wang D., Ali M. Automotive radars: A review of signal processing techniques. *IEEE Signal Processing Magazine*. 2017;34(2):22–35. <https://doi.org/10.1109/MSP.2016.2628914>
- Ермолаев В.Т., Семенов В.Ю., Флакман А.Г., Артюхин И.В., Шмонин О.А. Метод формирования виртуальных приемных каналов в автомобильном MIMO-радаре. *Радиотехника*. 2021;85(7):115–126.
- Li J., Stoica P. *MIMO Radar Signal Processing*. Wiley-IEEE Press; 2009. 448 p. ISBN 978-0-4701-7898-0
- Folster F., Rohling H., Lubbert U. An automotive radar network based on 77 GHz FMCW sensors. In: *IEEE International Radar Conference*. 2005. P. 871–876. <https://doi.org/10.1109/RADAR.2005.1435950>
- Артюхин И.В., Аверин И.М., Флакман А.Г., Рубцов А.Е. Алгоритм оценки углов прихода сигналов в системе распределенных некогерентных автомобильных радаров. *Журнал радиоэлектроники*. 2023;4:1–20. <http://jre.cplire.ru/jre/apr23/2/text.pdf>
- Gottinger M., Hoffmann M., Christmann M., Schutz M., Kirsch F., Gulden P., Vossiek M. Coherent Automotive Radar Networks: The Next Generation of Radar-Based Imaging and Mapping. *IEEE Journal of Microwaves*. 2021;1(1):149–163. <https://doi.org/10.1109/JMW.2020.3034475>
- Richards M.A. *Fundamentals of Radar Signal Processing*. 2nd edition. New York: McGraw-Hill; 2014. 656 p.
- Ермолаев В.Т., Флакман А.Г., Шмонин О.А. Применение концепции виртуальной антенной решетки в MIMO-радаре при наличии отражений от земной поверхности. *Известия вузов. Радиофизика*. 2020;63(3):240–249.
- Björnson E., Hoydis J., Sanguinetti L. Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency. *Found. Trends® Sign. Process.* 2017;11(3–4):154–655. <http://doi.org/10.1561/20000000093>
- Gentilho E., Scalassara P.R., Abrão T. Direction-of-Arrival Estimation Methods: A Performance-Complexity Tradeoff Perspective. *J. Sign. Process. Syst.* 2020;92(2):239–256. <https://doi.org/10.1007/s11265-019-01467-4>

REFERENCES

1. Giannini V., Goldenberg M., Eshrahi A., et al. 9.2 A 192-Virtual-Receiver 77/79GHz GMSK Code-Domain MIMO Radar System-on-Chip. In: *2019 IEEE International Solid-State Circuits Conference (ISSCC)*, San Francisco, CA, USA. 2019. P. 164–166. <https://doi.org/10.1109/ISSCC.2019.8662386>
2. Godara L.C. *Smart Antennas*. CRC Press; 2004. 472 p.
3. Tuncer T.E., Friedlander B. (Eds.). *Classical and Modern Direction-of-Arrival Estimation*. Academic Press, Inc.; 2009. 456 p. <https://doi.org/10.1016/C2009-0-19135-3>
4. Ermolaev V.T., Flaksman A.G., Elokhin A.V., et al. Minimal Polynomial Method for Estimating Parameters of Signals Received by an Antenna Array. *Acoust. Phys.* 2018;64(1):83–90. <https://doi.org/10.1134/S1063771018010050>
[Original Russian Text: Ermolaev V.T., Flaksman A.G., Elokhin A.V., Kuptsov V.V. Minimal Polynomial Method for Estimating Parameters of Signals Received by an Antenna Array. *Akusticheskii Zhurnal*. 2018;64(1):78–85 (in Russ.). <https://doi.org/10.7868/S0320791918010057>]
5. Ermolaev V.T., Flaksman A.G., Elokhin A.V., et al. Angular Superresolution of the Antenna-Array Signals Using the Root Method of Minimum Polynomial of the Correlation Matrix. *Radiophys. Quantum Electron.* 2018;61(3):232–241. <https://doi.org/10.1007/s11141-018-9884-5>
[Original Russian Text: Ermolaev V.T., Flaksman A.G., Elokhin A.V., Shmonin O.A. Angular Superresolution of the Antenna-Array Signals Using the Root Method of Minimum Polynomial of the Correlation Matrix. *Izvestiya Vysshikh Uchebnykh Zavedenii. Radiofizika*. 2018;61(3):261–272 (in Russ.).]
6. Rodionov A.A., Turchin V.I. Processing of Antenna-Array Signals on the Basis of the Interference Model Including a Rank-Deficient Correlation Matrix. *Radiophys. Quantum Electron.* 2017;60(1):54–64. <https://doi.org/10.1007/s11141-017-9776-0>
[Original Russian Text: Rodionov A.A., Turchin V.I. Processing of Antenna-Array Signals on the Basis of the Interference Model Including a Rank-Deficient Correlation Matrix. *Izvestiya Vysshikh Uchebnykh Zavedenii. Radiofizika*. 2017;60(1):60–71 (in Russ.).]
7. Patole S.M., Torlak M., Wang D., Ali M. Automotive radars: A review of signal processing techniques. *IEEE Signal Processing Magazine*. 2017;34(2):22–35. <https://doi.org/10.1109/MSP.2016.2628914>
8. Ermolaev V.T., Semenov V.Yu., Flaksman A.G., Artyukhin I.V., Shmonin O.A. The method of forming virtual receiving channels in the automobile MIMO-radar. *Radiotekhnika = J. Radioengineering*. 2021;85(7):115–126 (in Russ.).
9. Li J., Stoica P. *MIMO Radar Signal Processing*. Wiley-IEEE Press; 2009. 448 p. ISBN 978-0-4701-7898-0
10. Folster F., Rohling H., Lubbert U. An automotive radar network based on 77 GHz FMCW sensors. In: *IEEE International Radar Conference*. 2005. P. 871–876. <https://doi.org/10.1109/RADAR.2005.1435950>
11. Artyukhin I.V., Averin I.M., Flaksman A.G., Rubtsov A.E. Direction-of-Arrival Estimation Algorithm in Automotive Distributed Non-Coherent Multi-Radar Systems. *Zhurnal radioelektroniki = J. Radio Electronics*. 2023;4:1–20 (in Russ.). <https://doi.org/10.30898/1684-1719.2023.4.2>
12. Gottinger M., Hoffmann M., Christmann M., Schutz M., Kirsch F., Gulden P., Vossiek M. Coherent Automotive Radar Networks: The Next Generation of Radar-Based Imaging and Mapping. *IEEE Journal of Microwaves*. 2021;1(1):149–163. <https://doi.org/10.1109/JMW.2020.3034475>
13. Richards M.A. *Fundamentals of Radar Signal Processing*. 2nd edition. New York: McGraw-Hill; 2014. 656 p.
14. Ermolaev V.T., Flaksman A.G., Shmonin O.A. Using the Concept of a Virtual Antenna Array in a MIMO Radar in the Presence of Reflections from the Ground Surface. *Radiophys. Quantum Electron.* 2020;63(3):218–226. <https://doi.org/10.1007/s11141-021-10047-1>
[Original Russian Text: Ermolaev V.T., Flaksman A.G., Shmonin O.A. Using the Concept of a Virtual Antenna Array in a MIMO Radar in the Presence of Reflections from the Ground Surface. *Izvestiya Vysshikh Uchebnykh Zavedenii. Radiofizika*. 2020;63(3):240–249 (in Russ.).]
15. Björnson E., Hoydis J., Sanguinetti L. Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency. *Foundations and Trends® in Signal Processing*. 2017;11(3–4):154–655. <http://doi.org/10.1561/20000000093>
16. Gentilho E., Scalassara P.R., Abrão T. Direction-of-Arrival Estimation Methods: A Performance-Complexity Tradeoff Perspective. *J. Sign. Process. Syst.* 2020;92(2):239–256. <https://doi.org/10.1007/s11265-019-01467-4>

Об авторе

Артюхин Игорь Владимирович, электроник 1 категории, кафедра статистической радиофизики и мобильных систем связи, Радиофизический факультет, ФГАОУ ВО «Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского» (603022, Россия, Нижний Новгород, пр-т Гагарина, д. 23). E-mail: artjukhin@rf.unn.ru. Scopus Author ID 57216223873, <https://orcid.org/0009-0008-5139-6443>

About the author

Igor V. Artyukhin, Engineer, Department of Statistical Radiophysics and Mobile Communication Systems, Faculty of Radiophysics, National Research Lobachevsky State University of Nizhny Novgorod (23, Gagarina pr., Nizhny Novgorod, 603022 Russia). E-mail: artjukhin@rf.unn.ru. Scopus Author ID 57216223873, <https://orcid.org/0009-0008-5139-6443>

Математическое моделирование
Mathematical modeling

УДК 519.63

<https://doi.org/10.32362/2500-316X-2024-12-3-78-92>

EDN WBOETG



НАУЧНАЯ СТАТЬЯ

Анализ и прогнозирование динамики настроений пользователей интернет-ресурсов на основе уравнения Фоккера – Планка

Ю.П. Перова[®],
С.А. Лесько,
А.А. Иванов

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[®] Автор для переписки, e-mail: jul-np@yandex.ru

Резюме

Цели. Цель работы – вывод наблюдаемого на практике степенного закона распределения характеристик социально-динамических процессов из стационарного уравнения Фоккера – Планка и проверка возможности применения нестационарного уравнения Фоккера – Планка для описания динамики процессов в социальных системах.

Методы. При проведении исследований были использованы методы моделирования стохастических процессов, методы и модели теории графов, инструменты и технологии объектно-ориентированного программирования для разработки систем сбора данных из массмедиа-источников, методы имитационного моделирования.

Результаты. Наблюдаемое текущее состояние графа сети комментариев может быть описано с помощью вектора, элементами которого являются среднее значение коэффициента посредничества, среднее значение коэффициента кластеризации, доля пользователей в конкретном состоянии. Критическое состояние сети может быть задано базовым вектором. Зависимость от времени расстояния между базовым вектором и текущим вектором состояния образует временной ряд, значения которого можно рассматривать как «блуждающую точку», динамика перемещений которой описывается нестационарным уравнением Фоккера – Планка. Текущее состояние графа комментариев можно определить с помощью методов текстовой аналитики.

Выводы. Наблюдаемый на практике степенной закон зависимости стационарной плотности вероятности распределения новостей по числу комментариев может быть получен из решения стационарного уравнения Фоккера – Планка, а нестационарное уравнение может быть использовано для описания процессов в сложных сетевых структурах. Для описания состояний сети комментариев пользователей новостных массмедиа можно использовать векторное представление. Достижение или реализация желаемых, или нежелательных состояний всей социальной сети могут быть заданы на основе базовых векторов. Решение нестационарного уравнения Фоккера – Планка позволяет получить уравнение для плотности вероятности переходов между состояниями системы в единицу времени, которые хорошо согласуются с наблюдаемыми данными. Анализ полученной модели с использованием характеристик реального временного ряда для изменения графа комментариев читателей официальной страницы в социальной сети «ВКонтакте» информационного агентства «РИА Новости» и структурных параметров графа показывает ее адекватность.

Ключевые слова: социальные сети, моделирование социальных процессов, сетевой граф, характеристики сетевого графа, уравнение Фоккера – Планка, мониторинг, управление, нелинейная динамика, степенной закон распределения

• Поступила: 12.01.2023 • Доработана: 17.11.2023 • Принята к опубликованию: 08.04.2024

Для цитирования: Перова Ю.П., Лесько С.А., Иванов А.А. Анализ и прогнозирование динамики настроений пользователей интернет-ресурсов на основе уравнения Фоккера – Планка. *Russ. Technol. J.* 2024;12(3):78–92. <https://doi.org/10.32362/2500-316X-2024-12-3-78-92>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Analyzing and forecasting the dynamics of Internet resource user sentiments based on the Fokker–Planck equation

Julia P. Perova [®],
Sergey A. Lesko,
Andrey A. Ivanov

MIREA – Russian Technological University, Moscow, 119454 Russia
[®] Corresponding author, e-mail: jul-np@yandex.ru

Abstract

Objectives. The study aims to theoretically derive the power law observed in practice for the distribution of characteristics of sociodynamic processes from the stationary Fokker–Planck equation and apply the non-stationary Fokker–Planck equation to describe the dynamics of processes in social systems.

Methods. During the research, stochastic modeling methods were used along with methods and models derived from graph theory, as well as tools and technologies of object-oriented programming for the development of systems for collecting data from mass media sources, and simulation modeling approaches.

Results. The current state of the comment network graph can be described using a vector whose elements are the average value of the mediation coefficient, the average value of the clustering coefficient, and the proportion of users in a corresponding state. The critical state of the network can be specified by the base vector. The time dependence of the distance between the base vector and the current state vector forms a time series whose values can be considered as the “wandering point” whose movement dynamics is described by the non-stationary Fokker–Planck equation. The current state of the comment graph can be determined using text analysis methods.

Conclusions. The power law observed in practice for the dependence of the stationary probability density of news distribution by the number of comments can be obtained from solving the stationary Fokker–Planck equation, while the non-stationary equation can be used to describe processes in complex network structures. The vector representation can be used to describe the comment network states of news media users. Achieving or implementing desired or not desired states of the whole social network can be specified on the basis of base vectors. By solving the non-stationary Fokker–Planck equation, an equation is obtained for the probability density of transitions between system states per unit time, which agree well with the observed data. Analysis of the resulting model using the characteristics of the real time series to change the graph of comments of users of the RIA Novosti portal and the structural parameters of the graph demonstrates its adequacy.

Keywords: social networks, modeling of social processes, network graph, network graph characteristics, Fokker–Planck equation, monitoring, management, nonlinear dynamics, power law of distribution

• Submitted: 12.01.2023 • Revised: 17.11.2023 • Accepted: 08.04.2024

For citation: Perova J.P., Lesko S.A., Ivanov A.A. Analyzing and forecasting the dynamics of Internet resource user sentiments based on the Fokker–Planck equation. *Russ. Technol. J.* 2024;12(3):78–92. <https://doi.org/10.32362/2500-316X-2024-12-3-78-92>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Описание поведения пользователей социальных сетей и информационных ресурсов является одним из важнейших направлений математической социологии. С практической точки зрения создание моделей, описывающих динамику проявления пользовательских мнений и предпочтений, позволяет разрабатывать системы автоматизированного мониторинга общественного настроения и тенденций его изменения. Преимуществом таких систем по сравнению с традиционными методами изучения общественного мнения является их технологичность в реализации.

Следует отметить, что динамику изменения мнений и настроений пользователей сети Интернет можно в значительной степени отнести к стохастическим процессам. Присутствие человеческого фактора (множество людей с различными мнениями, предпочтениями и характером поведения) с одной стороны создает случайность изменений (в силу большого разнообразия поведенческих моделей пользователей), а с другой стороны вносит в динамику изменений элементы целенаправленности.

В [1] рассмотрена модель, описывающая пространственное и временное распространение информации в социальных сетях на основе стохастического дифференциального уравнения в частных производных. В работе была создана и исследована неавтономная диффузионная логистическая модель с граничными условиями Дирихле, которая показала, что на диффузию данных в социальных сетях сильно влияют коэффициент диффузии и внутренняя скорость роста (распространение информации или слухи можно рассматривать как своего рода вирусы, не обладающие физической формой).

В связи с этим наиболее перспективными для создания моделей динамики изменения общественного настроения на наш взгляд являются модели, которые можно создать на основе стохастических дифференциальных уравнений, например, уравнения Фоккера – Планка, которое учитывает как упорядоченные («снос»), так и случайные изменения («диффузия»). Уравнение Фоккера – Планка широко применяется для анализа и моделирования поведения временных рядов при описании процессов в сложных системах [2–5].

Следует отметить, что помимо уравнения Фоккера – Планка для моделирования на основе дифференциальных уравнений используются и другие подходы, например, уравнения Лиувилля [5, 6], уравнения диффузии [4, 7] и ряда других.

Помимо описания динамических процессов из уравнения Фоккера – Планка можно получить и стационарные решения, которые могут описывать состояние какой-либо системы в стационарном состоянии, когда, например, ее эволюция уже закончилась и изменения не происходят.

Для моделирования социальных процессов применяются не только модели на основе дифференциальных уравнений в частных производных, но, например, модели на основе теоретико-игровых подходов и принятие управленческих решений на их основе [8].

Динамика процессов в сетевых структурах неразрывно связана с их топологией, поэтому необходимо учитывать и их структурные характеристики. Например, в [9] представлена методология анализа тематик в социальной сети, которая включает сбор, обработку и классификацию информации, а также измерение времени между публикациями. Эти данные затем используются для создания шкалы времени.

Таким образом, создается граф, который позволяет отследить изменения популярности конкретных тем, обсуждаемых в социальных сетях. Этот граф также позволяет увидеть связанные события на основе настроений, присутствующих в обществе, а также определить периоды активных дискуссий по конкретным темам. А в [10] на основе графовой модели Кронекера проводилось изучение обнаружения сообществ на графах, обнаружения перекрывающихся сообществ, обнаружения сообществ в неполных сетях с отсутствующими связями и в полных сетях.

Кроме того, приобретает популярность для исследования структуры связей социальных сетей использование природоподобных алгоритмов. В частности, в [11] для оптимизации процесса решения проблем прогнозирования связей и обнаружения сообществ в анализе социальных сетей используются роевые алгоритмы. С увеличением размеров сетей поиск сходства между их узлами становится весьма трудоемким процессом.

Исследование процессов, происходящих в сложных системах с участием человеческого фактора, показывает, что очень часто для наблюдаемых характеристик параметров этих процессов выполняется степенной закон распределения $\rho(x) \sim x^{-\gamma}$ (где γ – характеристическая степень) [12–17], но в то же время вопрос теоретического обоснования возможности ее применения требует дальнейшего изучения. На наш взгляд, это обоснование очень важно. Выявление характера процессов, из которых возникает степенной закон, необходимо для более глубокого изучения поведения и анализа сложных социальных систем.

В связи с этим перспективным является исследование возможности применения уравнения Фоккера – Планка для разработки моделей динамики социальных процессов.

1. СБОР И ОБРАБОТКА ДАННЫХ

Для исследований были выбраны несколько новостных порталов и одно из сетевых сообществ социальной сети «ВКонтакте», посвященное обсуждению новостей информационного ресурса «РИА Новости»¹. Этот ресурс был выбран, исходя из его узнаваемости и популярности в российском обществе: он занимает 1 место среди медиаресурсов (за март 2022 г.) по версии «Brand-Analytics»², входит в топ-3 самых цитируемых информационных агентств в СМИ и социальных медиа (за март 2022 г.) и занимает 1 место³ по этим показателям в российском сегменте интернета.

Сначала с помощью специального программного приложения (парсера) был загружен интересующий диапазон новостей, начиная с 1 января 2019 г. по апрель 2022 г., с официальной страницы «РИА Новости» в социальной сети «ВКонтакте», используя разработанный парсер и программный интерфейс сети⁴. Внутри социальной сети у каждого поста есть свой уникальный адрес, где {owner_id} – уникальный идентификатор сообщества (в случае с «РИА новости» – это «-15755094»), а {post_id} – уникальный идентификатор поста (новости). У каждого поста (новости) присутствует ряд основных параметров, таких как: уникальный идентификатор поста в социальной сети; текст поста; дата и время публикации; количество просмотров

и комментариев пользователей. Комментарии имеют следующие параметры: уникальный идентификатор в сообществе социальной сети; уникальный идентификатор пользователя; текст комментария; дата и время появления; уровень иерархии комментария; связь по уровню комментирования с родительским комментарием (кто из пользователей комментировал кого из других пользователей при обсуждении новости).

Поскольку комментарии могли оставлять чат-боты, спамеры и недобросовестные пользователи, которые пишут комментарии на профессиональной основе, необходимо было ввести правила отчистки данных. К недобросовестным были отнесены те, кто написали за год более 7365 комментариев (в среднем более 20 за сутки) или писали с частотой более одного комментария в 5 мин.

При анализе полученных данных необходимо было определить, какому закону распределения подчиняется наблюдаемая плотность распределения. Для этого были рассмотрены три наиболее часто наблюдаемых закона распределения: Гаусса

$$\rho(x) = e^{-\frac{x^2}{2\sigma^2}} / \sigma\sqrt{2\pi}; \text{ экспоненциальное } \rho(x) = ae^{-\alpha x}$$

и степенное $\rho(x) \sim \beta x^{-\gamma}$. При обработке собранных данных с помощью линейаризации в соответствующих координатах было обнаружено, что наиболее близкая линейаризация наблюдается для степенного закона распределения (рис. 1). Для остальных законов линейаризация была хуже.

Прямая, проведенная на рис. 1, показывает, что линия тренда хорошо описывается выбранной нами линейной аппроксимацией $y = -0.76 - 1.48z$, где $y = \ln\{\rho(x)\}$, $z = \ln\{x\}$, $\ln\{\beta\} = -0.76$, а коэффициент корреляции равен 0.95.

Для подтверждения вывода о линейной аппроксимации можно исследовать поведение остатков и проверить гипотезу о том, что они нормально распределены со средним значением, равным нулю, и имеют однородную дисперсию. Вычисление остатков можно провести на основе реально наблюдаемых значений натурального логарифма доли комментаторов, давших данное число комментариев, и полученного нами уравнения. Рассчитанная величина математического ожидания для распределения остатков равна 0.25, а значение дисперсии равно 0.13. Проверка гипотезы о наклоне (двухвыборочный F-тест для дисперсий) показывает, что дисперсия остатков (рассчитанная относительно линии тренда) существенно меньше, чем дисперсия отклонения точек линейной регрессии от среднего значения величины наблюдаемых данных ($\sum y_i^2/n = \sum \ln\{\rho(x_i)\}^2/n$). Она равна 2.11 ($0.13 \ll 2.11$). Таким образом, из полученных данных можно сделать вывод, что распределение остатков очень близко к нормальному, и полученная регрессия является

¹ <https://vk.com/ria> (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

² <https://br-analytics.ru/mediatrends/media/?period=202203> (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

³ <https://www.mlg.ru/ratings/media/federal/11110/#internet> (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

⁴ <https://dev.vk.com/guide> (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

значимой, что подтверждает вывод о том, что натуральный логарифм доли комментаторов, написавших данные комментарии, линейно зависит от натурального логарифма числа комментариев, что подтверждает выполнение степенного закона.

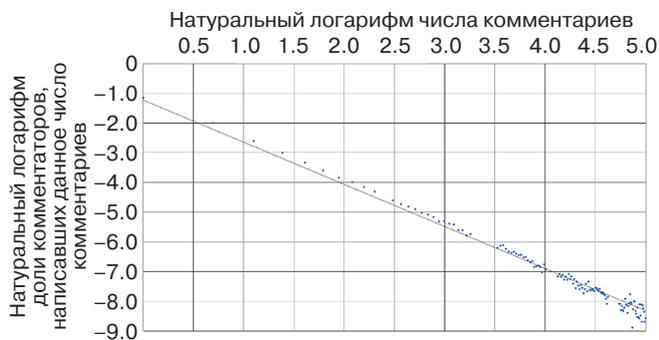


Рис. 1. Линейаризация наблюдаемых данных для степенного распределения доли комментаторов от числа сделанных ими комментариев

При проведении исследований также представляет интерес анализ динамики изменения числа комментариев к новостям, имеющим большой общественный интерес (за время просмотра такие новости набирают сотни комментариев), с течением времени.

Наблюдение этой динамики изменения показывает, что она может иметь как S-образный, так и двухступенчатый характер. В качестве демонстрации такой динамики комментирования пользователями «ВКонтакте» новостей ресурса «РИА Новости» выберем несколько новостей:

1. «Зеленский покинул Украину и переехал в Польшу, – заявил Володин» (см. рис. 2а, на котором динамика комментирования имеет S-образный характер)^{5, 6}, дата и время появления: 2022-03-04 16:13:27 UTC +03:00). Общее число комментариев составило 894. Число комментариев первого уровня (комментарий самой новости) составило 433, второго (комментарии комментариев первого уровня) – 461. Общее количество просмотров – 118764. Среднее время появления комментариев первого уровня – 73 минуты, второго уровня – 74 минуты.
2. «Скоро город будет освобожден, – заявил он» (см. рис. 2б, на котором динамика комментирования имеет двухступенчатый характер)^{7, 8}, дата и время появления: 2022-04-10 17:14:40 UTC +03:00).

⁵ https://vk.com/ria?w=wall-15755094_34243579 (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

⁶ <https://ria.ru/20220304/zelenskiy-1776545154.html> (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

⁷ https://vk.com/ria?w=wall-15755094_35202266 (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

⁸ <https://ria.ru/20220410/ukraina-1782778315.html> (in Russ.). Дата обращения 02.09.2023. / Accessed September 02, 2023.

Общее число комментариев составило 901. Число комментариев первого уровня составило 173, второго – 728. Общее количество просмотров – 173607. Среднее время появления комментариев первого уровня – 75 минуты, а второго уровня – 82 минуты.

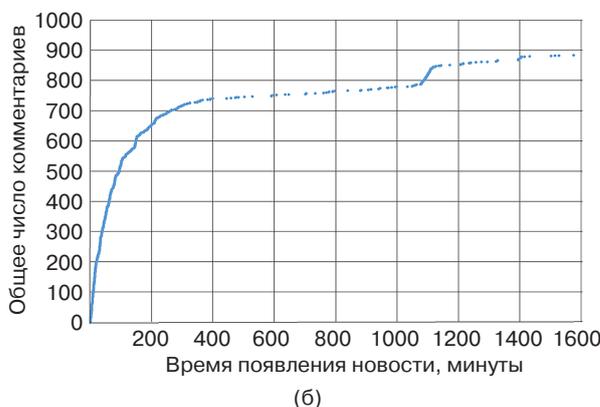
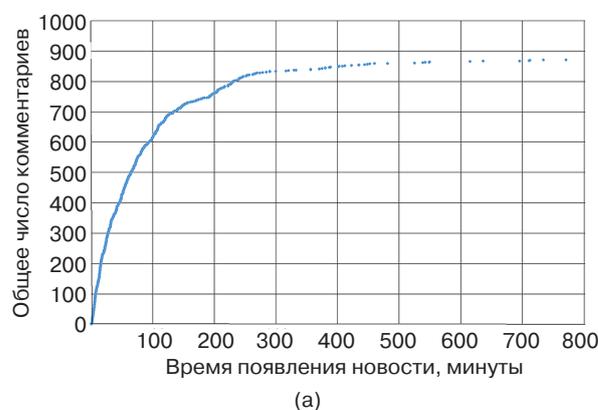


Рис. 2. Наблюдаемая динамика изменения числа комментариев к новостям

На наш взгляд, это может быть связано как с различием в среднем времени появления комментариев второго уровня (интервал времени между появлением комментария первого уровня и комментарием к данным комментариям), так и с соотношением между числом комментариев первого и второго уровней. Если для первой новости интервалы среднего времени появления комментариев первого и второго уровней практически совпадают, то для второй наблюдается небольшое увеличение интервала среднего времени появления комментариев второго уровня (происходит запаздывание по времени). Кроме того, для второй новости их число существенно превосходит число комментариев первого уровня.

Можно сформулировать следующую задачу дальнейшего теоретического исследования: какая природа процессов комментирования новостей и блогов, и какие особенности этих сложных социальных систем приводят к тому, что зависимость плотности вероятности распределения комментариев

от их числа описывается с помощью степенного закона, а динамика имеет во многих случаях сложный двухступенчатый характер?

2. РЕШЕНИЕ СТАЦИОНАРНОГО УРАВНЕНИЯ ФОККЕРА – ПЛАНКА

В общем виде уравнение Фоккера – Планка имеет вид:

$$\frac{\partial \rho(x, t)}{\partial t} = -\frac{\partial}{\partial x}[\mu(x)\rho(x, t)] + \frac{1}{2} \cdot \frac{\partial^2}{\partial x^2}[D(x)\rho(x, t)], \quad (1)$$

где $\rho(x, t)$ – зависящая от времени t плотность вероятности распределения по состояниям x (в нашем случае состояние x – это число комментариев, наблюдаемое в момент времени t); $D(x)$ – зависящий от состояния x коэффициент, определяющий случайное изменение состояния x («диффузия»); $\mu(x)$ – зависящий от состояния x коэффициент, определяющий целенаправленное изменение состояния x («снос»). Применительно к нашей модели $D(x)$ можно трактовать как действия пользователя, вызванные спонтанным импульсом, возникшим при прочтении новости или комментариев к ней других пользователей, когда описываемое в новости или блоге событие не является существенно важным, но пользователь готов потратить время на комментарий или ответить другому комментатору (у пользователя возникло спонтанное желание отреагировать на данную новость). Коэффициент $\mu(x)$ можно интерпретировать как целенаправленные действия, вызванные желанием отреагировать на существенно важные для пользователя новость или блог, а также дать комментарий на комментарий другого пользователя, если он затронул важную с точки зрения данного пользователя тему (пользователь постоянно интересуется данной темой).

Для построения модели необходимо сделать предположения о зависимости $D(x)$ и $\mu(x)$ от состояния x и рассмотреть два условия. Во-первых, учтем размерность членов, входящих в уравнение (1), а во-вторых, можно сделать предположение, что с ростом состояния x (ростом числа возможных комментариев, значимости новости или блога) величины $D(x)$ и $\mu(x)$ также должны увеличиваться. Физические соображения подсказывают, что все члены уравнения (1) должны иметь одинаковую размерность, которую имеет $\rho(x, t)$. И первое, и второе условие будет выполнено, если зависимости $D(x)$ и $\mu(x)$ от состояния x будут иметь вид: $\mu(x) = \mu_0 x$ и $D(x) = D_0 x^2$. Коэффициенты μ_0 и D_0 имеют размерность $1/t$.

Решение стационарного уравнения Фоккера – Планка:

$$-\frac{d}{dx}[\mu(x)\rho(x)] + \frac{1}{2} \cdot \frac{d^2}{dx^2}[D(x)\rho(x)] = 0$$

при следующих допущениях: $\mu(x) = \mu_0 x$ и $D(x) = D_0 x^2$ имеет вид $\rho(x) = [\gamma - 1]x^{-\gamma}$, что соответствует наблюдаемому на практике степенному закону распределения. По полученным из анализа наблюдаемых данных результатам $\gamma = 1.48$, а $\gamma - 1 = 0.48$, натуральный логарифм $(\gamma - 1)$ равен -0.73 , что с достаточно высокой точностью равно $\ln\{\beta\} = -0.76$ (см. полученное уравнение линеаризации: $y = -0.76 - 1.48z$, где $y = \ln\{\rho(x)\}$, $z = \ln\{x\}$, $\ln\{\beta\} = \ln\{\gamma - 1\} - 0.76$, а коэффициент корреляции равен 0.95). В целом это указывает на адекватность разработанной модели. Заметим, что члены стационарного уравнения для $\rho(x)$ также должны иметь одинаковую размерность. В решение стационарного уравнения входит выражение $2 \left[1 - \frac{\mu_0}{D_0} \right] = \gamma$, а оно от t не зависит.

Вывод степенного закона может быть описан следующим образом. Подставляем в стационарное уравнение Фоккера – Планка выражения для $\mu(x)$ и $D(x)$.

Подстановка $D(x)$ и $\mu(x)$ в уравнение (1) дает:

$$-\mu_0 x \frac{d\rho(x)}{dx} - \mu_0 \rho(x) + \frac{1}{2} D_0 x^2 \frac{d^2 \rho(x)}{dx^2} + 2D_0 x \frac{d\rho(x)}{dx} + D_0 x \rho(x) = 0.$$

Обозначим $2 \left[1 - \frac{\mu_0}{D_0} \right] = \gamma$, тогда

$$x^2 \frac{d^2 \rho(x)}{dx^2} + [2 + \gamma] x \frac{d\rho(x)}{dx} + \gamma \rho(x) = 0.$$

Решение полученного уравнения можно искать в виде: $\rho(x) = \sum_k C_k x^q$, где C_k – постоянные коэффициенты при соответствующих корнях характеристического уравнения, которое имеет вид: $q(q - 1) + [2 + \gamma]q + \gamma = 0$. Данное уравнение имеет два корня: $q_1 = -1$ и $q_2 = -\gamma$. Таким образом для $\rho(x)$ получаем: $\rho(x) = C_1 x^{-1} + C_2 x^{-\gamma}$. Постоянные коэффициенты C_1 и C_2 находятся с использованием условия нормировки функции $\rho(x)$:

$$\int_1^{\infty} \rho(x) dx = C_1 \ln x \Big|_1^{\infty} + C_2 \frac{x^{1-\gamma}}{1-\gamma} \Big|_1^{\infty} \equiv 1.$$

Интеграл вычисляется от 1 до ∞ , поскольку могут быть пользователи, которые дали очень большое число комментариев к новостям, но не может быть комментаторов, написавших менее одного комментария. Учитывая, что при $x \rightarrow \infty \ln x \Big|_{\infty} = \infty$, $C_1 = 0$ и соответственно $C_2 = \gamma - 1$, окончательно получаем: $\rho(x) = [\gamma - 1]x^{-\gamma}$.

3. РЕШЕНИЕ НЕСТАЦИОНАРНОГО УРАВНЕНИЯ ФОККЕРА – ПЛАНКА И АНАЛИЗ МОДЕЛИ

Приведем краткий вывод решения нестационарного уравнения Фоккера – Планка. Используя метод преобразований Лапласа для уравнения Фоккера – Планка, можно записать:

$$\begin{aligned} s\overline{G(s, x)} - \rho(x, 0) = \\ = -\frac{d}{dx} \left[\mu(x)\overline{G(s, x)} \right] + \frac{1}{2} \cdot \frac{d^2}{dx^2} \left[D(x)\overline{G(s, x)} \right]. \end{aligned} \quad (2)$$

Подставив в уравнение (2) соответствующие производные и зависимости $\mu(x)$ и $D(x)$, получим:

$$\begin{aligned} x^2 \frac{d^2 \overline{G(s, x)}}{dx^2} + 2 \left[2 - \frac{\mu_0}{D_0} \right] x \frac{d \overline{G(s, x)}}{dx} + \\ + 2 \left[1 - \frac{\mu_0 + s}{D_0} \right] \overline{G(s, x)} = -\frac{\delta(x - x_0)}{D_0}. \end{aligned} \quad (3)$$

Решение уравнения (3) будем искать в виде: $\overline{G(s, x)} = \sum_k C_k x^q$, где C_k – коэффициенты для корней характеристического уравнения, которое имеет вид:

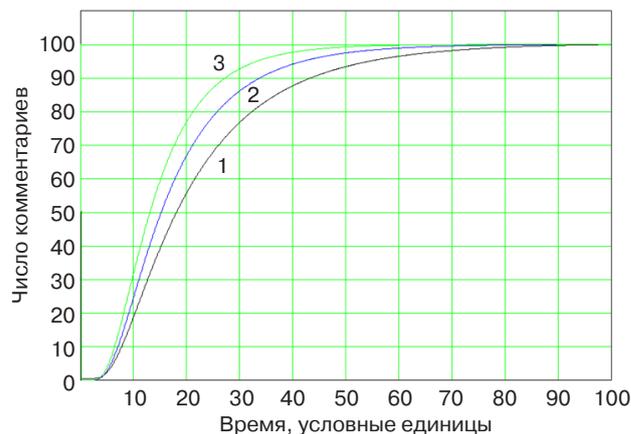
$$q(q-1) + 2 \left[2 - \frac{\mu_0}{D_0} \right] q + 2 \left[1 - \frac{\mu_0 + s}{D_0} \right] = 0.$$

Решение нестационарного уравнения Фоккера – Планка (3) при сделанных $\mu(x)$ и $D(x)$ допущениях имеет вид:

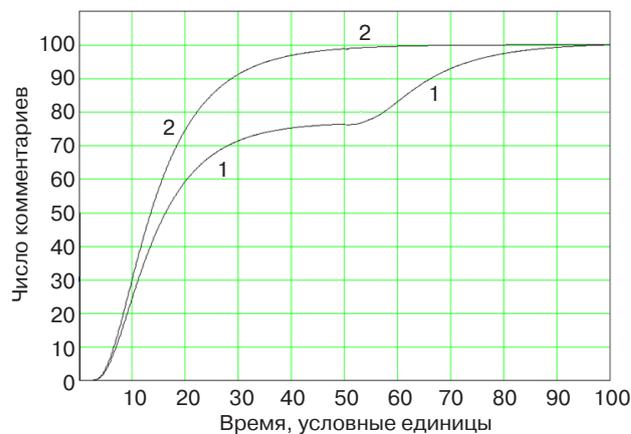
$$\begin{aligned} \rho(x, t) = \int \frac{\left[\frac{[\ln x]^2}{D_0 t} + \left[\frac{1}{2} - \frac{\mu_0}{D_0} \right] \ln x - 1 \right]}{\sqrt{2\pi D_0 t^3}} \times \\ \times e^{-\left[\frac{[\ln x]^2}{2D_0 t} + \left[\frac{3}{2} \frac{\mu_0}{D_0} \right] \ln x + \left[\frac{1}{2} - \frac{\mu_0}{D_0} \right]^2 \frac{D_0 t}{2} \right]} dt. \end{aligned} \quad (4)$$

Вероятность того, что число комментариев к моменту времени t достигнет некоторого числа L , можно найти по формуле: $P(L, t) = 1 - \int_0^L \rho(x, t) dx$. Зависимость числа комментариев от времени t будет описываться уравнением: $N(t) = P(L, t)L$.

Для анализа полученного решения проведем имитационное моделирование. В качестве примера выберем $L = 100$ и три набора значений μ_0 и D_0 : $\mu_0 = 0.45$ и $D_0 = 0.50$ условных единиц ($\mu_0 < D_0$ – кривая 1 на рис. 3а), $\mu_0 = 0.50$ и $D_0 = 0.50$ условных единиц ($\mu_0 = D_0$ – кривая 2 на рис. 3)



(а)



(б)

Рис. 3. Динамика изменения с течением времени числа комментариев к новости в имитационной модели на основе уравнения Фоккера – Планка:

(а) 1 процесс, (б) 2 параллельных процесса

и $\mu_0 = 0.55$ и $D_0 = 0.50$ условных единиц ($\mu_0 > D_0$ – кривая 3 на рис. 3а). Расчеты показывают, что с ростом μ_0 относительно D_0 скорость роста кривых для числа комментариев $N(t)$ при выбранных значениях параметров модели μ_0 , D_0 и L увеличивается (рис. 3а).

Двухступенчатую кривую можно получить, если использовать функцию плотности распределения, учитывающую время задержки τ :

$$\begin{aligned} \rho(x, t - \tau) = \int \frac{\left[\frac{[\ln x]^2}{D_0(t - \tau)} + \left[\frac{1}{2} - \frac{\mu_0}{D_0} \right] \ln x - 1 \right]}{\sqrt{2\pi D_0 [t - \tau]^3}} \times \\ \times e^{-\left[\frac{[\ln x]^2}{2D_0 [t - \tau]} + \left[\frac{3}{2} \frac{\mu_0}{D_0} \right] \ln x + \left[\frac{1}{2} - \frac{\mu_0}{D_0} \right]^2 \frac{D_0(t - \tau)}{2} \right]} dt. \end{aligned} \quad (5)$$

Подтверждение соответствия теоретической модели и наблюдаемых данных (рис. 2 и 3б) можно получить, если предположить, что могут одновременно

протекают два процесса с различными коэффициентами μ_0 и D_0 . Причем сумма парциальных долей процессов должна быть равна 1, т.е. $P_{\text{общ}}(L, t) = \alpha_1 P_1(L, t) + \alpha_2 P_2(L, t)$. В качестве примера моделирования выберем для процесса комментирования самой новости или блога следующие параметры модели: $\mu_{0,1} = 0.55$, $D_{0,1} = 0.50$, а для комментирования комментариев: $\mu_{0,2} = 0.50$, $D_{0,2} = 0.50$, $\tau = 50$ условных единиц, $\alpha_1 = 0.75$, $\alpha_2 = 0.25$ и $L = 100$ ($\mu_{0,1} > \mu_{0,2}$ было выбрано, исходя из предположения, что комментирование новости является более первичным процессом для пользователей, чем комментирование комментариев), где $\alpha_1 + \alpha_2 = 1$. Отметим, что это просто пример имитационного моделирования, чтобы показать свойства модели.

Как было указано ранее, плато на кривой 1 на рис. 3б может быть связано с существенным различием в среднем времени появления комментариев второго уровня (интервал времени между появлением комментария первого уровня и комментарием к данному комментарию), что может приводить к реализации двухступенчатой динамики в появлении комментариев.

На рис. 3б представлены результаты моделирования динамики изменения с течением времени числа комментариев $N(t)$ с учетом того, что два процесса могут протекать параллельно. Как видно из полученных результатов, наблюдается хорошее совпадение реальных данных и теоретических расчетов.

4. МОДЕЛИРОВАНИЕ ДИНАМИКИ НАСТРОЕНИЙ ПОЛЬЗОВАТЕЛЕЙ ИНТЕРНЕТ-МАССМЕДИА НА ОСНОВЕ УРАВНЕНИЯ ФОККЕРА – ПЛАНКА И ИЗМЕНЕНИЯ ПАРАМЕТРОВ ГРАФОВ СЕТЕЙ ИХ КОММЕНТАРИЕВ

Для описания состояний сети комментариев предлагается использовать векторное представление. Элементами векторов являются допустимые значения параметров сети (плотность, средняя величина коэффициента посредничества, среднее значение коэффициента кластеризации, эластичность и другие), а также такие характеристики как доля пользователей, которые могут быть отнесены на основе анализа текстов комментариев к одной из групп: 1) лоялист (безусловно поддерживает действия правительства и органов власти); 2) оппозиционер; 3) «тролль» (пользователь, использующий ресурс для того, чтобы устроить скандал); 4) неопределившийся или нейтральный. Достижение или реализация желаемых или нежелательных состояний всей социальной сети в целом могут быть на основе базовых векторов.

Изменение во времени величины расстояния между базовым вектором и вектором текущего состояния, можно рассматривать как «блуждания точки» на отрезке $[L_{\min}, L_{\max}]$ или как случайный (или

почти случайный) временной ряд. Некоторое заданное значение этого расстояния (состояние, при достижении которого необходимо принимать управленческие решения) можно рассматривать как ловушку или точку допустимого порога реализации, куда может с течением времени попасть «блуждающая точка». Это позволяет построить вероятностные социодинамические модели для прогноза динамики общественных настроений.

При традиционном описании поведения «блуждающей точки», как правило, используется диффузионная модель. Однако в данном случае она не может считаться надежной. Как правило, временные ряды, описывающие процессы в сложных системах (например, финансовые показатели фондовых и сырьевых бирж), не являются стационарными, что обусловлено разными причинами, в т.ч. наличием человеческого фактора. Их выборочные функции распределения имеют зависящее от времени математическое ожидание, что противоречит простой диффузионной модели и показывает нестационарность временных рядов.

В этой связи предполагается рассмотреть более сложные модели поведения «блуждающей точки», например, на основе уравнений Фоккера – Планка.

На рис. 4 представлена структура графа, полученного при обработке комментариев к новости. При цветовой визуализации полученных данных узлы графа, в зависимости от их состояния (отнесения к одному из 4 типов), могут быть размечены разными цветами (оппозиционеры – зелено-песочного цвета, «тролли» – красного, лоялисты – синего, неопределившиеся – зеленого цвета). Связи на рисунке показывают взаимное комментирование пользователей друг другом. Таким образом, по цвету узлов можно судить о их состоянии, а по ребрам графа – о взаимодействии.

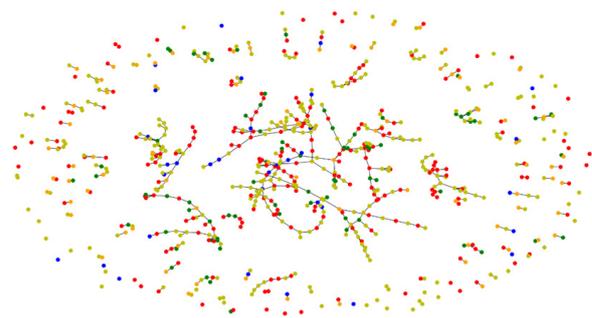


Рис. 4. Структура графа комментариев к рассматриваемой новости

На рис. 4 видно, что в данной структуре имеется много несвязанных одиночных вершин. Однако можно заметить и присутствие связанной компоненты графа, которая отдельно представлена на рис. 5. Замкнутые овалы на этом рисунке (это ядро графа на рис. 4) показывают комментирование пользователями самих себя.

Рассмотрим элементы вектора состояния сети, которые мы будем использовать в нашей модели.

1. Доля узлов, имеющих определенное состояние (например, тех, кто настроен отрицательно по отношению к какому-либо событию общественной жизни).
2. Коэффициент кластеризации – мера плотности связей данной вершины с соседними. Отношение реального количества связей, которые соединяют ближайших соседей данного узла i , к максимально возможному (такому, при котором все ближайшие соседи данного узла были бы соединены непосредственно друг с другом) называется коэффициентом кластеризации узла, его величина лежит на отрезке $[0, 1]$. Чем больше его величина, тем более значимым является данный узел в процессе обмена информацией.
3. Степень посредничества показывает отношение числа кратчайших путей между всеми парами узлов сети, проходящих через данный узел, к общему числу всех кратчайших путей в сети. Ее величина лежит на отрезке $[0, 1]$. Чем больше ее величина, тем значительнее роль данной вершины в обмене информацией.

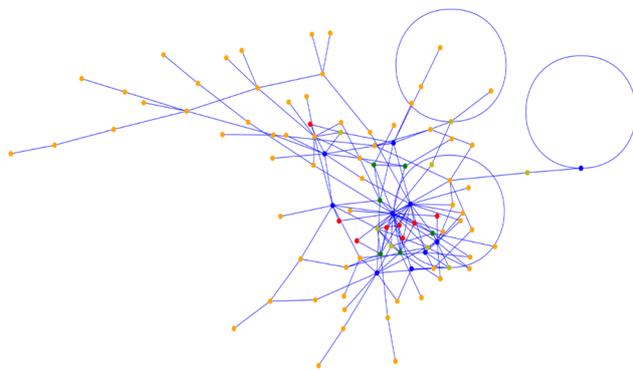


Рис. 5. Граф взаимосвязей пользователей по комментариям

Определим величины элементов базового вектора состояния сети (обозначим его θ). Они задают пороговые значения, переход через которые является с точки зрения управления состояниями нежелательным. Учитывая, что в любом сообществе всегда существует от 0.10 до 0.15 доли участников, всегда несогласных по любому вопросу, примем долю тех, кто настроен отрицательно по отношению к рассматриваемому событию, равной 0.12. Желаемое среднее значение (по всем узлам) коэффициента кластеризации такой сети тоже примем небольшим, например, равным 0.05, и значение средней степени посредничества узлов в такой сети тоже равным 0.05. Таким образом, базовый вектор будет иметь вид: $\theta = (0.12; 0.05; 0.05)$.

Заметим, что число параметров, с помощью которых можно описывать состояние сети, может быть больше. Были выбраны только те, которые, на наш взгляд, являются наиболее значимыми. Кроме того, выбранные параметры являются нормированными (лежат на отрезке $[0, 1]$), поэтому в одинаковой мере влияют на расчет метрики расстояний.

В предлагаемом нами подходе различные графы комментирования новостей на выбранном ресурсе по определенным тематикам в течении суток могут быть объединены в единую структуру через связи между узлами, которые относятся к пользователям. Таким образом можно выделить большой граф, который будет описывать активность пользователей данного сетевого информационного ресурса в течение суток. Далее можно определить элементы вектора текущего состояния, который описывает его характеристики.

Изменения компонентов этого вектора за каждые сутки в течение определенного времени будут образовывать временной ряд.

Для описания изменения значения расстояния между величиной текущего вектора состояния и заданного базового вектора с течением времени рассмотрим решение нестационарного уравнения Фоккера – Планка, что может позволить построить вероятностные социодинамические модели для прогнозирования динамики общественных настроений.

Сформулируем краевую задачу, решение которой будет описывать процесс изменения значения расстояния между величиной текущего вектора состояния графа сети комментариев и заданного базового вектора с течением времени.

Первое граничное условие. При выборе первого граничного условия будем исходить из следующих соображений: состояние $x = L_{\min}$ (левая граница отрезка возможных состояний) определяет состояние, переход через которое необходимо избежать (область, находящаяся левее данного состояния на отрезке, является нежелательной). Вероятность обнаружить такое состояние системы может быть не равна 0, а плотность вероятности, которая определяет поток в состоянии $x = L_{\min}$, нужно взять равной 0, т.к. состояния не должны выходить за эту границу (здесь реализуется условие отражения). Таким образом:

$$\rho(x, t)|_{x=L_{\min}} = 0. \quad (6a)$$

Второе граничное условие. Ограничим область возможных состояний справа некоторой величиной $x = L_{\max}$ (метрика, используемая при расчетах, не может быть больше величины вектора,

элементы которого имеют максимальные значения в пространстве выбранных координат). Вероятность обнаружить такое состояние с течением времени будет отлична от нуля. Однако плотность вероятности, определяющую поток в состоянии $x = L_{\max}$, необходимо положить равной нулю (расстояние между текущим и базовым вектором состояний ограничено максимальными значениями возможных координат в используемом векторном пространстве (реализуется условие отражения от границы)):

$$\rho(x, t)|_{x=L_{\max}} = 0. \quad (66)$$

Для формулировки краевой задачи необходимо задать начальное условие. Так как в момент времени $t = 0$ состояние системы (расстояние между базовым вектором и вектором текущего состояния) может быть равно некоторому значению x_0 , то начальное условие можно задать в виде:

$$\rho(x, t = 0) = \begin{cases} \int \delta(x - x_0) dx = 1, & x = x_0, \\ 0, & x \neq x_0. \end{cases} \quad (67)$$

Кратко описание решения данной краевой задачи имеет следующий вид. Используя метод преобразований Лапласа для уравнения Фоккера – Планка, можно записать:

$$\begin{aligned} s\overline{G(s, x)} - \rho(x, 0) = \\ = -\frac{d}{dx} [\mu(x)s\overline{G(s, x)}] + \frac{1}{2} \cdot \frac{d^2}{dx^2} [D(x)s\overline{G(s, x)}]. \end{aligned} \quad (7)$$

Подставив в уравнение (7) соответствующие производные и зависимости $\mu(x)$ и $D(x)$, получим:

$$\begin{aligned} x^2 \frac{d^2 \overline{G(s, x)}}{dx^2} + 2 \left[2 - \frac{\mu_0}{D_0} \right] x \frac{d \overline{G(s, x)}}{dx} + \\ + 2 \left[1 - \frac{\mu_0 + s}{D_0} \right] \overline{G(s, x)} = -\frac{\delta(x - x_0)}{D_0}. \end{aligned} \quad (8)$$

Решение данного уравнения ищем в виде: $\overline{G(s, x)} = \sum_k C_k x^q$, где C_k – коэффициенты для корней q характеристического уравнения:

$$q(q-1) + 2 \left[2 - \frac{\mu_0}{D_0} \right] q + 2 \left[1 - \frac{\mu_0 + s}{D_0} \right] = 0.$$

Решение будем искать в виде двух функций: $\overline{G_1(s, x)}$ на отрезке $[L, x_0]$ и $\overline{G_2(s, x)}$ на отрезке $[x_0, 1]$ с использованием условия сшивки в точке $x = x_0$ функций $\overline{G_1(s, x)}$ и $\overline{G_2(s, x)}$. Наличие в уравнении δ -функции приводит к тому, что решение, оставаясь непрерывным в точке $x = x_0$, испытывает в ней разрыв производной.

При $L_{\min} \leq x \leq x_0$

$$\overline{G_1(s, x)} = C_1 x^{-[1+\alpha]+\beta\sqrt{k+s}} + C_2 x^{-[1+\alpha]-\beta\sqrt{k+s}}.$$

При $x_0 \leq x \leq L_{\max}$

$$\overline{G_2(s, x)} = C_3 x^{-[1+\alpha]+\beta\sqrt{k+s}} + C_4 x^{-[1+\alpha]-\beta\sqrt{k+s}}.$$

После соответствующих преобразований получим:

- при $L_{\min} \leq x \leq x_0$

$$\begin{aligned} \overline{G_1(s, x)} = \frac{x_0^\alpha x^{-[1+\alpha] \operatorname{sh} \left\{ \left[\beta \ln \left(\frac{L_{\max}}{x_0} \right) \right] \sqrt{k+s} \right\}}}{\beta D_0 \sqrt{k+s}} \times \\ \times \frac{\operatorname{sh} \left\{ \left[\beta \ln \left(\frac{x}{L_{\min}} \right) \right] \sqrt{k+s} \right\}}{\operatorname{sh} \left\{ \left[\beta \ln \left(\frac{L_{\max}}{L_{\min}} \right) \right] \sqrt{k+s} \right\}}; \end{aligned}$$

- при $x_0 \leq x \leq L_{\max}$

$$\begin{aligned} \overline{G_2(s, x)} = \frac{x_0^\alpha x^{-[1+\alpha] \operatorname{sh} \left\{ \left[\beta \ln \left(\frac{x_0}{L_{\min}} \right) \right] \sqrt{k+s} \right\}}}{\beta D_0 \sqrt{k+s}} \times \\ \times \frac{\operatorname{sh} \left\{ \left[\beta \ln \left(\frac{x}{L_{\max}} \right) \right] \sqrt{k+s} \right\}}{\operatorname{sh} \left\{ \left[\beta \ln \left(\frac{L_{\max}}{L_{\min}} \right) \right] \sqrt{k+s} \right\}}. \end{aligned}$$

Сделаем обратные преобразования Лапласа:

$$\overline{G_1(s, x)} = \frac{A(s)}{B(s)} = \sum_{n=1}^M \frac{A(s_n)}{B(s_n)} \cdot e^{s_n t},$$

где s_n – простые, отличные от нуля корни $B(s)$,

$$B(s) = \sqrt{k+s} \cdot \operatorname{sh} \left\{ \left[\beta \ln L \right] \sqrt{k+s} \right\},$$

$$A(s) = \operatorname{sh} \left\{ \left[\beta \ln x_0 \right] \sqrt{k+s} \right\} \cdot \operatorname{sh} \left\{ \left[\beta \ln \frac{x}{L} \right] \sqrt{k+s} \right\}.$$

После соответствующих преобразований получим при $L_{\min} \leq x \leq x_0$:

$$\rho_1(x, t) = 2 \cdot \frac{x_0^\alpha x^{-[1+\alpha]} e^{-kt}}{D_0 \beta^2 \ln\left(\frac{L_{\max}}{L_{\min}}\right)} \cdot \sum_{n=1}^M \frac{\sin\left\{\pi n \frac{\ln\left(\frac{L_{\max}}{x_0}\right)}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)}\right\} \cdot \sin\left\{\pi n \frac{\ln\left(\frac{x}{L_{\min}}\right)}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)}\right\}}{\cos \pi n} \cdot e^{-\frac{\pi^2 n^2 t}{\left[\beta \ln\left(\frac{L_{\max}}{L_{\min}}\right)\right]^2}}.$$

Аналогично при $x_0 \leq x \leq L_{\max}$

$$\rho_2(x, t) = -4 \cdot \frac{x_0^\alpha x^{-[1+\alpha]} e^{-kt}}{D_0 \beta^2 \ln L} \cdot \sum_{n=1}^M \frac{\sin\left\{\pi n \frac{\ln\left(\frac{x_0}{L}\right)}{\ln L}\right\} \cdot \sin\left\{\pi n \frac{\ln x}{\ln L}\right\}}{\cos \pi n} \cdot e^{-\frac{\pi^2 n^2 t}{[\beta \ln L]^2}}.$$

С учетом $\alpha = \frac{1-2 \cdot \frac{\mu_0}{D_0}}{2} = \frac{1}{2} - \frac{\mu_0}{D_0}$, $\beta^2 = \frac{2}{D_0}$ и $k = \frac{\alpha^2}{\beta^2}$ получим:

- при $L_{\min} \leq x \leq x_0$:

$$\rho_1(x, t) = -2 \cdot \frac{x_0^\alpha x^{-[1+\alpha]} \cdot e^{-\frac{D_0 \alpha^2}{2} t}}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)} \cdot \sum_{n=1}^M \frac{\sin\left\{\pi n \frac{\ln\left(\frac{L_{\max}}{x_0}\right)}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)}\right\} \cdot \sin\left\{\pi n \frac{\ln\left(\frac{x}{L_{\min}}\right)}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)}\right\}}{\cos \pi n} \cdot e^{-\frac{\pi^2 n^2 D_0 t}{2 \left[\ln\left(\frac{L_{\max}}{L_{\min}}\right)\right]^2}}; \quad (9a)$$

- при $x_0 \leq x \leq L_{\max}$:

$$\rho_2(x, t) = 2 \cdot \frac{x_0^\alpha x^{-[1+\alpha]} \cdot e^{-\frac{D_0 \alpha^2}{2} t}}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)} \cdot \sum_{n=1}^M \frac{\sin\left\{\pi n \frac{\ln\left(\frac{x_0}{L_{\min}}\right)}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)}\right\} \cdot \sin\left\{\pi n \frac{\ln\left(\frac{x}{L_{\max}}\right)}{\ln\left(\frac{L_{\max}}{L_{\min}}\right)}\right\}}{\cos \pi n} \cdot e^{-\frac{\pi^2 n^2 D_0 t}{2 \left[\ln\left(\frac{L_{\max}}{L_{\min}}\right)\right]^2}}, \quad \alpha = \frac{1}{2} - \frac{\mu_0}{D_0}. \quad (9б)$$

Вероятность того, что к моменту времени t состояние системы будет находиться на отрезке от L_{\min} до L_{\max} , т.е. пороговое состояние (θ) не будет достигнуто, можно вычислить следующим образом:

$$P(\theta, t) = \int_{L_{\min}}^{x_0} \rho_2(x, t) dx + \int_{x_0}^{L_{\max}} \rho_1(x, t) dx. \quad (10)$$

Вероятность $Q(\theta, t)$ того, что пороговое состояние θ к моменту времени t окажется достигнутым или превзойденным, вычисляется по формуле:

$$Q(\theta, t) = 1 - P(\theta, t). \quad (11)$$

Определение границ отрезка возможных состояний от L_{\min} до L_{\max} обсудим позже.

Используя средства sentiment-анализа, все комментарии пользователей в течение суток могут быть размечены на позитивные и негативные, и можно таким образом получить временной ряд. На рис. 6, представлен временной ряд суточной активности пользователей (имеющих негативное отношение) официальной страницы «РИА Новости» в социальной сети «ВКонтакте» по комментированию новостей с января 2013 г. по декабрь 2020 г. Активность определялась как отношение общего числа уникальных комментариев, оставленных пользователями ко всем новостям в течение суток, к общему

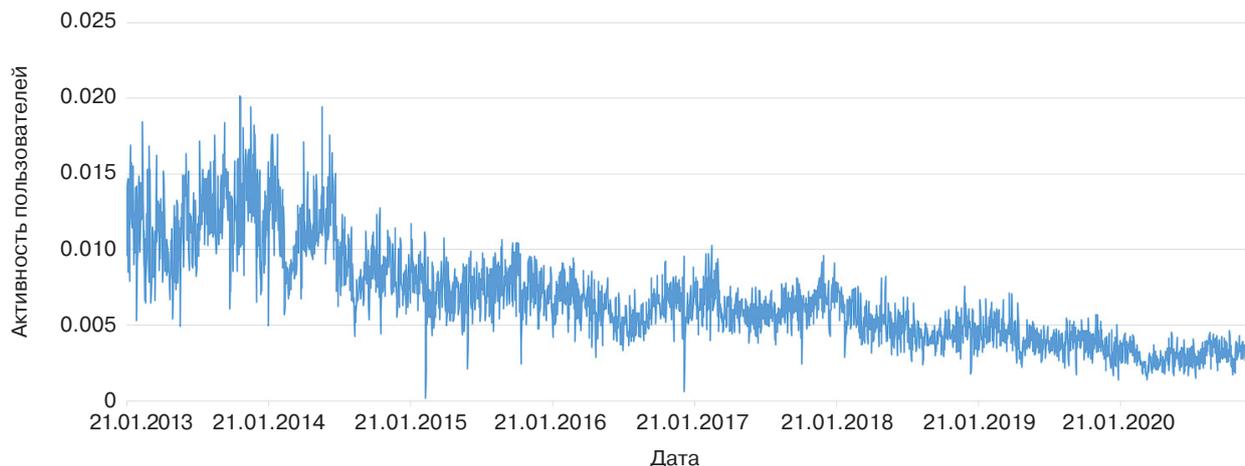


Рис. 6. Временной ряд активности пользователей портала «РИА Новости» по комментированию новостей с января 2013 г. по декабрь 2020 г.

количеству уникальных просмотров пользователями в течение суток всех опубликованных новостей.

Необходимо отметить, что посетители в целом очень неохотно оставляют комментарии. Общее число тех, кто их делает, не превышает 1–2% от всех просмотров портала.

Можно построить объединенный за сутки граф комментариев, определить его характеристики и вектор состояния (он изменяется от суток к суткам).

Обсудим выбор границ отрезка возможных состояний от L_{\min} до L_{\max} . Для графа комментариев пользователей портала «РИА Новости», элементы вектора текущего состояния в момент времени 15 мая 2019 г. (который принимаем за $t = 0$) будут равны $X(t) = (0.0035; 0.07; 0.12)$: доля негативных комментариев равна 0.0035; среднее значение коэффициента кластеризации равно 0.07; значение средней степени посредничества узлов равно 0.12.

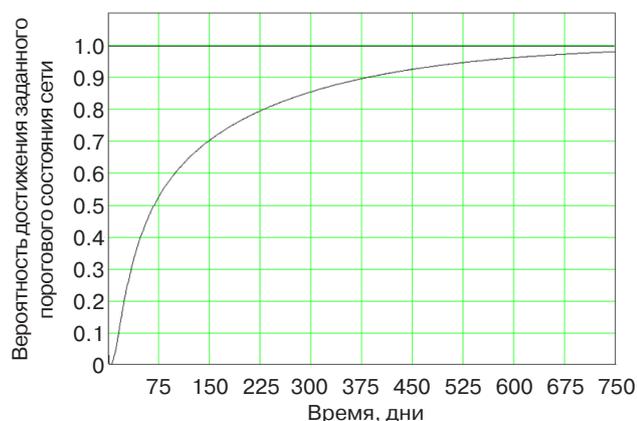


Рис. 7. Зависимость вероятности достижения заданного порогового состояния сети от времени для рассмотренного примера в момент времени 15 мая 2019 г.

Базовый вектор желаемого состояния зададим в виде $\theta = (0.0025; 0.05; 0.05)$: доля негативных комментариев равна 0.0025; среднее значение коэффициента кластеризации равно 0.05; значение средней степени посредничества узлов равно 0.05. Значение доли негативных комментариев, равное 0.0025, выбрано для проверки модели, т.к. известно, когда это значение будет действительно наблюдаться на графике временного ряда. Сравнив результаты моделирования прогноза и реальное значение, можно оценить адекватность модели. Расстояние между заданным базовым вектором желаемого состояния $\theta = (0.0025; 0.05; 0.05)$ и вектором текущего состояния $X(t)$ в момент времени $t = 0$ будет равно $x_0 = 0.073$. Проведя анализ динамики временного ряда изменения состояния сети за несколько предыдущих дней и используя уравнения модели, можно решить обратную задачу и определить величины параметров модели μ_0 и D_0 . В данном случае $\mu_0 = 0.0002$ и $D_0 = 0.0009$.

Правая граница отрезка возможных состояний L_{\max} может быть задана как расстояние между базовым вектором (θ) и вектором максимально возможных значений параметров сети $X(t) = (1; 1; 1)$. В рассматриваемом случае $L_{\max} = 1.70$. Левую границу для надежности рассмотрения можно определить, например, как половину длины заданного базового вектора (в данном случае $|\theta| = 0.071$), таким образом, L_{\min} будет равно 0.035.

Как показывают полученные результаты, если не оказывать на сеть пользователей никакого воздействия, то в сложившихся условиях требуемое состояние может быть достигнуто с вероятностью 0.8 на 225 день и с вероятностью 0.9 на 375 день (рис. 7), что примерно и произошло в действительности, судя по наблюдаемым данным.

5. АЛГОРИТМ ПРОГНОЗИРОВАНИЯ ДОСТИЖЕНИЯ ЗАДАННОГО СОСТОЯНИЯ ГРАФА СЕТИ КОММЕНТАРИЕВ ПОЛЬЗОВАТЕЛЕЙ НОВОСТНЫХ МАССМЕДИА

Проведенные исследования и разработанная модель динамики изменения настроений пользователей интернет-массмедиа на основе уравнения Фоккера – Планка и изменения параметров графов сетей их комментариев позволяют создать алгоритм их прогнозирования, который состоит из следующих шагов:

- а) Сбор с привязкой к дате и времени текстовых комментариев и метаданных пользователей по определенной тематике с новостных сетевых медиа-ресурсов.
- б) Обработка данных с использованием средств текстовой аналитики и сентимент-анализ, получение графа комментариев пользователей по определенной тематике и расчет его характеристик (плотность сети, средняя величина коэффициента посредничества, среднее значение коэффициента кластеризации, эластичность, доля пользователей с тем или иным настроением).
- в) Задание значения элементов базового вектора, который будет определять достижение желаемого или нежелательного состояния (θ), и формирование на основе обработанных данных и заданного вектора временного ряда изменения графа комментариев пользователей по определенной тематике с течением времени.
- г) Задание длительности шага τ (час, день, неделя и т.д.). По значениям временного ряда на нескольких шагах для заданного τ определение с помощью численных расчетов с использованием наблюдаемых данных и уравнений (5) и (6) параметров модели μ_0 и D_0 .
- д) Принятие последнего среднего значения метрики расстояний между базовым вектором и вектором текущего состояния сети за начальное состояние x_0 и с использованием полученных значений μ_0 и D_0 , а также уравнений (5) и (6). Проведение расчетов и получение зависимости от времени вероятности достижения желаемого или нежелательного состояния. Далее можно задать величину вероятности (например, 0.95) и оценить время достижения заданного уровня вероятности (сделать прогноз по времени).

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

Результаты, полученные в работе:

1. Наблюдаемое на практике стационарное распределение новостей по числу комментариев к ним соответствует степенному закону: $\rho(x) = [\gamma - 1]x^{-\gamma}$, где $\rho(x)$ – доля новостей в общем их числе,

имеющая x комментариев, а γ – показатель степени.

2. Динамика изменения с течением времени числа комментариев к новости или блогу может иметь как S-образный вид, так и двухступенчатый, что может быть связано с существенным различием в среднем времени появления комментариев второго уровня (интервал времени между появлением комментария первого уровня и комментарием к данному комментарию), т.е. величиной средней задержки.
3. Наблюдаемый на практике степенной закон зависимости стационарной плотности вероятности распределения новостей по числу комментариев (состояниям x) может быть получен из решения стационарного уравнения Фоккера – Планка, если при его выводе сделать ряд допущений. В частности, предположить, что коэффициент $\mu(x)$, отвечающий в уравнении Фоккера – Планка за целенаправленное изменение состояния системы x (x – текущее число комментариев к новости), линейно зависит от состояния x , а коэффициент $D(x)$, отвечающий за случайное изменение, зависит от x квадратично. Все это позволяет предположить, что уравнение Фоккера – Планка может быть использовано для описания процессов в сложных сетевых структурах.
4. Решение нестационарного уравнения Фоккера – Планка при допущениях о линейной зависимости $\mu(x)$ от состояния x и квадратичной зависимости $D(x)$ от состояния x позволяет получить уравнение для плотности вероятности переходов между состояниями системы в единицу времени, которое хорошо согласуется с наблюдаемыми данными с учетом влияния времени задержки между появлением комментария первого уровня и комментарием к данному комментарию.
5. Разработанные на основе уравнения Фоккера – Планка модели хорошо согласуются с наблюдаемыми данными, что позволяет создать алгоритмы мониторинга и прогнозирования эволюции общественного мнения пользователей новостных информационных ресурсов.
6. Параметры графов сетей комментариев могут быть получены при использовании как готовых инструментов и библиотек языка Python, применяемых для анализа сложных сетей, так и с помощью дополнительно разработанного программного обеспечения.
7. Анализ полученной модели с использованием характеристик реального временного ряда для изменения графа комментариев пользователей портала «РИА Новости» и структурных параметров графа показывает ее адекватность и непротиворечивость. Как показывают полученные

результаты, если не оказывать на сеть пользователей никакого воздействия, то в сложившихся условиях требуемое состояние по числу негативно настроенных комментаторов может быть достигнуто с вероятностью 0.8 на 225 день, и с вероятностью 0.9 на 375 день, что примерно и произошло в действительности, судя по наблюдаемым данным.

8. Созданная на основе уравнения Фоккера – Планка модель позволяет разработать алгоритм прогнозирования достижения заданного состояния графа сети комментариев пользователей новостных массмедиа, что является практически очень важным.

В заключении отметим, что сложный характер динамики процессов в сложных социальных системах можно описывать не только на основе моделей, созданных на основе уравнения Фоккера – Планка. Например, в работах [18–21] представлены разработанные авторами модели описания стохастической динамики изменения состояний в сложных социальных системах, учитывающие процессы самоорганизации и наличие памяти. Для создания модели были рассмотрены графические схемы вероятностей переходов между возможными состояниями описываемых систем с учетом предыдущих

состояний, что позволяет учитывать память и описывать не только марковские, но и немарковские процессы.

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке Российского научного фонда, грант № 22-21-00109 «Разработка моделей прогнозирования динамики социальных настроений на основе анализа временных рядов текстового контента социальных сетей с использованием уравнений Фоккера – Планка и нелинейной диффузии».

ACKNOWLEDGMENTS

The study is financially supported by the Russian Science Foundation, grant No. 22-21-00109 “Development of models for forecasting the dynamics of social sentiments based on analyzing time series of text content of social networks using the Fokker–Planck equations and nonlinear diffusion.”

Вклад авторов. Все авторы в равной степени внесли свой вклад в исследовательскую работу.

Authors' contribution. All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Du B., Lian X., Cheng X. Partial differential equation modeling with Dirichlet boundary conditions on social networks. *Bound. Value Probl.* 2018;2018(1):50. <https://doi.org/10.1186/s13661-018-0964-4>
2. Lux T. Inference for systems of stochastic differential equations from discretely sampled data: a numerical maximum likelihood approach. *Ann. Finance.* 2013;9(2):217–248. <http://doi.org/10.1007/s10436-012-0219-9>
3. Hurn A., Jeisman J., Lindsay K. Teaching an Old Dog New Tricks: Improved Estimation of the Parameters of Stochastic Differential Equations by Numerical Solution of the Fokker–Planck Equation. In: Dungey M., Bardsley P. (Eds.). *Proceedings of the Australasian Meeting of the Econometric Society.* 2006. The Australian National University, Australia. P. 1–36.
4. Elliott R.J., Siu T.K., Chan L. A PDE approach for risk measures for derivatives with regime switching. *Ann. Finance.* 2007;4(1):55–74. <http://dx.doi.org/10.1007/s10436-006-0068-5>
5. Орлов Ю.Н., Федоров С.Л. Генерация нестационарных траекторий временного ряда на основе уравнения Фоккера–Планка. *ТРУДЫ МФТИ.* 2016;8(2):126–133. [Orlov Y.N., Fedorov S.L. Nonstationary time series trajectories generation on the basis of the Fokker–Planck equation. *TRUDY MFTI = Proceedings of MIPT.* 2016;8(2):126–133 (in Russ.).]
6. Chen Y., Cosimano T.F., Himonas A.A., Kelly P. An Analytic Approach for Stochastic Differential Utility for Endowment and Production Economies. *Comput. Econ.* 2013;44(4):397–443. <http://doi.org/10.1007/s10614-013-9397-4>
7. Savku E., Weber G.-W. Stochastic differential games for optimal investment problems in a Markov regime-switching jump-diffusion market. *Ann. Oper. Res.* 2020;132(6):1171–1196. <https://doi.org/10.1007/s10479-020-03768-5>
8. Красников К.Е. Математическое моделирование некоторых социальных процессов с помощью теоретико-игровых подходов и принятие на их основе управленческих решений. *Russ. Technol. J.* 2021;9(5):67–83. <https://doi.org/10.32362/2500-316X-2021-9-5-67-83> [Krasnikov K.E. Mathematical modeling of some social processes using game-theoretic approaches and making managerial decisions based on them. *Russ. Technol. J.* 2021;9(5):67–83 (in Russ.). <https://doi.org/10.32362/2500-316X-2021-9-5-67-83>]
9. Kim S.L., Hinders M.K. Dynamic wavelet fingerprint for differentiation of tweet storm types. *Soc. Netw. Anal. Min.* 2020;10(1):4. <https://doi.org/10.1007/s13278-019-0617-3>
10. Hoffmann T., Peel L., Lambiotte R., Jones N.S. Community detection in networks without observing edges. *Sci. Adv.* 2020;6(4):1478. <https://doi.org/10.1126/sciadv.aav1478>

11. Pulipati S., Somula R., Parvathala B.R. Nature inspired link prediction and community detection algorithms for social networks: a survey. *Int. J. Syst. Assur. Eng. Manag.* 2021. <https://doi.org/10.1007/s13198-021-01125-8>
12. Dorogovtsev S.N., Mendes J.F.F. Evolution of networks. *Adv. Phys.* 2002;51(4):1079–1187. <http://doi.org/10.1080/00018730110112519>
13. Newman M.E.J. The structure and function of complex networks. *SIAM Rev.* 2003;45(2):167–256. <https://doi.org/10.1137/S003614450342480>
14. Dorogovtsev S.N., Mendes J.F.F., Samukhin A.N. Generic scale of the scale-free growing networks. *Phys. Rev. E.* 2001;63(6):062101. <https://doi.org/10.1103/PhysRevE.63.062101>
15. Golder S., Wilkinson D., Huberman B. Rhythms of Social Interaction: Messaging Within a Massive Online Network. In: Steinfield C., Pentland B.T., Ackerman M., Contractor N. (Eds.). *Communities and Technologies*. 2007. https://doi.org/10.1007/978-1-84628-905-7_3
16. Kumar R., Novak J., Tomkins A. Structure and evolution of online social networks. In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and data Mining, KDD '06*. 2006. P. 611–617. <https://doi.org/10.1145/1150402.1150476>
17. Mislove A., Marcon M., Gummadi K.P., Druschel P., Bhattacharjee B. Measurement and analysis of online social networks. In: *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, IMC'07*. 2007. P. 29–42. <https://doi.org/10.1145/1298306.1298311>
18. Zhukov D., Khvatova T., Zaltsman A. Stochastic Dynamics of Influence Expansion in Social Networks and Managing Users' Transitions from One State to Another. In: *Proceedings of the 11th European Conference on Information Systems Management (ECISM 2017)*. 2017. P. 322–329.
19. Zhukov D., Khvatova T., Millar C., Zaltzman A. Modelling the stochastic dynamics of transitions between states in social systems incorporating self-organization and memory. *Technol. Forecast. Soc. Change.* 2020;158:120134. <https://doi.org/10.1016/j.techfore.2020.120134>
20. Zhukov D., Khvatova T., Istratov L. A stochastic dynamics model for shaping stock indexes using self-organization processes, memory and oscillations. In: *Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics (ECIAIR 2019)*. 2019. P. 390–401.
21. Zhukov D.O., Zaltzman A.D., Khvatova T.Yu. Forecasting Changes in States in Social Networks and Sentiment Security Using the Principles of Percolation Theory and Stochastic Dynamics. In: *Proceedings of the 2019 IEEE International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS)*. 2019. Article number 8928295. P. 149–153. <https://doi.org/10.1109/ITQMIS.2019.8928295>

Об авторах

Перова Юлия Петровна, старший преподаватель, кафедра телекоммуникаций, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: jul-np@yandex.ru. Scopus Author ID 57431908700, <https://orcid.org/0000-0003-4028-2842>

Лесько Сергей Александрович, д.т.н., доцент, профессор кафедры «Информационно-аналитические системы кибербезопасности», Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: sergey@testor.ru. Scopus Author ID 57189664364, <https://orcid.org/0000-0002-6641-1609>

Иванов Андрей Андреевич, магистрант, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: heliosgoodgame@gmail.com. <https://orcid.org/0009-0002-7199-2871>

About the authors

Julia P. Perova, Senior Lecturer, Department of Telecommunications, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: jul-np@yandex.ru. Scopus Author ID 57431908700, <https://orcid.org/0000-0003-4028-2842>

Sergey A. Lesko, Dr. Sci. (Eng.), Docent, Professor of the Cybersecurity Information and Analytical Systems Department, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: sergey@testor.ru. Scopus Author ID 57189664364, <https://orcid.org/0000-0002-6641-1609>

Andrey A. Ivanov, Student, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: heliosgoodgame@gmail.com. <https://orcid.org/0009-0002-7199-2871>

Математическое моделирование
Mathematical modeling

УДК 330.4

<https://doi.org/10.32362/2500-316X-2024-12-3-93-103>

EDN YSWUJG



НАУЧНАЯ СТАТЬЯ

Анализ подходов к определению тренда в структуре временного ряда

У.С. Мохнаткина,
Д.В. Парфенов,
Д.А. Петрусевич @

МИРЭА – Российский технологический университет, Москва, 119454 Россия

@ Автор для переписки, e-mail: petrusevich@mirea.ru, petrdenis@mail.ru

Резюме

Цели. Основная цель – сравнить качество прогнозирования моделей временных рядов, по-разному описывающих тренд, и сформировать заключение о применимости каждого подхода при описании тренда в зависимости от свойств временного ряда.

Методы. Тренд может рассматриваться как склонность рассматриваемой величины к возрастанию или убыванию в долгосрочной перспективе. Также встречается подход, при котором тренд является функцией некоторого вида, отражающей закономерности в поведении рассматриваемого временного ряда (речь идет о закономерностях, характеризующих поведение ряда для всего рассматриваемого периода, а не краткосрочные особенности). В работе рассматривается разложение STL, построение моделей ARIMA, использование моделей ACD (усредненного условного смещения) и другие подходы. Хотя разложение на тренд, сезонность, остаток и является общеупотребительной практикой, многие комбинации, представленные в вычислительном эксперименте, построены впервые (например, использование ряда Фурье для моделирования тренда, совмещение модели сезонности и модели тренда на основе алгоритма ACD). Во второй части работы представлен вычислительный эксперимент, в котором модели, использующие различные подходы к понятию тренда, его выделению и обработке, сравниваются по значению функции максимального правдоподобия и по прогнозу на тестовый период для динамических рядов макроэкономической статистики РФ; цены акций Сбербанка РФ на Московской бирже временного периода 2000–2021 гг.

Результаты. Во всех экспериментах один из наиболее точных прогнозов сделан при помощи метода LOESS. Для сезонных рядов достаточно точные результаты показывает моделирование тренда на основе многочлена и сезонности на основе функций ARIMA, совмещение модели тренда на основе алгоритма ACD и сезонности на основе ETS и моделирование на основе ряда Фурье.

Выводы. Метод LOESS для групп сезонных и несезонных рядов дает наилучший результат по всем показателям, поэтому можно рекомендовать именно этот метод для получения наиболее точных результатов для рядов различной природы. Моделирование тренда с помощью разложения в ряд Фурье приводит к достаточно точным результатам на временных рядах различной природы. Для сезонных рядов один из лучших результатов дает комбинация моделирования тренда на основе многочлена и сезонности в виде модели ARIMA.

Ключевые слова: динамические ряды, макроэкономическая статистика, ARIMA, ACD, временные ряды, тренд, функция максимального правдоподобия, моделирование тренда

• Поступила: 21.06.2023 • Доработана: 08.02.2024 • Принята к опубликованию: 09.04.2024

Для цитирования: Мохнаткина У.С., Парфенов Д.В., Петрусевич Д.А. Анализ подходов к определению тренда в структуре временного ряда. *Russ. Technol. J.* 2024;12(3):93–103. <https://doi.org/10.32362/2500-316X-2024-12-3-93-103>

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Analysis of approaches to identification of trend in the structure of the time series

Ulyana S. Mokhnatkina,
Denis V. Parfenov,
Denis A. Petrushevich @

MIREA – Russian Technological University, Moscow, 119454 Russia

@ *Corresponding author, e-mail: petrushevich@mirea.ru, petrdenis@mail.ru*

Abstract

Objectives. The study set out to compare the forecasting quality of time series models that describe the trend in different ways and to form a conclusion about the applicability of each approach in describing the trend depending on the properties of the time series.

Methods. A trend can be thought of as the tendency of a given quantity to increase or decrease over the long term. There is also an approach in which a trend is viewed as some function, reflecting patterns in the behavior of the time series. In this case, we discuss the patterns that characterize the behavior of the series for the entire period under consideration, rather than short-term features. The experimental part involves STL decomposition, construction of ARIMA models (one of the stages of preparation for which includes differentiation, i.e., removal of the trend and transition to a weakly stationary series), construction of ACD models (average conditional displacement) and other approaches. Time-series models based on various trend models are compared with respect to the value of the maximum likelihood function. Many of the combinations have not been constructed before (Fourier series as a trend model, combination of ACD model for trend with seasonal models). Example forecasts of macroeconomic statistics of the Russian Federation and stock prices of Sberbank on the Moscow Exchange in the time range of 2000–2021 are presented.

Results. In the experiments, The LOESS method obtained the best results. A combination of polynomial model for trend description and ARIMA for seasonally description and combination of ACD algorithm for trend and ETS for seasonal model obtained good forecasts in case of seasonal time series, while Fourier time series as a trend model also achieved close quality of prediction.

Conclusions. Since the LOESS method for groups of seasonal and non-seasonal series gives the best results for all indicators, this method can be recommended for obtaining the most accurate results for series of different nature. Trend modeling using Fourier series decomposition leads to quite accurate results for time series of different natures. For seasonal series, one of the best results is given by the combination of modeling a trend on the basis of a polynomial and seasonality in the form of the ARIMA model.

Keywords: dynamic series, macroeconomic statistics, ARIMA, ACD, time series, trend, maximum likelihood function, trend modeling

• Submitted: 21.06.2023 • Revised: 08.02.2024 • Accepted: 09.04.2024

For citation: Mokhnatkina U.S., Parfenov D.V., Petrusevich D.A. Analysis of approaches to identification of trend in the structure of the timeseries. *Russ. Technol. J.* 2024; 12(3):93–103. <https://doi.org/10.32362/2500-316X-2024-12-3-93-103>

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

В рамках представленной работы анализируются различные подходы к определению тренда временного ряда, существующие в современной науке. Следует отметить, что понятие тренда может быть задано с помощью нескольких определений, единого подхода к определению тренда не существует. Способы выделения тренда могут зависеть как от характеристик ряда, так и от выбранного исследователем подхода. Под трендом обычно понимают возрастание или убывание временного ряда на прогнозируемом участке. В медицине тренд, в основном, рассматривается как общее направление изменения среднего уровня характеристик в наборе данных [1]. Это определение можно обобщить до наличия постоянного однонаправленного изменения исследуемой величины [2–4]. Важно отметить наличие шумовой составляющей временного ряда, способной влиять на значения ряда как в меньшую, так и в большую сторону. По этой причине исследователей интересует компонента долгосрочного тренда, т.е. характеристики изменения исследуемой величины в течение длительного периода времени [5–8]. Также встречается способ выделения тренда в функциональном виде, что является анализом процесса на более глубоком уровне, где используется та же терминология.

В рамках работы рассмотрено несколько методов выделения тренда, приведены прогнозы на тестовый период моделей временных рядов (для случаев, где модель позволяет сделать прогноз), проведено сравнение различных подходов по качеству прогноза на тестовый период. В качестве обрабатываемых данных используются временные ряды макроэкономической статистики РФ¹, а также данные курса акций Сбербанка РФ на Московской бирже². Все рассматриваемые модели настраивались на периоде

обучения 2000–2020 гг. (из него удалены кризисные годы 2008 и 2014, данные склеены). Модели, исследуемые в работе, не зависят от конкретного периода времени, и выводы можно распространить и на другие временные процессы.

Исследователь обычно имеет под рукой множество моделей различной природы (на основе нейронных сетей, стандартные ARIMA³, ETS⁴, GARCH⁵ и т.п.), каждая из которых делает свой прогноз на целевой период. Модели, построенные в вычислительном эксперименте, могут использоваться (как показывает точность их прогнозов на тестовый период) в общем наборе предсказателей. Представленный подход актуален в связи с возможностью построить не применяемые ранее комбинации моделей для описания тренда и сезонной составляющей (если она есть) временного ряда. Это позволит добавить в существующий комплекс применяемых моделей и методов новые и объяснить, в каких случаях они могут лучше описывать динамику временного процесса по сравнению с другими моделями. Борьба за улучшение качества прогноза временного ряда остается важной вне зависимости от широты инструментария исследователя.

Цель работы состоит в построении различных моделей тренда, сопоставлении точности их прогнозов на тестовый период со стандартными моделями и построении выводов о возможности дальнейшего использования различных моделей для тренда при прогнозировании временных рядов.

В рамках концепции STL (англ., seasonal trend decomposition based on locally estimated scatterplot smoothing) [9] тренд T_t – это детерминированная часть временного ряда y_t , который может содержать также сезонную составляющую S_t и шумовую составляющую R_t . Ряд можно представить в аддитивной (1) или мультипликативной (2) форме:

¹ Единый архив экономических и социологических данных. Динамические ряды макроэкономической статистики РФ. Индекс денежных доходов населения; реального объема сельскохозяйственного производства. <http://sophist.hse.ru/hse/nindex.shtml>. Дата обращения 03.11.2023. [Unified archive of economic and sociological data. Dynamic series of macroeconomic statistics of the Russian Federation. Index of money incomes of the population; real volume of agricultural production. <http://sophist.hse.ru/hse/nindex.shtml> (in Russ.). Accessed November 03, 2023.]

² Курс акций ПАО Сбербанк (SBER). <https://www.moex.com/ru/issue.aspx?board=TQBR&code=SBER>. Дата обращения 03.11.2023. [Sberbank (SBER) stock price. <https://www.moex.com/ru/issue.aspx?board=TQBR&code=SBER> (in Russ.). Accessed November 03, 2023.]

³ Autoregressive integrated moving average model – модель авторегрессии и скользящего среднего или модель Бокса – Дженкинса.

⁴ Exponential time smoothing – экспоненциальное сглаживание.

⁵ Generalized autoregressive conditional heteroscedastic – авторегрессионная условная гетероскедастичность.

$$y_t = S_t + T_t + R_t, \quad (1)$$

$$y_t = S_t \times T_t \times R_t, \quad (2)$$

где t – время.

Подходы (1) и (2) эквивалентны, если возможно произвести замену переменных и преобразование единиц измерения на основе логарифмирования [9]:

$$\ln y_t = \ln(S_t \times T_t \times R_t) = \ln S_t + \ln T_t + \ln R_t.$$

Для выяснения наличия тренда можно использовать подход на основе статистических гипотез, в частности, критерий Манна – Кендала [10, 11].

Методы оценки тренда временного ряда делятся на параметрические и непараметрические. В параметрических методах под трендом понимают функцию одной или нескольких переменных. Примерами параметрических методов являются способы оценки тренда в виде функции, где параметры, характеризующие тренд, можно вычислять с помощью метода наименьших квадратов, полиномиальной подгонки, логарифмической, степенной, экспоненциальной, гармонической, логистической, кусочно-линейной функций, модели авторегрессии (модель AR⁶) и т.д. К непараметрическим методам можно отнести модель скользящего среднего (модель MA⁷), метод медианного сглаживания, метод экспоненциального сглаживания (модель ETS), применение частотных фильтров и т.д. В экспериментальной части работы рассматривается подбор логарифмической, линейной и экспоненциальной функций для описания тренда, а также применение байесовского подхода [12]. Они описывают монотонные участки роста и падения исследуемой величины, поэтому данные методы возможно применять только для монотонного временного ряда. В случае, когда исходный ряд не является монотонным, требуется произвести предварительную обработку, разбив ряд на монотонные части, и описывать каждую часть отдельно [13–17].

В экспериментальной части работы также исследуется метод описания тренда с помощью полиномов второго порядка и выше. Функция многочлена в общем случае не является монотонной, что потенциально приводит к вычислительным проблемам, а также встает вопрос выбора конкретного вида функции (степени многочлена). Данная проблема свойственна также сплайн-интерполяции. Кроме прочего, при полиномиальной интерполяции функция может сильно отклоняться от фиксированных значений в узлах. Например, известна проблема функции $1/(1+x^2)$ [18].

Рассмотрим алгоритм усредненного условного смещения ACD (average conditional displacement) [15] для оценки тренда и решения задач на его основе [12, 13]. По сравнению со сплайн-интерполяцией [19, 20] для описания поведения некоторой функции на заданном отрезке алгоритм ACD и родственные ему обладают полезным свойством сохранения монотонности. Схожие идеи также находят применение в работах на основе других подходов [14].

Кроме указанных выше способов описания тренда, широко применяются математические модели временных рядов, построенные на основе нейронных сетей (LSTM⁸, GRU⁹ и т.д.) [21–25]. Часто такие модели лучше описывают долгосрочные закономерности в данных по сравнению с моделями ARIMA, основанными на статистике [24, 26]. Известны работы, где модели разных типов интегрируются в единый механизм [22, 23]. Также в настоящее время исследователи делают попытки анализировать с помощью нейронных сетей характеристики и особенности временного ряда [26].

В экспериментальной части работы представлены построение моделей тренда на основе различных методов, объединение моделей с информацией о сезонности процесса, проверка качества прогнозов и соответствующие выводы о способности полученной модели настраиваться на значения временного ряда и о качестве прогноза на тестовый период.

РАССМАТРИВАЕМЫЕ МЕТОДЫ ВЫДЕЛЕНИЯ ТРЕНДА

Одна из самых распространенных моделей для описания временного ряда – ARIMA(p, d, q) [9], состоящая из авторегрессионной части AR (для модели порядка p значения ряда X ставятся в зависимость от p своих предыдущих значений):

$$X_t = c + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p},$$

где $\varphi_i, i = \overline{1, p}$ – коэффициенты функции; и из части скользящего среднего MA порядка q [9]:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

где $\theta_j, j = \overline{1, q}$ – коэффициенты функции. Порядок d обозначает число дифференцирований ряда.

Важно отметить, что при построении модели от тренда избавляются, переходя к стационарной временной разности (многократное дифференцирование ряда, пока не выполнится тест

⁶ Autoregressive model.

⁷ Moving average.

⁸ Long-short term memory – долгосрочная короткая память.

⁹ Gated recurrent unit – управляемый рекуррентный блок.

на стационарность) [9, 27]. Таким образом, описание модели и процесс прогнозирования производятся с преобразованным стационарным временным рядом. Модель ARIMA остается одной из самых распространенных в области прогнозирования временных рядов. В вычислительной части исследования ее результаты сравниваются с прогнозами других моделей.

В алгоритме усредненного условного смещения [15] для оценки тренда и решения задач на его основе [16, 17] участки временного ряда приближаются монотонными функциями вида:

$$f(x) = y_l + k(x - x_l), \quad (3)$$

где y_l – значение функции $f(x)$ в крайнем левом положении сегмента при $x = x_l$, k – коэффициент наклона прямой $f(x)$ (рисунок). Так как оценка тренда строится с помощью последовательного вычисления монотонных сегментов, то не возникает упомянутых проблем, как в случае полиномиальной интерполяции или сплайн-интерполяции.

В рамках вычислительного эксперимента качество прогнозов моделей ARIMA и ACD сравнивается с результатами моделей на основе нейронных сетей (LSTM, GRU) и моделей с применением беггинга [28] к данным временных рядов. Беггинг подразумевает формирование множества псевдовыборок на основе данных ряда. Итоговый прогноз получается с помощью усреднения или взвешенного усреднения прогнозов на тестовый период, построенных для каждой псевдовыборки [28].

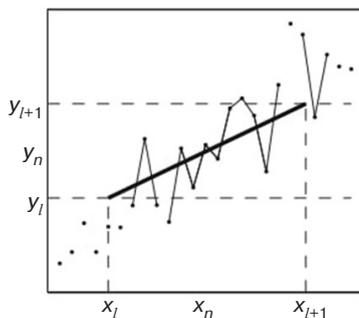


Рисунок. Схема построения монотонного сегмента по отрезку временного ряда согласно алгоритму ACD [15]

Для оценки прогнозов моделей рассматриваются меры близости вектора прогноза и вектора реальных значений временного ряда [9]:

$$\text{RMSE} = \sqrt{\frac{\sum (\tau(t) - ts(t))^2}{N}}, \quad (4)$$

$$\text{MAE} = \frac{\sum |\tau(t) - ts(t)|}{N}.$$

Здесь RMSE (root mean square error) – среднеквадратичная ошибка; MAE (mean absolute error) – средняя абсолютная ошибка; $\tau(t)$ – реальные значения временного ряда; $ts(t)$ – прогноз математической модели; N – длина отрезка прогнозирования (часто совпадает с длиной интервала сезонности). В основном, речь идет о временных рядах с годичной сезонностью, поэтому $N = 12$ месяцев).

ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

В эксперименте строятся и сравниваются модели временного ряда. Для моделирования используются ряды: денежных доходов населения, реального объема сельскохозяйственного производства по данным макроэкономической статистики РФ (ежемесячные показатели, безразмерные) и акции Сбербанка на московской бирже (ежемесячные показатели, рубли). Последний год в данных используется как тестовый период, на который все модели делают прогноз. Оставшаяся часть данных используется для обучения и настройки моделей. Модели временных рядов, которые участвуют в эксперименте, используют только данные временного ряда для настройки, не учитывая внешние факторы, поэтому перед настройкой моделей мы исключили данные вокруг кризисов 2008 и 2014 гг. Данные предыдущего и следующего периодов относительно кризисного года склеиваются. Графики ряда денежных доходов, его автокорреляционной (autocorrelation function, ACF) и частичной автокорреляционной функций (partial autocorrelation function, PACF) [9] представлены и подробно описаны в [29, рис. 3 и 4]. Графики для ряда реального объема сельскохозяйственного производства представлены и подробно описаны в [29, рис. 5 и 6].

При сравнении моделей используется функция максимального правдоподобия, оценки MAE/RMSE, которые показывают близость прогноза к реальным данным тестового периода. Результаты обработки индекса денежных доходов населения представлены в табл. 1 (жирным шрифтом отмечены лучшие модели по различным критериям).

Первые пять моделей тренда представляют собой многочлены, коэффициенты которых оцениваются с помощью метода наименьших квадратов. На прогноз накладывается модель сезонности. Следующая пятерка моделей отличается тем, что прогноз по построенной полиномиальной модели тренда делается с помощью ARIMA.

При прогнозировании с помощью гиперболической/показательной/логарифмической функции тренд моделируется с помощью соответствующей функции. На основе полученной регрессионной модели тренд прогнозируется на тестовый период с последующим наложением информации о сезонности.

Таблица 1. Модели денежных доходов населения по данным макроэкономической статистики РФ и их прогнозы на тестовый период

Оценка тренда	LLF	MAE	RMSE
Многочлен 1 степени $y = a_0 + a_1x$	-364.650	0.192	0.193
Многочлен 2 степени $y = a_0 + a_1x + a_2x^2$	-363.084	0.042	0.046
Многочлен 3 степени $y = a_0 + \sum_{i=1}^3 a_i x^i$	-362.139	0.055	0.062
Многочлен 4 степени $y = a_0 + \sum_{i=1}^4 a_i x^i$	-362.746	0.016	0.020
Многочлен 5 степени $y = a_0 + \sum_{i=1}^5 a_i x^i$	-363.487	0.081	0.086
Многочлен 1 степени + ARIMA	-364.151	0.145	0.147
Многочлен 2 степени + ARIMA	-362.863	0.024	0.027
Многочлен 3 степени + ARIMA	-362.453	0.026	0.033
Многочлен 4 степени + ARIMA	-362.668	0.014	0.019
Многочлен 5 степени + ARIMA	-362.746	0.016	0.020
Гиперболическая функция $y = a_0 + a_{-1} / x$	-361.679	0.105	0.107
Логарифмическая функция $y = a_0 + a_1 \ln x$	-363.174	0.050	0.054
Экспоненциальная функция $y = \exp(a_0 + a_1 x)$	-364.967	0.221	0.222
Интерполяция сплайнами	-362.688	–	–
Разложение в ряд Фурье	-362.783	0.020	0.027
Экспоненциальное сглаживание	-363.073	0.040	0.044
Метод LOESS	-362.719	0.015	0.019
Алгоритм ACD	-362.959	0.030	0.034
Тренд ACD + ARIMA	-362.198	0.060	0.074
Тренд ACD + ETS	-361.777	0.010	0.012

Коэффициенты модели также вычисляются с помощью метода наименьших квадратов.

Разложение в ряд Фурье используется для описания модели тренда на обучающий период. Прогнозирование осуществляется с помощью ARIMA аналогично представленным выше методам с наложением модели сезонности.

Интерполяцию функции, заданной таблично (что, в целом, соответствует измерениям, по которым строится временной ряд), можно провести с помощью сплайнов. Интересно, что такой метод интерполяции обладает высокой точностью, но на его основе не удастся сделать прогноз. При этом можно сравнить то, насколько сплайны и другие модели подстроены под данные ряда.

При использовании модели STL (метод LOESS¹⁰) [30, 31] ряд делится на составляющие: тренд, сезонность, шум.

При применении алгоритма ACD подход аналогичен. Модель тренда строится на основе ACD, производится прогноз на тестовый период.

Также исследуются гибридные модели, в которых на модель тренда накладывается сезонность и делается прогноз с помощью ETS или ARIMA.

Вычислительный эксперимент следует методике, используемой в работе [29]. В первой части используется временной ряд ежемесячных показателей

¹⁰ Locally estimated scatterplot smoothing – метод локального сглаживания диаграммы рассеивания.

денежных доходов населения. Ряд имеет годовую сезонность. Тестовым периодом является 2018 год. Результаты приведены в табл. 1 (точность – 0,01). Перед проведением эксперимента данные ряда были преобразованы в диапазон [0, 1]. Лучшие модели по точности прогноза на тестовый период или по значению логарифма функции максимального правдоподобия (LLF, log-likelihood function) выделены жирным шрифтом.

Самые точные результаты (столбцы MAE и RMSE) получаются при рассмотрении гибридной модели с оценкой тренда алгоритмом ACD с информацией о сезонности временного ряда и ее прогнозом на тестовый период с помощью ETS. Хорошие показатели по точности прогноза на тестовый период у метода оценки тренда с помощью многочлена четвертой степени с добавлением информации о сезонности и случайной составляющей на основе модели ARIMA(0, 4, 1). Значения функции правдоподобия меньше по сравнению с предыдущим методом. Такие же значения ошибок и качества модели для оценки тренда с помощью метода LOESS. Данные модели имеют почти те же показатели точности подстройки под данные ряда (LLF), что и при моделировании ряда на основе сплайнов.

Следующий по показателям точности результатов и качеству подстройки под исходный ряд – способ оценки тренда с помощью многочлена четвертой степени, смоделированного и спрогнозированного методом линейной регрессии. Такие же показатели имеет метод оценки тренда с помощью многочлена пятой степени с прогнозом на тестовый период методом ARIMA(3, 5, 1).

Неплохие оценки по точности прогноза и по функции правдоподобия при моделировании тренда с помощью разложения в ряд Фурье, многочленов второй и третьей степени с прогнозированием с помощью метода ARIMA. Чуть ниже точность прогноза при моделировании тренда с помощью алгоритма ACD.

Классические методы ARIMA и ETS имеют менее точные показатели прогноза на тестовый период, но далеко не худшие.

Наиболее низкие значения для логарифма функции максимального правдоподобия имеют методы оценки тренда с помощью экспоненциальной функции, а также с помощью многочлена первой степени. Однако данные модели имеют худшие результаты точности прогноза на тестовый период. Худший по всем показателям – метод моделирования тренда с помощью гиперболической функции.

Аналогичный вычислительный эксперимент проведен для временного ряда ежемесячных показателей индекса реального объема сельскохозяйственного производства (подробный анализ характеристик ряда приведен в работе [29]). Данные

2000–2020 гг. используются для обучения. Ряд имеет годовую сезонность. Тестовым периодом является 2021 год. Результаты приведены в табл. 2. Лучшие модели по точности прогноза на тестовый период или по значению логарифма функции максимального правдоподобия выделены жирным шрифтом.

Самые низкие значения для логарифма функции максимального правдоподобия имеет гибридная модель с оценкой тренда алгоритмом ACD с информацией о сезонности временного ряда и ее прогнозом на тестовый период с помощью ETS. Однако данный метод имеет не самые точные результаты прогнозов на тестовый период (столбцы MAE и RMSE).

Лучшие результаты по показателю MAE для прогноза на тестовый период у метода оценки тренда с помощью логарифмической и экспоненциальной функции (лучшие модели выделены жирным шрифтом). Близкие к ним значения у моделей с разложением в ряд Фурье, экспоненциальным сглаживанием, гиперболической функцией, а также с методом LOESS.

Лучшие результаты по показателю RMSE для прогноза на тестовый период у метода оценки тренда с помощью разложения в ряд Фурье. Также неплохие значения ошибок у моделирования тренда с помощью экспоненциальной функции, экспоненциального сглаживания, метода LOESS и с помощью многочлена первой степени с прогнозированием как на основе ARIMA, так и на основе линейной регрессии.

Наименьшее значение функции правдоподобия после гибридной модели с трендом ACD и прогнозом ETS имеет способ моделирования тренда с помощью многочлена пятой степени с прогнозированием линейной регрессией. Немного больше данное значение для гибридной модели с трендом ACD и информацией о сезонности временного ряда, спрогнозированной ARIMA. Также можно отметить модели с оценкой тренда с помощью многочленов третьей и четвертой степени. Однако данная группа моделей имеет одни из худших показателей для точности прогноза.

Важно отметить, что при оценке тренда с помощью многочленов точность снижается с ростом степени многочлена. Данное свойство сохраняется и при объединении оценки тренда с помощью многочленов с моделью ARIMA.

В представленных выше примерах обрабатываются сезонные временные ряды. Рассмотрим работу различных методов оценки тренда на несезонном временном ряду акций Сбербанка РФ¹¹ (руб.).

¹¹ Курс акций ПАО Сбербанк (SBER). <https://www.moex.com/ru/issue.aspx?board=TQBR&code=SBER>. Дата обращения 03.11.2023. [Sberbank (SBER) stock price. <https://www.moex.com/ru/issue.aspx?board=TQBR&code=SBER> (in Russ.). Accessed November 03, 2023.].

Таблица 2. Модели индекса реального объема сельскохозяйственного производства по данным макроэкономической статистики РФ и их прогнозы на тестовый период

Оценка тренда	LLF	MAE	RMSE
Многочлен 1 степени $y = a_0 + a_1x$	-395.597	0.084	0.097
Многочлен 2 степени $y = a_0 + a_1x + a_2x^2$	-395.871	0.096	0.105
Многочлен 3 степени $y = a_0 + \sum_{i=1}^3 a_i x^i$	-396.132	0.112	0.117
Многочлен 4 степени $y = a_0 + \sum_{i=1}^4 a_i x^i$	-396.173	0.114	0.119
Многочлен 5 степени $y = a_0 + \sum_{i=1}^5 a_i x^i$	-397.175	0.171	0.185
Многочлен 1 степени + ARIMA	-395.567	0.083	0.096
Многочлен 2 степени + ARIMA	-395.734	0.090	0.100
Многочлен 3 степени + ARIMA	-395.830	0.093	0.103
Многочлен 4 степени + ARIMA	-395.836	0.094	0.103
Многочлен 5 степени + ARIMA	-396.031	0.102	0.107
Гиперболическая функция $y = a_0 + a_{-1} / x$	-394.600	0.079	0.124
Логарифмическая функция $y = a_0 + a_1 \ln x$	-395.058	0.069	0.101
Экспоненциальная функция $y = \exp(a_0 + a_1 x)$	-395.310	0.075	0.095
Интерполяция сплайнами	-396.028	-	-
Преобразования Фурье	-395.648	0.076	0.089
Экспоненциальное сглаживание	-395.343	0.078	0.096
Метод LOESS	-395.379	0.079	0.097
Алгоритм ACD	-396.643	0.145	0.153
Тренд ACD + ARIMA	-397.140	0.170	0.195
Тренд ACD + ETS	-400.076	0.121	0.121

Данные 2000–2021 гг. используются для обучения. Тестовым периодом является 2022 год. Графики функций ACF и PACF представлены и подробно описаны в [29, рис. 7]. Результаты моделирования и прогнозирования представлены в табл. 3. Лучшие модели по точности прогноза на тестовый период или по значению логарифма функции максимального правдоподобия выделены жирным шрифтом.

Для биржевых данных сезонность отсутствует, поэтому результат прогноза зависит полностью от модели тренда и случайной составляющей.

Наилучшим образом поведение ряда описывают модели функции сплайнов и метода LOESS. При хороших показателях функции правдоподобия модель тренда с помощью метода LOESS имеет самые

точные прогнозы. Ошибка составляет 0.004 для MAE и 0.006 для RMSE. Также низкие значения LLF соответствуют моделированию тренда с помощью многочлена пятой степени с прогнозом линейной регрессии и с помощью разложения в ряд Фурье. Данные модели также имеют одни из лучших показателей точности после модели метода LOESS.

Остальные модели, в отличие от моделирования сезонных рядов, хуже подстраиваются под поведение ряда. Для них характерно ухудшение точности по мере возрастания значений LLF.

При моделировании тренда с помощью многочлена при возрастании степени многочлена, начиная с четвертой степени, точность растет, что отчасти может быть связано с переобучением. Это относится

Таблица 3. Модели временного ряда акций Сбербанка РФ и их прогнозы на тестовый период

Оценка тренда	LLF	MAE	RMSE
Многочлен 1 степени $y = a_0 + a_1x$	-4979.271	0.105	0.128
Многочлен 2 степени $y = a_0 + a_1x + a_2x^2$	-4950.327	0.199	0.215
Многочлен 3 степени $y = a_0 + \sum_{i=1}^3 a_i x^i$	-4881.779	0.440	0.465
Многочлен 4 степени $y = a_0 + \sum_{i=1}^4 a_i x^i$	-4986.211	0.108	0.114
Многочлен 5 степени $y = a_0 + \sum_{i=1}^5 a_i x^i$	-5019.458	0.079	0.119
Многочлен 1 степени + ARIMA	-4993.657	0.070	0.093
Многочлен 2 степени + ARIMA	-4995.075	0.067	0.089
Многочлен 3 степени + ARIMA	-4997.347	0.063	0.083
Многочлен 4 степени + ARIMA	-5006.537	0.067	0.085
Многочлен 5 степени + ARIMA	-5006.588	0.062	0.079
Гиперболическая функция $y = a_0 + a_{-1} / x$	-4880.421	0.443	0.451
Логарифмическая функция $y = a_0 + a_1 \ln x$	-4923.877	0.288	0.300
Экспоненциальная функция $y = \exp(a_0 + a_1 x)$	-4992.214	0.072	0.094
Интерполяция сплайнами	-5013.558	–	–
Преобразования Фурье	-5014.470	0.038	0.054
Экспоненциальное сглаживание	-4976.059	0.117	0.145
Метод LOESS	-5013.578	0.004	0.006
Алгоритм ACD	-4996.171	0.077	0.086
Тренд ACD + ARIMA	-4987.216	0.145	0.162
Тренд ACD + ETS	-4979.289	0.170	0.191

и к комбинации моделирования ряда с помощью многочлена и ARIMA, но показатели точности улучшаются при росте степени многочлена, начиная с первой степени.

ЗАКЛЮЧЕНИЕ

При моделировании тренда для сезонных рядов наложение на него информации о сезонной и случайной составляющих влияет на качество прогнозов. Лучшие результаты показывают способы моделирования тренда с помощью разложения в ряд Фурье и метод LOESS. Также достаточно точные результаты у комбинации моделирования тренда с помощью многочлена с методом ARIMA для сезонности.

При использовании многочленов для оценки тренда с прогнозом линейной регрессии показатели хуже, чем при использовании комбинации многочлена с моделью ARIMA. При этом динамика поведения показателей точности у них одинаковая.

Интересно заметить, что алгоритм ACD лучше всего проявил себя на данных денежных доходов населения. Этот временной ряд имеет неоднородную дисперсию. Прогноз данных с использованием алгоритма ACD может быть весьма полезен для гетероскедастических рядов.

Моделирование тренда с помощью экспоненциальной и логарифмической функций не продемонстрировало выдающихся результатов. Данные методы также сложнее с вычислительной точки зрения

по сравнению с моделями многочленов. Модель логарифмической функции имеет ограничения на значения данных, т.к. вещественного логарифма отрицательного аргумента не существует. Модель гиперболической функции является одной из худших как по показателям функции правдоподобия, так и по оценкам точности.

В отличие от оценки тренда с помощью экспоненциальной функции, выделение тренда с использованием экспоненциального сглаживания привело к одному из лучших результатов. Недостатком данного метода является неопределенный коэффициент сглаживания.

При работе с несезонными временными рядами качество прогноза зависит лишь от трендовой и шумовой составляющих. Лучше всего под данные подстраиваются модели с разложением в ряд Фурье,

методом LOESS и функцией сплайнов. Однако на основе сплайнов трудно сделать прогноз вследствие их ориентации на интерполяцию данных.

Следует отметить, что метод LOESS для группы несезонных рядов также дает наилучший или близкий к лучшему результат по всем показателям, поэтому можно рекомендовать именно этот метод получения наиболее точных результатов для рядов различной природы. Также можно отметить моделирование тренда с помощью разложения в ряд Фурье, т.к. данный подход также приводит к достаточно точным результатам на временных рядах различной природы.

Вклад авторов. Все авторы в равной степени внесли свой вклад в исследовательскую работу.

Authors' contribution. All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Allen R. Time series methods in the monitoring of intracranial pressure. Part 1: Problems, suggestion for a monitoring scheme and review of appropriate techniques. *J. Biomed. Eng.* 1983;5(1):5–18. [https://doi.org/10.1016/0141-5425\(83\)90073-0](https://doi.org/10.1016/0141-5425(83)90073-0)
2. Blom J.A., Ruyter J.F., Saranammi F., Beneken J.W. Detection of trends in monitored variables. In: Carson E.R., Cramp D.G. (Eds.). *Computer and Controls in Clinical Medicine*. New York: Plenum; 1985. P. 153–174. https://doi.org/10.1007/978-1-4613-2437-9_6
3. Challis R.E., Kitney R.I. Biomedical signal processing (in four parts). Part I: Time domain methods. *Med. Biol. Eng. Comput.* 1990;28(6):509–524. <https://doi.org/10.1007/bf02442601>
4. Haimowitz I.J., Kohane I.S. Automated trend detection with alternative temporal hypotheses. In: *Proceedings of the 13th International Joint Conference of Artificial Intelligence IJCAI-93*. 1993. P. 146–151.
5. Helsel D.R., Hirsch R.M., Ryberg K.R., Archfield S.A. *Statistical Methods in Water Resources*. USGS Science Publishing Network, Reston Publishing Service Center; 2018. 458 p. ISBN 978-1-4113-4348-1. <https://doi.org/10.3133/tm4a3>
6. Ding H., Li Z., Ren Q., Chen H., Song M., Wang Y. Single-variable method for predicting trends in chlorophyll a concentration based on the similarity of time series. *Ecological Indicators*. 2022;14096):109027. <https://doi.org/10.1016/j.ecolind.2022.109027>
7. Yao J., Wang P., Wang G., Shrestha S., Xue B., Sun W. Establishing a time series trend structure model to mine potential hydrological information from hydrometeorological time series data. *Sci. Total Environ.* 2020;698:134227. <https://doi.org/10.1016/j.scitotenv.2019.134227>
8. De Leo F., De Leo A., Besio G., Briganti R. Detection and quantification of trends in time series of significant wave heights: An application in the Mediterranean Sea. *Ocean Eng.* 2020;202:107155. <https://doi.org/10.1016/j.oceaneng.2020.107155>
9. Hyndman R.J., Athanasopoulos G. *Forecasting: Principles and Practice*. 3rd ed. OTexts; 2021. 442 p. ISBN-13 978-0-9875-0713-6
10. Mann H.B. Nonparametric tests against trend. *Econometrica*. 1945;13(3):2453–259. <https://doi.org/10.2307/1907187>
11. Kendall M.G. *Rank Correlation Methods*. 2nd ed. Hafner Publishing Co.; 1955. 196 p.
12. Kohns D., Bhattacharjee A. Nowcasting growth using Google Trends data: A Bayesian Structural Time Series model. *Int. J. Forecast.* 2022;39(3):1384–1412. <https://doi.org/10.1016/j.ijforecast.2022.05.002>
13. Yahyaoui H., Al-Daihani R. A novel trend based SAX reduction technique for time series. *Expert Systems with Applications*. 2019;130(C):113–123. <https://doi.org/10.1016/j.eswa.2019.04.026>
14. Xie Y., Liu S., Huang S., Fang H., Ding M., Huang C., Shen T. Local trend analysis method of hydrological time series based on piecewise linear representation and hypothesis test. *J. Clean. Prod.* 2022;339(1):130695. <https://doi.org/10.1016/j.jclepro.2022.130695>
15. Vamoş C., Crăciun M. *Automatic Trend Estimation*. Dordrecht, Heidelberg, New York, London: Springer; 2013. 131 p. <https://doi.org/10.1007/978-94-007-4825-5>
16. Feng Y., Zhou C. Forecasting financial market activity using a semiparametric fractionally integrated Log-ACD. *Int. J. Forecast.* 2015;31(2):349–363. <http://doi.org/10.1016/j.ijforecast.2014.09.001>
17. Allen D., Chan F., McAleer M., Peiris S. Finite sample properties of the QMLE for the Log-ACD model: Application to Australian stocks. *J. Econometrics*. 2008;147(1):163–185. <https://doi.org/10.1016/j.jeconom.2008.09.020>

18. Epperson J. On the Runge example. *The American Mathematical Monthly*. 1987;94(4):329–341. <https://doi.org/10.2307/2323093>
19. Drozdov I., Petrushevich D. Water pollution time series analysis. *IOP Conf. Ser.: Mater. Sci. Eng.* 2021;1047(1):012095. <http://doi.org/10.1088/1757-899X/1047/1/012095>
20. Petrushevich D. Review of missing values procession methods in time series data. *J. Phys.: Conf. Ser.* 2021;1889(3):032009. <http://doi.org/10.1088/1742-6596/1889/3/032009>
21. Wang P., Zheng X., Ai G., Liu D., Zhu B. Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos, Solitons & Fractals*. 2020;140:110214. <https://doi.org/10.1016/j.chaos.2020.110214>
22. Kumar B., Sunil P., Yadav N. A novel hybrid model combining β SARMA and LSTM for time series forecasting. *Appl. Soft Comput.* 2023;134:110019. <https://doi.org/10.1016/j.asoc.2023.110019>
23. Abebe M., Noh Y., Kang Y.-J., Seo C., Kim D., Seo J. Ship trajectory planning for collision avoidance using hybrid ARIMA-LSTM models. *Ocean Eng.* 2022;256:111527. <https://doi.org/10.1016/j.oceaneng.2022.111527>
24. Arunkumar K.E., Kalaga D.V., Kumar M.S., Kawaji M., Brenza T.M. Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. *Alexandria Eng. J.* 2022;61(10):7585–7603. <https://doi.org/10.1016/j.aej.2022.01.011>
25. Ning Y., Kazemi H., Tahmasebi P. A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. *Comput. Geosci.* 2022;164:105126. <https://doi.org/10.1016/j.cageo.2022.105126>
26. Anghinoni L., Zhao L., Ji D., Pan H. Time series trend detection and forecasting using complex network topology analysis. *Neural Netw.* 2019;117:295–306. <https://doi.org/10.1016/j.neunet.2019.05.018>
27. Box G., Jenkins G., Reinsel G.C. *Time Series Analysis: Forecasting and Control*. John Wiley and Sons; 2008. 784 p. ISBN-13 978-0470272848
28. Petropoulos F., Hyndman R.J., Bergmeir C. Exploring the sources of uncertainty: Why does bagging for time series forecasting work? *Eur. J. Oper. Res.* 2018;268(2):545–554. <https://doi.org/10.1016/j.ejor.2018.01.045>
29. Грамович Я.В., Мусатов Д.Ю., Петрусевич Д.А. Применение беггинга в прогнозировании временных рядов. *Russ. Technol. J.* 2024;12(1):101–110. <https://doi.org/10.32362/2500-316X-2024-12-1-101-110>
[Gramovich I.V., Musatov D.Yu., Petrushevich D.A. Implementation of bagging in time series forecasting. *Russ. Technol. J.* 2024;12(1):101–110. <https://doi.org/10.32362/2500-316X-2024-12-1-101-110>]
30. Zhao K., Wulder M.A., Hu T., Bright R., Wu Q., Qin H., Li Y., Toman E., Mallick B., Zhang X., Brown M. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote Sens. Environ.* 2019;232:111181. <https://doi.org/10.1016/j.rse.2019.04.034>
31. Li J., Li Z.-L., Wu H., You N. Trend, seasonality, and abrupt change detection method for land surface temperature time-series analysis: Evaluation and improvement. *Remote Sens. Environ.* 2022;280:113222. <https://doi.org/10.1016/j.rse.2022.113222>

Об авторах

Мохнаткина Ульяна Станиславовна, студент, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: atlantika@live.ru. <https://orcid.org/0009-0008-8756-7267>

Парфенов Денис Васильевич, к.т.н., доцент, доцент кафедры высшей математики, Институт искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: parfenov@mirea.ru. Scopus Author ID 57217119805, SPIN-код РИНЦ 7463-3220, <https://orcid.org/0009-0004-0905-3827>

Петрусевич Денис Андреевич, к.ф.-м.н., доцент кафедры высшей математики, Институт искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: petrushevich@mirea.ru, petrdenis@mail.ru. Scopus Author ID 55900513600, ResearcherID AAA-6661-2020, SPIN-код РИНЦ 7999-6345, <https://orcid.org/0000-0001-5325-6198>

About the authors

Ulyana S. Mokhnatkina, Student, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: atlantika@live.ru. <https://orcid.org/0009-0008-8756-7267>

Denis V. Parfenov, Cand. Sci. (Eng.), Associate Professor, Higher Mathematics Department, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: parfenov@mirea.ru. Scopus Author ID 57217119805, RSCI SPIN-code 7463-3220, <https://orcid.org/0009-0004-0905-3827>

Denis A. Petrushevich, Cand. Sci. (Phys.-Math.), Associate Professor, Higher Mathematics Department, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: petrushevich@mirea.ru, petrdenis@mail.ru. Scopus Author ID 55900513600, ResearcherID AAA-6661-2020, RSCI SPIN-code 7999-6345, <https://orcid.org/0000-0001-5325-6198>

Отпечатано в «МИРЭА – Российский технологический университет».
119454, РФ, г. Москва, пр-т Вернадского, д. 78.
Подписано в печать 31.05.2024 г.
Формат 60 × 90/8. Печать цифровая.
Уч.-изд. листов 13.00.
Тираж 100 экз. Заказ № 1192.

Подписку на печатную версию
Russian Technological Journal можно оформить
через ООО «Агентство «Книга-Сервис», www.akc.ru.
Подписной индекс: **79641**.

Printed in MIREA – Russian Technological University.
78, Vernadskogo pr., Moscow, 119454 Russian
Federation.
Signed to print May 31, 2024.
Format 60 × 90/8. Digital print.
C.p.l. 13.00.
100 copies. Order No. 1192.

Subscription to the *Russian Technological Journal* printed version can be made through
the *Kniga-Servis* Agency, www.akc.ru.
Subscription index: **79641**.

