

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



ISSN 2782-3210 (Print)
ISSN 2500-316X (Online)

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
2023, том 11, № 4

Russian Technological Journal
2023, Vol. 11, No. 4

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



Russian Technological Journal 2023, том 11, № 4

Дата опубликования 31 июля 2023 г.

Научно-технический рецензируемый журнал освещает вопросы комплексного развития радио-технических, 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 (ПТУ 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 2023, Vol. 11, No. 4

Publication date July 31, 2023.

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	хабилированный доктор наук, профессор, проректор Вильнюсского университета, Вильнюс, Литва. Scopus Author ID 7003687871, juras.banyas@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	Habilitation 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

- В.В. Беликов, И.А. Прокуронов*
7 Построение верификатора стойкости пароля с использованием классических методов машинного обучения и рекуррентной LSTM нейронной сети
- Е.И. Зайцев, Е.В. Нурматова*
16 О подходе к управлению знаниями и разработке мультиагентной системы представления и обработки знаний

Роботизированные комплексы и системы. Технологии дистанционного зондирования неразрушающего контроля

- М.Э. Беляков, С.А.К. Диане*
26 Алгоритмы визуального анализа внешней среды автономным мобильным роботом в задаче уборки территории

Современные радиотехнические и телекоммуникационные системы

- В.К. Битюков, А.И. Лавренов, Д.А. Малицкий*
36 Анализ пульсаций DC/DC-преобразователя, построенного по Zeta-топологии, с использованием его предельной непрерывной математической модели
- А.А. Парамонов, В.М. Нгуен, М.Т. Нгуен*
49 Многозадачная нейронная сеть в задаче распознавания вида QAM- и PSK-модуляции в условиях параметрической априорной неопределенности

Математическое моделирование

- В.В. Аристов, А.В. Строганов, А.Д. Ястребов*
59 Моделирование пространственного распространения волн пандемии COVID-19 в России на основе кинетико-переносного описания
- Д.А. Карпов, В.И. Струченков*
72 Комбинированные алгоритмы аппроксимации для интерактивного проектирования дорожных трасс в системах автоматизированного проектирования
- С.Е. Савотченко*
84 Модели волноводов, сочетающих градиентные и нелинейно-оптические слои
- В.Б. Федоров, С.Г. Харламов, А.И. Стариковский*
94 Восстановление смазанного фотографического изображения движущегося объекта, получаемого на пределе разрешающей способности

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

- Д.Х. Михайлиди, А.В. Рагуткин, Д.О. Скобелев, А.Б. Сухатерин*
105 Организация инжинирингового центра для импортозамещения в промышленности

Contents

Information systems. Computer sciences. Issues of information security

- Vladimir V. Belikov, Ivan A. Prokuronov*
7 Password strength verification based on machine learning algorithms and LSTM recurrent neural networks
- Evgeniy I. Zaytsev, Elena V. Nurmatova*
16 Approach to knowledge management and the development of a multi-agent knowledge representation and processing system

Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing

- Maksim E. Beliaikov, Sekou Abdel Kader Diane*
26 Algorithms for the visual analysis of an environment by an autonomous mobile robot for area cleanup

Modern radio engineering and telecommunication systems

- Vladimir K. Bityukov, Alexey I. Lavrenov, Daniil A. Malitskiy*
36 Analysis of the DC/DC Zeta topology converter ripples by applying its limiting continuous mathematical model
- Aleksei A. Paramonov, Van Minh Nguyen, Minh Tuong Nguyen*
49 Multi-task neural network for solving the problem of recognizing the type of QAM and PSK modulation under parametric a priori uncertainty

Mathematical modeling

- Vladimir V. Aristov, Andrey V. Stroganov, Andrey D. Yastrebov*
59 Modeling of spatial spread of COVID-19 pandemic waves in Russia using a kinetic-advection model
- Dmitry A. Karpov, Valery I. Struchenkov*
72 Combined approximation algorithms for interactive design of road routes in CAD
- Sergey E. Savotchenko*
84 Models of waveguides combining gradient and nonlinear optical layers
- Victor B. Fedorov, Sergey G. Kharlamov, Anatoly I. Starikovskiy*
94 Restoration of a blurred photographic image of a moving object obtained at the resolution limit

Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems

- Dmitry Kh. Mikhailidi, Alexander V. Ragutkin, Dmitry O. Skobelev, Alexey B. Sukhaterin*
105 Organization of an engineering center for industrial import substitution

УДК 004.89
<https://doi.org/10.32362/2500-316X-2023-11-4-7-15>



НАУЧНАЯ СТАТЬЯ

Построение верификатора стойкости пароля с использованием классических методов машинного обучения и рекуррентной LSTM нейронной сети

В.В. Беликов^{1, @},
И.А. Прокуронов²

¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия

² СФБ Лаборатория, Москва, 127083 Россия

@ Автор для переписки, e-mail: belikov_v@mirea.ru

Резюме

Цель. Аутентификация с использованием паролей является одним из наиболее распространенных способов проверки подлинности в компьютерных системах. Существующие атаки на пароли, включающие в себя, в т.ч. атаки перебора и атаки по словарю, требуют не только защиты учетных данных пользователя на этапе эксплуатации паролей, но и определения требований к паролю, позволяющих повысить стойкость пароля к атакам, минимизируя возможность их реализации злоумышленником. Важной задачей при этом становится разработка верификатора, осуществляющего проверку пароля на стойкость и позволяющего исключить задание пользователем паролей, подверженных взлому. Построение верификатора с использованием методов машинного обучения позволяет алгоритмам самим формулировать требования к сложности пароля в произвольно комплексной форме, отталкиваясь только от инцидентов, имеющих для каждой категории стойкости списков известных паролей.

Методы. Предложены алгоритмы машинного обучения с учителем: метод опорных векторов, случайный лес, бустинг, рекуррентная LSTM (long short-term memory) нейронная сеть. В эксперименте для предобработки данных применены метод простой индексации символов с последующей обработкой embedding-слоем и метод TF-IDF (term frequency-inverse document frequency). Для выбора гиперпараметров алгоритмов была использована кросс-валидация.

Результаты. Проведен анализ рекомендаций и требований к паролям в международных и отечественных стандартах и возможности их реализации в виде верификатора стойкости пароля в различных операционных системах. Приведены результаты эксперимента на существующем наборе помеченных по уровню стойкости паролей. Проведена их оценка с использованием masco f1-меры.

Выводы. Использование рекуррентной LSTM нейронной сети выделено как одно из наиболее перспективных направлений для построения верификатора стойкости пароля.

Ключевые слова: компьютерная безопасность, стойкость пароля, машинное обучение с учителем, рекуррентная нейронная сеть, LSTM

• Поступила: 11.12.2022 • Доработана: 01.02.2023 • Принята к опубликованию: 02.05.2023

Для цитирования: Беликов В.В., Прокуронов И.А. Построение верификатора стойкости пароля с использованием классических методов машинного обучения и рекуррентной LSTM нейронной сети. *Russ. Technol. J.* 2023;11(4):7–15. <https://doi.org/10.32362/2500-316X-2023-11-4-7-15>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Password strength verification based on machine learning algorithms and LSTM recurrent neural networks

Vladimir V. Belikov ^{1, @},
Ivan A. Prokuronov ²

¹ MIREA – Russian Technological University, Moscow, 119454 Russia

² SFB Laboratory, Moscow, 127083 Russia

@ Corresponding author, e-mail: belikov_v@mirea.ru

Abstract

Objectives. One of the most commonly used authentication methods in computer systems, password authentication is susceptible to various attacks including brute-force and dictionary attacks. This susceptibility requires not only the strict protection of user credentials, but also the definition of criteria for increasing a password's strength to minimize the possibility of its exploitation by an attacker. Thus, an important task is the development of a verifier for checking passwords for strength and prohibiting the user from setting passwords that are susceptible to cracking. The use of machine learning methods to construct a verifier involves algorithms for formulating requirements for password complexity based on lists of known passwords available for each strength category.

Methods. The proposed supervised machine learning algorithms comprise support vector machines, random forest, boosting, and long short-term memory (LSTM) recurrent neural network types. Embedding and term frequency-inverse document frequency (TF-IDF) methods are used for data preprocessing, while cross-validation is used for selecting hyperparameters.

Results. Password strength recommendations and requirements from international and Russian standards are described. The existing methods of password strength verification in various operating systems are analyzed. The experimental results based on existing datasets comprising passwords having an associated level of strength are presented.

Conclusions. A LSTM recurrent neural network is highlighted as one of the most promising areas for building a password strength verifier.

Keywords: cybersecurity, password strength, supervised machine learning, recurrent neural network, LSTM

• Submitted: 11.12.2022 • Revised: 01.02.2023 • Accepted: 02.05.2023

For citation: Belikov V.V., Prokuronov I.A. Password strength verification based on machine learning algorithms and LSTM recurrent neural networks. *Russ. Technol. J.* 2023;11(4):7–15. <https://doi.org/10.32362/2500-316X-2023-11-4-7-15>

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

Аутентификация с использованием паролей является одним из наиболее распространенных способов проверки подлинности в компьютерных системах [1]. Существующие атаки на пароли (в т.ч. атаки перебора, атаки по словарю и атаки с помощью радужных таблиц) требуют не только защиты учетных данных пользователя на этапе эксплуатации паролей, но и определения требований к паролю при его задании. Эти требования должны повышать стойкость пароля к указанным атакам, минимизируя возможность их реализации злоумышленником [2]. Важной задачей при этом становится разработка верификатора, осуществляющего проверку пароля на стойкость и позволяющего исключить задание пользователем паролей, подверженных взлому [3].

АНАЛИЗ СУЩЕСТВУЮЩИХ ПОДХОДОВ К ПОСТРОЕНИЮ ВЕРИФИКАТОРА СТОЙКОСТИ ПАРОЛЯ

Наиболее известные правила задания стойкого пароля представлены [4] в следующих документах:

1. Специальная публикация Национального института стандартов и технологий (NIST) «Рекомендации по цифровой идентификации. Аутентификация и управление жизненным циклом»¹.
2. Методический документ Федеральной службы по техническому и экспортному контролю (ФСТЭК) «Меры защиты информации в государственных информационных системах»².

Эти руководящие принципы содержат рекомендации для пользователей при создании паролей, а также требования и рекомендации для верификаторов (веб-сайты, программное обеспечение и т.д.), которые содержат в себе систему проверки и обработки паролей.

В первом документе пароли фигурируют под названием «запомненные секреты» (от англ. *memorized secrets*). Среди основных положений можно выделить следующие:

1. Запомненные секреты (пароли) должны быть длиной не менее 8 символов, если они были выбраны пользователем.
2. Запомненные секреты (пароли), случайно сгенерированные компьютером или верификатором,

должны иметь длину не менее 6 символов, а также могут состоять полностью из цифр.

3. Если компьютер или верификатор запрещает выбранный сохраненный секрет (пароль) на основании того, что он содержится в принятом ранее черном списке скомпрометированных значений, пользователь должен выбрать другой сохраненный секрет (пароль).

В документе ФСТЭК рекомендованная минимальная длина пароля составляет 6 символов для достижения требований четвертого (самого простого) уровня защищенности персональных данных [5].

При обработке запросов на установление и изменение сохраненных секретов верификаторы должны сравнивать предполагаемые секреты со списком часто используемых, ожидаемых или скомпрометированных паролей. Например, список может включать в себя:

1. Пароли, полученные из баз взломанных паролей.
2. Словарные слова.
3. Повторяющиеся или последовательные символы (например, «qqqqqq», «qwerty12345»).
4. Слова, зависящие от контекста, такие как название службы, имя пользователя и их производные (например, «mireastudent», «ivanivanov»).

Еще одним способом проверки пароля на его стойкость является использование специальных сервисов, таких как, например, сервис, предложенный на официальном сайте компании Kaspersky³. Этот сервис позволяет получить информацию о стойкости пароля, о встречаемости пароля в утекших базах, а также узнать сколько потребуется времени, чтобы взломать пароль с помощью атаки грубой силы (рис. 1).

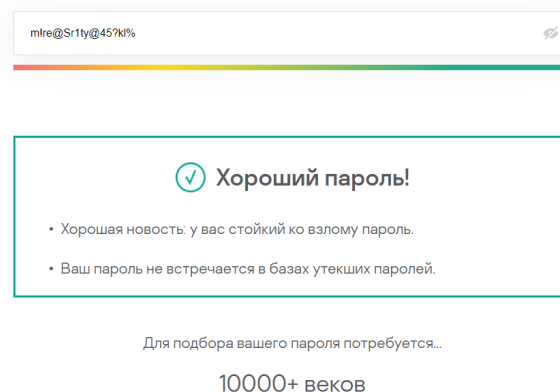


Рис. 1. Демонстрация проверки пароля на стойкость на сайте Kaspersky

¹ <https://pages.nist.gov/800-63-3/sp800-63b.html>. Дата обращения 01.02.2023. / Accessed February 01, 2023.

² <https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-11-fevralya-2014-g> (in Russ.). Дата обращения 01.02.2023. / Accessed February 01, 2023.

³ <https://password.kaspersky.com/ru/> (in Russ.). Дата обращения 01.02.2023. / Accessed February 01, 2023.

Сложность выбираемых пользователем паролей может быть охарактеризована с использованием концепции теории информации, основоположником которой считается американский математик Клод Шеннон. Хотя энтропию, как меру информационной емкости системы, можно легко вычислить для данных, имеющих детерминированные функции распределения, оценка энтропии для паролей, выбранных пользователем, является сложной задачей, и предыдущие попытки сделать это не привели к получению точного результата. По этой причине, например, в публикации NIST используется подход, основанный в первую очередь на длине пароля.

Для обеспечения проверки пароля на стойкость в различных операционных системах существуют специальные модули, которые оценивают пароль по каким-либо критериям, придерживаясь общепринятых рекомендаций, а также рекомендаций, которые могут быть установлены администратором системы. В общем случае это называется политикой паролей, которая представляет собой некий набор правил, повышающий безопасность учетных записей пользователей, поощряя их к использованию более надежных паролей.

Долгое время в операционных системах Linux использовался РАМ-модуль (РАМ – подключаемые модули аутентификации, pluggable authentication modules) *pam_cracklib*, который предназначался для проверки пароля по словарным словам [6]. В последних версиях Linux данный модуль был заменен на модуль *pam_pwquality*, основанный на модуле *pam_cracklib* и полностью обратно совместимый с его опциями. Этот модуль делает подход к созданию политик более простых паролей для проверки того, что пользователи принимают рекомендации и требования администратора по созданию пароля.

Примером политики паролей, которую можно задать с использованием указанного модуля, может быть следующий список требований:

1. Минимальная длина пароля – 10 символов.
2. В новом пароле должно присутствовать 6 новых символов, которых не было в старом пароле.
3. Пароль должен содержать строчные буквы, заглавные буквы, специальные символы, а также цифры.
4. В пароле не должно содержаться более двух подряд повторяющихся символов.
5. В пароле не более шести подряд идущих символов должны быть из одного класса.
6. Присутствие проверки GESOC.
7. Запретить слова «mirea», «security», «admin», «password», «cyber».

Содержание файла конфигурации модуля */etc/security/pwquality.conf*, отражающего описанную политику безопасности, приведено на рис. 2.

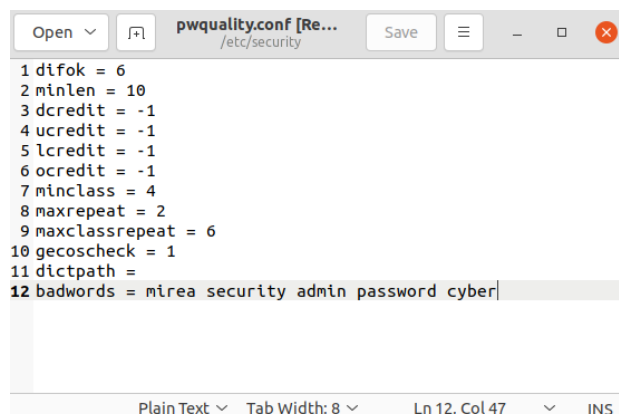


Рис. 2. Вид файла конфигурации *pwquality.conf*

Для проверки корректности заданной политики паролей можно воспользоваться программой *pwscore*, которая входит в пакет *libpwquality*. Пример тестирования пароля, содержащего запрещенное слово, и пароля, удовлетворяющего всем требованиям, с использованием утилиты *pwscore*, выставляющей баллы от 0 до 100, приведен на рис. 3.

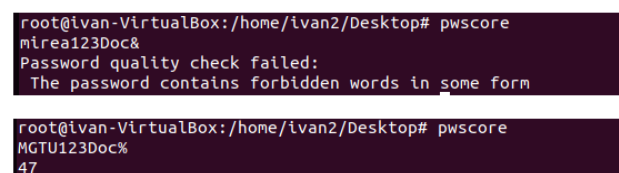


Рис. 3. Тестирование пароля

В операционных системах семейства Windows настройка политики паролей задается в разделе «Локальная политика безопасности» утилиты *gpedit.msc* [7] (рис. 4).

Имея в своем арсенале всего лишь восемь настраиваемых параметров, операционная система Windows позволяет строго устанавливать и менять политику паролей, включая или включая и комбинируя установленные значения этих параметров. Одним из самых важных при этом является пункт «Пароль должен отвечать требованиям сложности». Включение данного пункта определяет то, что задаваемые пароли должны удовлетворять следующим требованиям:

1. Пароль не должен содержать в себе имя учетной записи пользователя или частей полного имени пользователя длиной более двух подряд идущих символов.
2. Минимальная длина пароля 6 символов.
3. Пароль должен содержать символы, как минимум, из трех классов: латинские заглавные буквы, латинские строчные буквы, цифры, специальные символы. В Windows 11 был добавлен четвертый класс, который требует наличие любого символа Unicode.

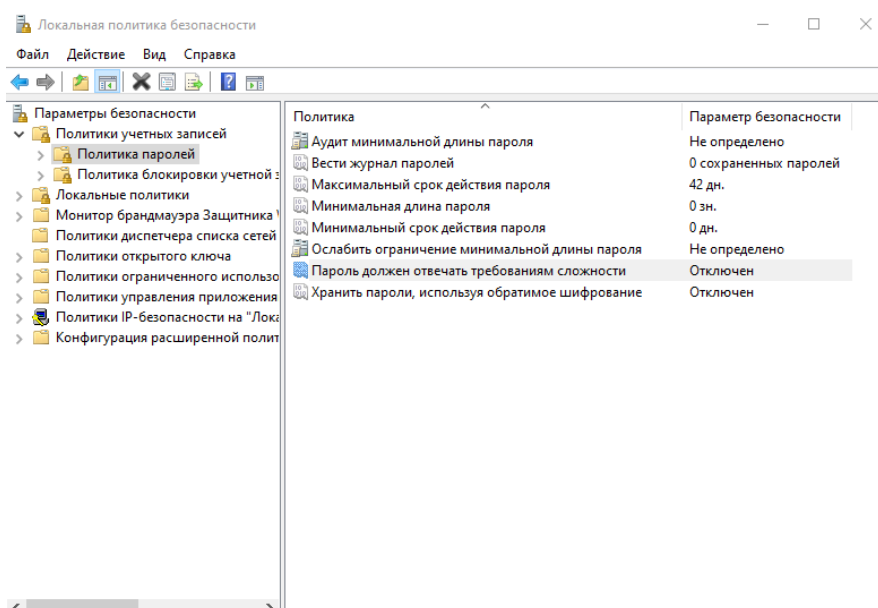


Рис. 4. Окно настройки политики паролей в Windows

Таким образом, анализ существующих подходов к построению верификатора стойкости паролей показывает, что наиболее распространенным решением является выделение правил, в которых акцент делается прежде всего на длину пароля и его наличие в имеющихся словарях паролей. Такое решение обладает рядом недостатков, в т.ч. сложностью реализации более комплексных требований, отражающих свойства высокой стойкости паролей. В отличие от этого, построение верификатора с использованием машинного обучения позволяет алгоритмам самим формулировать эти требования в произвольно комплексной форме, отталкиваясь только от инцидентов – имеющихся для каждой категории стойкости списков известных паролей. Одним из свойств указанного подхода является его универсальность, т.к. под каждый конкретный случай не требуется создавать и настраивать политику паролей (количество специальных символов, количество цифр и т.д.). Машина имитирует смесь многих алгоритмов проверки надежности пароля. Обученную машину можно использовать как отдельный модуль (например, ПАМ-модуль в Linux), который всегда будет точно определять стойкость паролей. Например, в случае если разработчик создает какую-либо новую социальную сеть, ему будет достаточно обучить машину на основании датасета, собранного из других популярных социальных сетей, где политика паролей считается эффективной: TikTok⁴, Twitter⁵ (запрещена

в Российской Федерации), Facebook⁶ (запрещена в Российской Федерации) и т.п. После этого машина встраивается в виде верификатора, который оценивает пароль и выдает соответствующую информацию пользователю при регистрации. В таком случае, если введенному паролю будет поставлен высший балл стойкости, пользователь может быть уверен в том, что его пароль действительно стойкий без привязки к конкретным требованиям и рекомендациям.

МЕТОДОЛОГИЯ И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА

Задача верификации стойкости пароля может быть представлена как задача классификации, где объектом данных является пароль, а классом – его уровень стойкости.

В качестве набора данных для эксперимента использовался датасет одной из самых крупных утечек паролей – хостинга 000webhost [8]. С помощью инструмента PARS [9], в котором находятся множество счетчиков, оценивающих стойкость пароля, независимым разработчиком была создана база, содержащая пароли и их оценку на основании трех разных алгоритмов, реализованных компаниями Twitter, Microsoft⁷, Battle.net⁸. Сам датасет содержит в себе лишь те пароли, которые были оценены всеми тремя

⁴ <https://www.tiktok.com/>. Дата обращения 01.02.2023. / Accessed February 01, 2023.

⁵ <https://twitter.com/>. Дата обращения 01.02.2023. / Accessed February 01, 2023.

⁶ <http://facebook.com/>. Дата обращения 01.02.2023. / Accessed February 01, 2023.

⁷ <https://www.microsoft.com/>. Дата обращения 01.02.2023. / Accessed February 01, 2023.

⁸ <https://www.battle.net>. Дата обращения 01.02.2023. / Accessed February 01, 2023.

измерителями одинаково. В наборе представлено три класса оценки пароля (0 – низкий, 1 – средний, 2 – высокий). Датасет содержит в себе несбалансированные данные: 496649 паролей с оценкой «1», 89662 паролей с оценкой «0», 83113 сильных паролей.

Примеры паролей для каждого уровня стойкости представлены в табл. 1. Распределения длины пароля в целом по набору данных и по отдельным уровням стойкости представлены на рис. 5 и 6.

Таблица 1. Примеры паролей из учебного набора данных

Стойкость пароля	Примеры паролей
0	wewes19 asdas95
1	oyeleye1 80188063JA
2	JFRTgxTQyNQTh9ZD d7a6AoTMxMw0dLVy

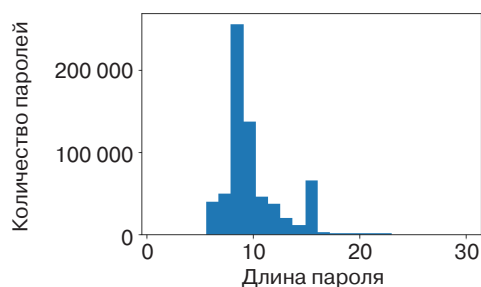


Рис. 5. Гистограмма длины пароля в наборе данных

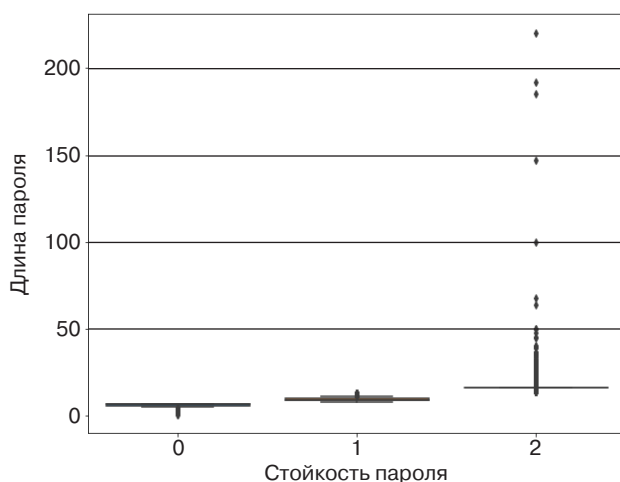


Рис. 6. Распределение длины пароля в зависимости от стойкости в наборе данных

Для проведения эксперимента набор после проведения случайного перемешивания разбивался на обучающий (80%) и тестовый (20%).

Для представления текстового пароля в виде вектора числовых значений в целях учета не только наличия, но и веса каждого отдельного символа,

использовался метод TF-IDF (term frequency-inverse document frequency) [10]:

$$TF = \frac{\text{число вхождений символа в пароль}}{\text{общее количество символов в пароле}},$$

$$IDF = \log_2 \frac{\text{общее количество паролей}}{\text{количество паролей, в которых встречается символ}},$$

$$TFIDF = TF \cdot IDF,$$

После применения к набору паролей метода TF-IDF был получен набор, где каждому паролю соответствует вектор длины, совпадающей с размером словаря, а каждому элементу вектора — вес TF-IDF-символа, порядковый номер которого в словаре соответствует порядковому номеру элемента в векторе.

Для обучения классических алгоритмов использовалось уравнивание количества экземпляров в каждом классе с использованием технологии undersampling, для выбора гиперпараметров — k -fold кросс-валидация [11] с $k = 5$. Полученные результаты обучения алгоритмов на обучающем наборе данных, оцененные с использованием макро $f1$ -меры [12], представлены в табл. 2.

Таблица 2. Результаты обучения, оцененные на обучающем наборе методом кросс-валидации

Используемый алгоритм	Найденное значение гиперпараметра	Значение макро $f1$ -меры
Метод опорных векторов	Коэффициент регуляризации: 1	0.806
Случайный лес	Максимальная глубина дерева: 32	0.903
Бустинг	Максимальная глубина дерева: 32	0.946

Вместе с тем, использование классических методов машинного обучения осложняется необходимостью преобразования паролей с использованием алгоритма TF-IDF, что является затратной процедурой и требует пересчета весов при обновлении набора паролей. Кроме этого, указанные алгоритмы представляют пароли в виде неупорядоченного набора символов (упрощение, используемое в подходе «Bag of Words»⁹) и не учитывают взаимное расположение

⁹ Упрощенное представление текста в виде мешка (мультимножества) его слов без какого-либо учета грамматики и порядка слов, но с сохранением информации об их количестве. [A simplified representation of the text as a bag (multiset) of its words without any regard to grammar or word order but with retention of information about their number.]

символов. Это частично может быть исправлено использованием моделей биграмм и триграмм, что в свою очередь значительно усложнит процесс обучения. В отличие от этого, рекуррентные нейронные сети работают напрямую с последовательностями символов произвольной длины [13]. Так, для рекуррентной нейронной сети пароли «PASSWORD» и «AWSSODPR» будут определяться двумя разными векторами, тогда как при использовании TF-IDF – одним. Поэтому для оценки стойкости пароля в эксперименте в качестве метода, альтернативного классическим алгоритмам машинного обучения, использовалась также реализованная с использованием библиотеки PyTorch [14] нейронная сеть, состоящая из следующих слоев:

1. Слой embedding, используемый для преобразования пароля, состоящего из символов, в вектор числовых значений.
2. Слой long short-term memory (LSTM) – особая разновидность архитектуры рекуррентных нейронных сетей, способных к обучению долгосрочным зависимостям, что важно при работе с паролями большой длины [15].
3. Линейный слой, используемый для преобразования внутреннего состояния LSTM-слоя в оценки принадлежности пароля к категориям стойкости (category scores).

Результаты сравнения работы алгоритма бустинга и рекуррентной LSTM-сети на тестовом наборе приведены в табл. 3.

Таблица 3. Результаты обучения, оцененные на тестовом наборе

Показатель	Бустинг	Рекуррентная LSTM-сеть
Верность	0.972	0.9995
Точность	0.971	0.9994
Полнота	0.971	0.9990
Макро f1-мера	0.971	0.9992

Исходный код эксперимента можно увидеть на сайте¹⁰.

Следует отметить, что используемый в эксперименте подход никак не снижает возможности хакера в случае применения фишинговых атак, атак кейлоггеров, а также атак «злоумышленник посередине». Однако, если обратиться к исследованию Data Breach Investigation Report от 2020 г. [16], то можно отметить, что 89% всех разновидностей взлома связаны с каким-либо видом злоупотребления учетными данными (атаки грубой силы и их подвиды,

а также атаки, нацеленные на повторное использование людьми учетных данных). Данный факт говорит о том, что подход к определению стойкости пароля с помощью машинного обучения минимизирует риски по большинству векторов возможных атак, что также объясняет значимое количество современных исследований по этому направлению [17–21].

Таким образом, можно сделать вывод о том, что рекуррентная LSTM-сеть опережает традиционные методы машинного обучения, приближая значения показателя качества классификации стойкости паролей к единице.

ЗАКЛЮЧЕНИЕ

В статье предложен подход к построению верификатора стойкости пароля с применением методов машинного обучения, проведено сравнение нескольких алгоритмов на наборе данных паролей, помеченных по уровню стойкости. Показано, что такой подход обладает рядом преимуществ по сравнению с классическими методами верификации стойкости пароля, работающими без применения технологий машинного обучения. Так, использование методов машинного обучения для построения верификатора позволяет формулировать требования к сложности пароля в произвольно комплексной форме, отталкиваясь только от инцидентов. Кроме этого, предложенный в статье подход позволяет лучше противостоять атакам, являющимся подвидом атак грубой силы, а также атакам с помощью радужных таблиц. В первом случае это достигается тем, что сложный пароль, стойкость которого оценена с помощью алгоритма машинного обучения, делает перебор пароля невозможным в связи с огромными временными затратами, которые потребуются злоумышленнику для выполнения этой задачи. Во втором случае атаковать сложный пароль будет возможно, только применив огромную радужную таблицу, которая, в свою очередь, потребует использования значительного объема ресурсов злоумышленника, что делает такую атаку нерелевантной.

Среди рассмотренных методов машинного обучения особенную эффективность продемонстрировали рекуррентные нейронные сети. Стоит отметить, что обучение представления текстового пароля в виде вектора числовых значений, заключающееся в нахождении весов слоя embedding, происходит одновременно с обучением всей сети для максимизации точности классификации. Это позволяет выбирать самой нейронной сети такую векторизацию пароля, которая эффективна именно для решаемой задачи. Кроме этого, нейронные сети работают с последовательностью символов, а не только с их наличием в пароле, что позволяет отказаться от упрощения, используемого в подходе «Bag of Words».

¹⁰ https://github.com/james116blue/password_strength_verifier. Дата обращения 01.02.2023. / Accessed February 01, 2023.

Вышесказанное позволяет выделить использование рекуррентной нейронной сети как одно из наиболее перспективных направлений исследований для построения верификатора стойкости пароля.

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

- Conklin A., Dietrich G., Walz D. Password-based authentication: a system perspective. In: *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*. 2004; IEEE. <https://doi.org/10.1109/HICSS.2004.1265412>
- Dell'Amico M., Michiardi P., Roudier Y. Password strength: An empirical analysis. In: *2010 Proceedings IEEE INFOCOM*. 2010; IEEE. <https://doi.org/10.1109/INFCOM.2010.5461951>
- Chakrabarti S., Singhal M. Password-based authentication: Preventing dictionary attacks. *Computer*. 2007;40(6): 68–74. <https://doi.org/10.1109/MC.2007.216>
- Shay R., Komanduri S., Kelley P.G., Leon P.G., Mazurek M.L., Bauer L., Christin N., Cranor L.F. Encountering stronger password requirements: user attitudes and behaviors. In: *Proceedings of the Sixth Symposium on Usable Privacy and Security*. 2010; Article 2. <https://doi.org/10.1145/1837110.1837113>
- Селифанов В.В. Оценка эффективности системы защиты информации государственных информационных систем от несанкционированного доступа. *Интеграция науки, общества, производства и промышленности: сборник статей Международной научно-практической конференции*. 2016. С. 109–113. [Selifanov V.V. Evaluation of the efficiency of the information protection system of state information systems from unauthorized access. In: *Integration of Science, Society, production and Industry: Collection of Articles of the International Scientific and Practical Conference*. 2016. P. 109–113 (in Russ.).]
- Ferreira J.F., Johnson S.A., Mendes A., Brooke P.J. Certified password quality: a case study using Coq and Linux pluggable authentication modules. In: *Integrated Formal Methods. IFM 2017. Lecture Notes in Computer Science*. V. 10510. Springer International Publishing; 2017. P. 407–421. https://doi.org/10.1007/978-3-319-66845-1_27
- Alshare K.A., Lane P.L., Lane M.R. Information security policy compliance: a higher education case study. *Information & Computer Security*. 2018;26(1):91–108. <https://doi.org/10.1108/ICS-09-2016-0073>
- AlSabah M., Oligeri G., Riley R. Your culture is in your password: An analysis of a demographically-diverse password dataset. *Computers & Security*. 2018;77: 427–441. <https://doi.org/10.1016/j.cose.2018.03.014>
- Ji S., Yang S., Wang T., Liu C., Lee W.H., Beyah R. Pars: A uniform and open-source password analysis and research system. In: *ACSAC' 15: Proceedings of the 31st Annual Computer Security Applications Conference*. 2015. P. 321–330. <https://doi.org/10.1145/2818000.2818018>
- Aizawa A. An information-theoretic perspective of TF-IDF measures. *Information Processing & Management*. 2003;39(1):45–65. [https://doi.org/10.1016/S0306-4573\(02\)00021-3](https://doi.org/10.1016/S0306-4573(02)00021-3)
- Bishop C.M. *Pattern Recognition and Machine Learning*. New York: Springer; 2006. 738 p.
- Lever J., Krzywinski M., Altman N. Classification evaluation: It is important to understand both what a classification metric expresses and what it hides. *Nat. Methods*. 2016;13(8):603–604. <https://doi.org/10.1038/nmeth.3945>
- Medsker L.R., Jain L.C. (Eds.). *Recurrent Neural Networks. Design and Applications*. CRC Press; 2001. P. 64–67.
- Imambi S., Prakash K.B., Kanagachidambaresan G.R. PyTorch. In: Prakash K.B., Kanagachidambaresan G.R. (Eds.). *Programming with TensorFlow*. EAI/Springer Innovations in Communication and Computing (book series). Springer; 2021. P. 87–104. https://doi.org/10.1007/978-3-030-57077-4_10
- Yu Y., Si X., Hu C., Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* 2019;31(7):1235–1270. https://doi.org/10.1162/neco_a_01199
- Jartelius M. The 2020 Data Breach Investigations Report—a CSO's perspective. *Network Security*. 2020;2020(7): 9–12. [https://doi.org/10.1016/S1353-4858\(20\)30079-9](https://doi.org/10.1016/S1353-4858(20)30079-9)
- Sarkar S., Nandan M. Password Strength Analysis and its Classification by Applying Machine Learning Based Techniques. In: *2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA)*. IEEE, 2022. P. 1–5. <https://doi.org/10.1109/ICCSEA54677.2022.9936117>
- Sakya S.S., Mauparna M.N. Building a Multi-class Password Strength Generator and Classifier Model by Augmenting Supervised Machine Learning Techniques. Preprint. 2022. <https://doi.org/10.21203/rs.3.rs-1820885/v1>
- Murmu S., Kasyap H., Tripathy S. PassMon: A Technique for Password Generation and Strength Estimation. *J. Network Syst. Manage.* 2022;30(1):13. <https://doi.org/10.1007/s10922-021-09620-w>
- Tran L., Nguyen T., Seo C., Kim H., Choi D. A Survey on Password Guessing. *arXiv preprint arXiv:2212.08796*. 2022. <https://doi.org/10.48550/arXiv.2212.08796>
- Xiao Y., Zeng J. Dynamically generate password policy via Zipf distribution. *IEEE Transactions on Information Forensics and Security*. 2022;17:835–848. <https://doi.org/10.1109/TIFS.2022.3152357>

Об авторах

Беликов Владимир Вячеславович, к.воен.н., доцент, доцент кафедры информационной безопасности № 252 Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: belikov_v@mirea.ru. Scopus Author ID 57983605100, <https://orcid.org/0000-0003-1423-1072>

Прокуронов Иван Андреевич, специалист инженерно-криптографического анализа, ООО «СФБ Лаборатория» (127083, Россия, Москва, ул. Мишина, д. 56, стр. 2). E-mail: miltumultik@gmail.com. <https://orcid.org/0000-0003-0999-311X>

About the authors

Vladimir V. Belikov, Cand. Sci. (Military), Assistant Professor, Department of Information Security, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: belikov_v@mirea.ru. Scopus Author ID 57983605100, <https://orcid.org/0000-0003-1423-1072>

Ivan A. Prokuronov, Cryptographic Analysis Specialist, SFB Laboratory (56/2, Mishina ul., Moscow, 127083 Russia). E-mail: miltumultik@gmail.com. <https://orcid.org/0000-0003-0999-311X>

УДК 004.89
<https://doi.org/10.32362/2500-316X-2023-11-4-16-25>



НАУЧНАЯ СТАТЬЯ

О подходе к управлению знаниями и разработке мультиагентной системы представления и обработки знаний

Е.И. Зайцев[®],
Е.В. Нурматова

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

Резюме

Цели. Мультиагентная система представления и обработки знаний (МСПОЗ) – это распределенная система искусственного интеллекта, предназначенная для решения проблем, которые трудно или невозможно решить с помощью монолитной интеллектуальной системы. Решение сложных проблем в МСПОЗ осуществляется интеллектуальными программными агентами, которые инкапсулируют в программных классах когнитивные структуры данных, методы логического вывода и машинного обучения. Интеллектуальные программные агенты МСПОЗ способны рационально действовать в условиях неполноты и нечеткости поступающей информации. Целями работы являются исследование и разработка моделей, методов, программных модулей и инструментальных программных средств, которые позволяют создать высокоэффективную МСПОЗ.

Методы. В работе использовались методы агентного моделирования, позволяющие формально описывать и программно имитировать рациональное поведение интеллектуальных агентов, методы экспертных оценок, математический аппарат теории автоматов, марковские цепи, нечеткая логика, нейронные сети, алгоритмы машинного обучения с подкреплением.

Результаты. Разработаны структурная схема МСПОЗ, мультиагентный решатель, схема управления доступом к микросервисам. Предложены методы распределения интеллектуальных программных агентов по узлам МСПОЗ, а также алгоритмы оптимизации логической структуры распределенной базы знаний (РБЗ), позволяющие повысить эффективность объемных, стоимостных и временных характеристик МСПОЗ.

Выводы. Предложен подход к разработке и использованию интеллектуальных программных агентов, который объединяет механизмы рассуждений на основе знаний с нейросетевыми моделями. Разработаны структура МСПОЗ, схема управления РБЗ, методы оптимизации РБЗ, определения доступности используемых агентами микросервисов, обеспечения надежности и скоординированного функционирования вычислительных узлов системы, а также инструментальные программные средства, позволяющие упростить процесс проектирования и реализации МСПОЗ. Полученные результаты демонстрируют эффективность представленного подхода к управлению знаниями и разработке высокопроизводительной проблемно-ориентированной МСПОЗ.

Ключевые слова: мультиагентная система, интеллектуальные программные агенты, мультиагентный интеллектуальный решатель, система представления и обработки знаний, обучение с подкреплением

• Поступила: 24.10.2022 • Доработана: 27.01.2023 • Принята к опубликованию: 02.05.2023

Для цитирования: Зайцев Е.И., Нурматова Е.В. О подходе к управлению знаниями и разработке мультиагентной системы представления и обработки знаний. *Russ. Technol. J.* 2023;11(4):16–25. <https://doi.org/10.32362/2500-316X-2023-11-4-16-25>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Approach to knowledge management and the development of a multi-agent knowledge representation and processing system

Evgeniy I. Zaytsev[@],
Elena V. Nurmatova

MIREA – Russian Technological University, Moscow, 119454 Russia

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

Abstract

Objectives. A multi-agent knowledge representation and processing system (MKRPS) comprises a distributed artificial intelligence system designed to solve problems that are difficult or impossible to solve using monolithic systems. Solving complex problems in an MKRPS is accomplished by communities of intelligent software agents that use cognitive data structures, logical inference, and machine learning. Intelligent software agents are able to act rationally under conditions of incompleteness and ambiguity of incoming information. The aim of the present work is to identify models and methods, as well as software modules and tools, for use in developing a highly efficient MKRPS.

Methods. Agent-based modeling methods were used to formally describe and programmatically simulate the rational behavior of intelligent agents, expert evaluation methods, the mathematical apparatus of automata theory, Markov chains, fuzzy logic, neural networks, and reinforcement learning.

Results. An MKRPS structure diagram, a multi-agent solver, and microservices access control diagram were developed. Methods for distribution of intelligent software agents on the MKRPS nodes are proposed along with algorithms for optimizing the logical structure of the distributed knowledge base (DKB) to improve the performance of the MKRPS in terms of volume, cost and time criteria.

Conclusions. The proposed approach to the development and use of intelligent software agents combines knowledge-based reasoning mechanisms with neural network models. The developed MKRPS structure and DKB control diagram includes described methods for optimizing the DKB, determining the availability of microservices used by the agents, ensuring the reliability assurance and coordinated functioning of the computing nodes of the system, as well as instrumental software tools to simplify the design and implementation of the MKRPS. The results demonstrate the effectiveness of the presented approach to knowledge management and the development of a high-performance problem-oriented MKRPS.

Keywords: multi-agent system, intelligent software agents, multi-agent intelligent solver, knowledge representation and processing system, reinforcement learning

• Submitted: 24.10.2022 • Revised: 27.01.2023 • Accepted: 02.05.2023

For citation: Zaytsev E.I., Nurmatova E.V. Approach to knowledge management and the development of a multi-agent knowledge representation and processing system. *Russ. Technol. J.* 2023;11(4):16–25. <https://doi.org/10.32362/2500-316X-2023-11-4-16-25>

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 мультиагентной системе представления и обработки знаний (МСПОЗ) решение сложных плохо формализуемых проблем осуществляется интеллектуальными программными агентами, которые способны рационально действовать в условиях неопределенности, неполноты и нечеткости поступающей информации [1–5]. Принимая решения и осуществляя рациональные действия, программные агенты используют базу знаний и событийно-управляемые микросервисы, спроектированные как отдельные взаимодействующие процессы с легковесными межпроцессными связями. Взаимодействие агентов с микросервисами осуществляется через события, в качестве которых могут выступать как простые уведомления, так и сложные структуры с поддержкой состояний.

Решение сложных проблем в МСПОЗ осуществляется путем декомпозиции проблем на подзадачи, которые совместно решают реактивные и когнитивные прикладные программные агенты. Используется как горизонтальная декомпозиция, приводящая к созданию многосвязной системы с плоской структурой, так и вертикальная декомпозиция, которая создает иерархическую систему с несколькими уровнями.

Благодаря реализации в МСПОЗ методов машинного обучения с подкреплением (reinforcement learning, RL) поведение прикладных программных агентов становится более рациональным при многократном решении задач. Для обучения прикладных программных агентов в МСПОЗ используется алгоритм Actor-Critic [6–10].

С целью повышения производительности в МСПОЗ реализованы специализированные прикладные программные интерфейсы (application programming interface, API) и системные программные модули, ассоциированные с системными программными агентами. Системные программные агенты осуществляют планирование и управление вычислительными ресурсами МСПОЗ, а также обеспечивают мобильность прикладных программных агентов. Прикладные агенты могут перемещаться по узлам МСПОЗ вместе с необходимым для них окружением. В отличие от контейнеров, реализуемых на основе пространств имен (например, платформой

Docker¹), для поддержки технологии мобильных агентов в МСПОЗ используются специализированные модули LibOS (library operating system).

Производительность МСПОЗ во многом определяется выбранным способом структуризации, хранения и обработки знаний [11, 12]. Был разработан высокопроизводительный проблемно-ориентированный мультиагентный решатель, и оптимизирована логическая структура распределенной базы знаний (РБЗ), имеющая минимальное общее время обработки запросов и транзакций.

СТРУКТУРА МСПОЗ

Структурная схема МСПОЗ изображена на рис. 1. На каждом вычислительном узле МСПОЗ функционируют прикладные (интеллектуальные) программные агенты двух типов: реактивные и когнитивные [13]. Прикладные программные агенты взаимодействуют между собой, а также с системными программными агентами, входящими в состав внешней разделяемой библиотеки пользовательского уровня LibOS, которая ориентирована на экзоядерную архитектуру операционной системы (ОС). Системные программные агенты используются для планирования и управления вычислительными ресурсами, распределения нагрузки и мониторинга системы.

Для представления агентно-ориентированных абстракций (целей, желаний, намерений, убеждений агентов) и реализации логического вывода используются специальные программные методы и когнитивные структуры данных (cognitive data structures, CDS), ассоциированные с когнитивными программными агентами.

На рис. 2 показан пример диаграммы состояний и переходов когнитивного программного агента одного из узлов МСПОЗ.

Как следует из диаграммы, когнитивный программный агент может находиться в одном из пяти состояний, два из которых являются составными (composite state), т.е. имеют вложенные состояния. Изменение состояния программного агента происходит в результате некоторого события. Возможен

¹ <https://www.docker.com/>. Дата обращения 20.03.2023. / Accessed March 20, 2023.

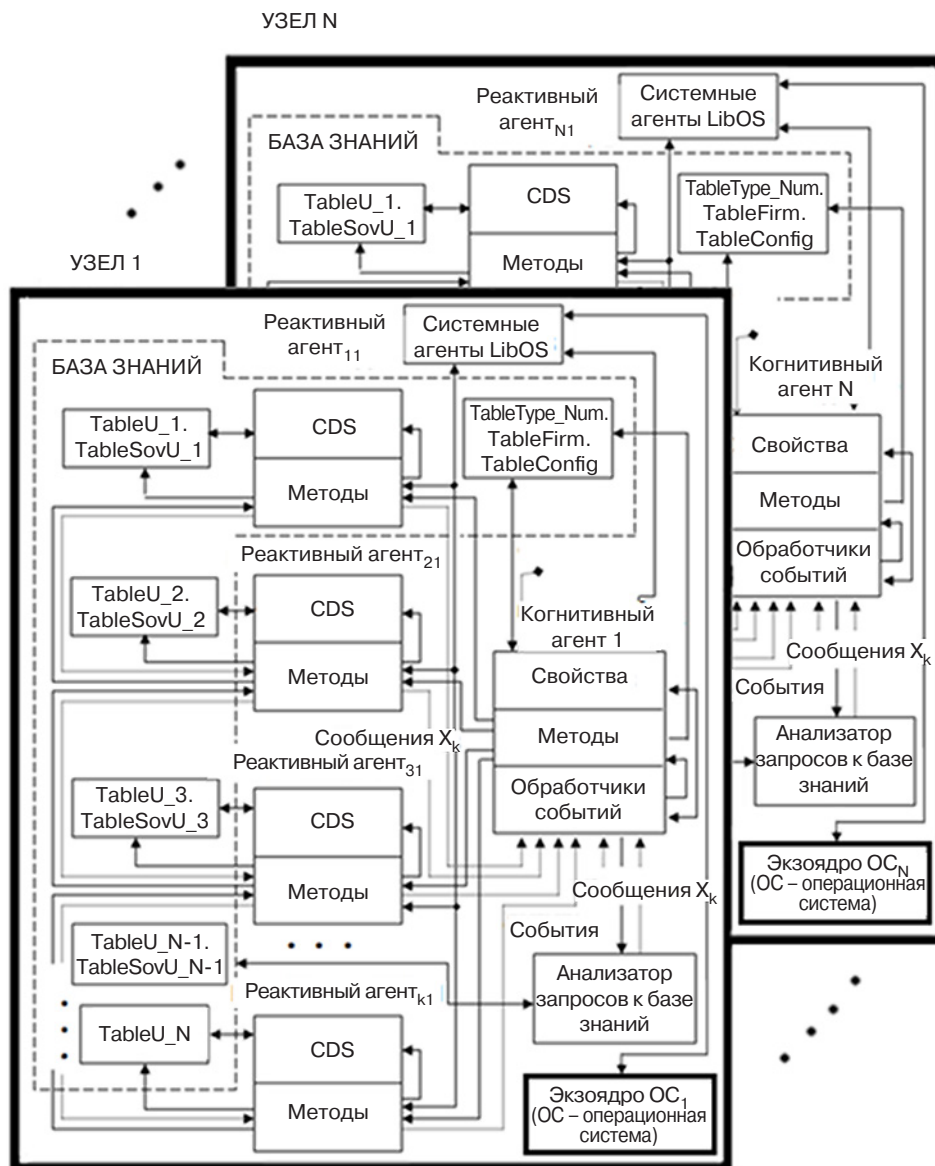


Рис. 1. Структура МСПОЗ

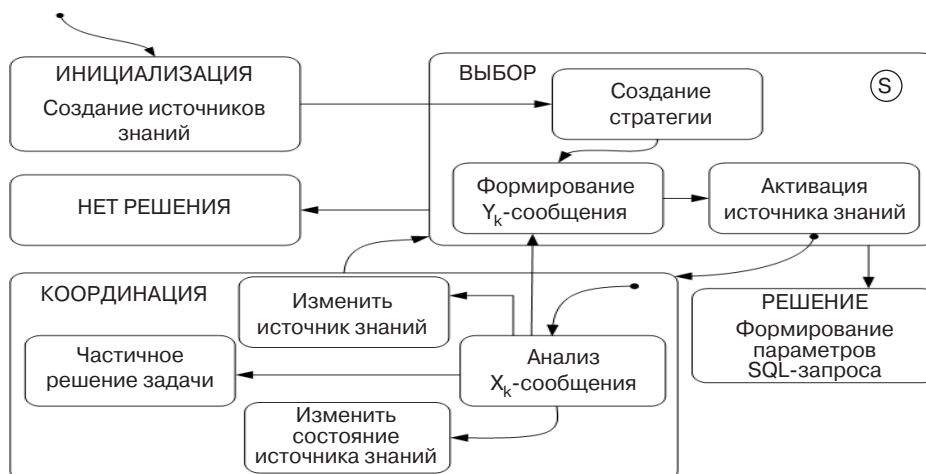


Рис. 2. Диаграмма состояний и переходов когнитивного агента

переход в новое состояние без совершения события, который осуществляется сразу после выполнения действий (или активностей), связанных с предыдущим состоянием.

Из состояния «Инициализация» когнитивный агент переходит в составное состояние «Выбор». В этом состоянии генерируется стратегия и выбирается необходимый источник знаний с учетом установленных на этапе инициализации связей, а также поступающих от других программных агентов осведомительных сообщений. Далее когнитивный программный агент переходит в составное состояние «Координация», в котором активизируются новые источники знаний и координируются действия реактивных программных агентов.

Если реактивные программные агенты не находят согласованного решения (состояние «Нет решения»), когнитивный программный агент возвращается во вложенное состояние частичного решения задачи. Если решение найдено (состояние «Решение»), то полученные на данном этапе данные используются для формирования новых запросов к базе знаний.

Для работы с базой знаний реализованы четыре типа методов, формирующих запросы и обрабатывающих результаты этих запросов:

- анализ (метод ANS) – реализует логический анализ событий;
- ассоциация (метод ASS) – используется для получения ответов на запросы о связях между объектами и событиями;
- сравнение событий или объектов (метод CMP);
- спецификация объектов (метод VAL).

Для спецификации объектов могут использоваться как четкие, так и нечеткие запросы к базе

знаний. При реализации нечетких запросов могут использоваться различные типы функций принадлежности, которые выбираются инженером по знаниям в результате вычислительных экспериментов.

Когнитивные программные агенты координируют работу реактивных программных агентов, ассоциированных с локальными источниками знаний. Пример диаграммы взаимодействия реактивных программных агентов одного из узлов МСПОЗ представлен на рис. 3.

На диаграмме взаимодействия время движется сверху вниз, программные компоненты доступа к базе знаний и агенты представлены вертикальными линиями. Сообщения между агентами (компонентами) отмечены горизонтальными стрелками. При получении сообщения каждый агент (источник знаний) осуществляет вызов соответствующего программного метода (функции-члена класса $F()$ или $S()$) и возвращает результат. Слева на рисунке приведены комментарии. В данном случае каждый реактивный программный агент с номером k взаимодействует только со своим ближайшим соседом с номером $(k - 1)$. Программные агенты этой группы последовательно выполняют возложенные на них задачи в рамках одного процесса (без переключения контекста). Приоритеты реактивных программных агентов узла МСПОЗ устанавливаются в соответствии с порядковым номером агента. Первый реактивный агент использует высокоприоритетные фреймы, связанные с таблицами TableU_1 и TableSovU_1. Программный агент с номером k имеет самый низкий приоритет и связан с таблицей TableU_N.

Если узлом МСПОЗ является мультипроцессорная система, то агенты этого узла могут действовать одновременно. Для организации параллельных

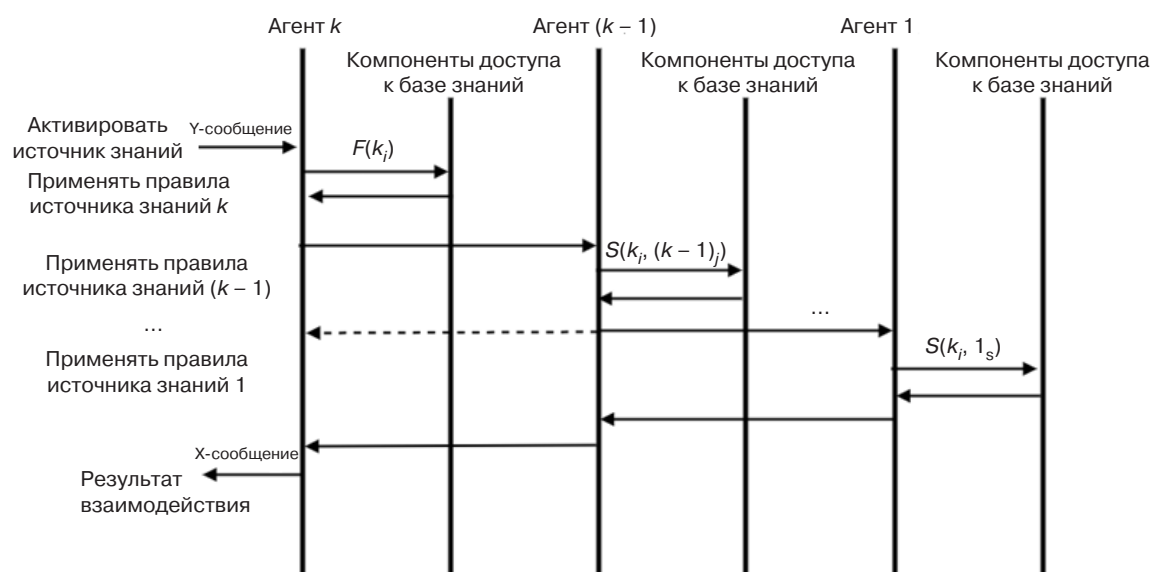


Рис. 3. Схема взаимодействия реактивных программных агентов

вычислений прикладные программные агенты каждого узла МСПОЗ распределяются по группам с использованием матриц совместимости и включения.

Матрица совместимости S имеет следующий вид:

$$S = \begin{bmatrix} 0 & s_{12} & s_{13} & \dots & s_{1M} \\ s_{21} & 0 & & \dots & s_{2M} \\ s_{31} & s_{32} & 0 & \dots & s_{3M} \\ \dots & \dots & \dots & \dots & \dots \\ s_{M1} & s_{M2} & s_{M3} & \dots & 0 \end{bmatrix} \begin{matrix} S_1 \\ S_2 \\ S_3 \\ \dots \\ S_M \end{matrix},$$

где $s_{ij} = 1$, если агенты A_i и A_j должны работать параллельно, иначе $s_{ij} = 0$; S_i – i -я строка матрицы совместимости S ; M – число агентов.

Для распределения программных агентов узла по группам используется матрица включения R :

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1M} \\ r_{21} & r_{22} & \dots & r_{2M} \\ \dots & \dots & \dots & \dots \\ r_{H1} & r_{H2} & \dots & r_{HM} \end{bmatrix} \begin{matrix} R_1 \\ R_2 \\ \dots \\ R_H \end{matrix},$$

где H – число групп; $r_{ij} = 1$, если агент A_i включен в группу Y_j . Агент A_i входит в группу Y_j , если $S_i \cap R_j = \emptyset$, т.е. строки матрицы не пересекаются.

Для оптимального разбиения множества агентов на подмножества при использовании матриц совместимости и включения необходимо учитывать структуру узла МСПОЗ, функциональные особенности программных агентов и их требования к вычислительным ресурсам.

В мультиагентном обучении с подкреплением (multi-agent reinforcement learning, MARL) окружающая среда зависит от всех программных агентов. В отличие от централизованного обучения, при котором программные агенты полностью контролируют вычислительный процесс, при этом политики распределяются центральным агентом, в децентрализованной модели, используемой в МСПОЗ, агенты независимы и могут обмениваться опытом и политиками. В децентрализованной модели выполнение и обучение реализуются локально, что позволяет прикладным программным агентам адаптироваться к местному восприятию окружающей среды (рис. 4).

Децентрализованная модель мультиагентного обучения с подкреплением позволяет использовать стандартные алгоритмы RL. Прикладные программные агенты МСПОЗ обучаются через серию поощрений и наказаний на основе алгоритма Actor-Critic, в котором стратегия генерирует действия, а функция ценности критикует эти действия.

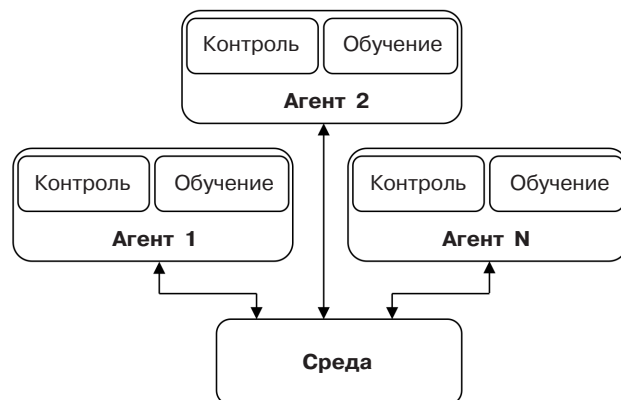


Рис. 4. Децентрализованная модель мультиагентного обучения с подкреплением

Поскольку для каждого программного агента имеются актор и критик, агенты могут иметь разные стратегии (политики) и вознаграждения. Агенты МСПОЗ сотрудничают для оптимизации общей долгосрочной цели. Аппроксимация функций в МСПОЗ реализуется на базе нейронных сетей, которые моделируют как функции политики, так и функции ценности.

Для эффективной реализации мультиагентного обучения с подкреплением необходимо решить проблемы многомерных и мультимодальных целей, масштабируемости, нестабильности и оптимальности [7].

В процессе решения подзадач программные агенты используют микросервисы, которые для повышения надежности и производительности дублируются на разных узлах МСПОЗ. Системные программные агенты распределяют вычислительную нагрузку и управляют микросервисами на основе данных, которые предоставляют агенты-мониторы (рис. 5).

СИНТЕЗ РАСПРЕДЕЛЕННЫХ КОГНИТИВНЫХ СТРУКТУР ДАННЫХ

Выполнение операций со структурами данных в МСПОЗ, будь то пополнение, либо извлечение данных, осуществляется по ключам вследствие большой размерности создаваемых когнитивных структур данных. Для этого вся логическая структура CDS должна быть разбита на ряд кластеров, которые имеют наименьшую взаимосвязь при различных ограничениях. В данном случае примем ограничения на размерность кластеров и ограничения на степень семантической близости логических записей, входящих в кластеры. При этом будет учитываться тип используемых систем хранения данных.

Введем бинарный параметр Z_{kj}^i , который характеризует использование k -м запросом i -й группы данных, отнесенной к j -й логической записи. Расчет данного параметра основывается на бинарной переменной a_{ik} , равной единице, если i -я группа данных

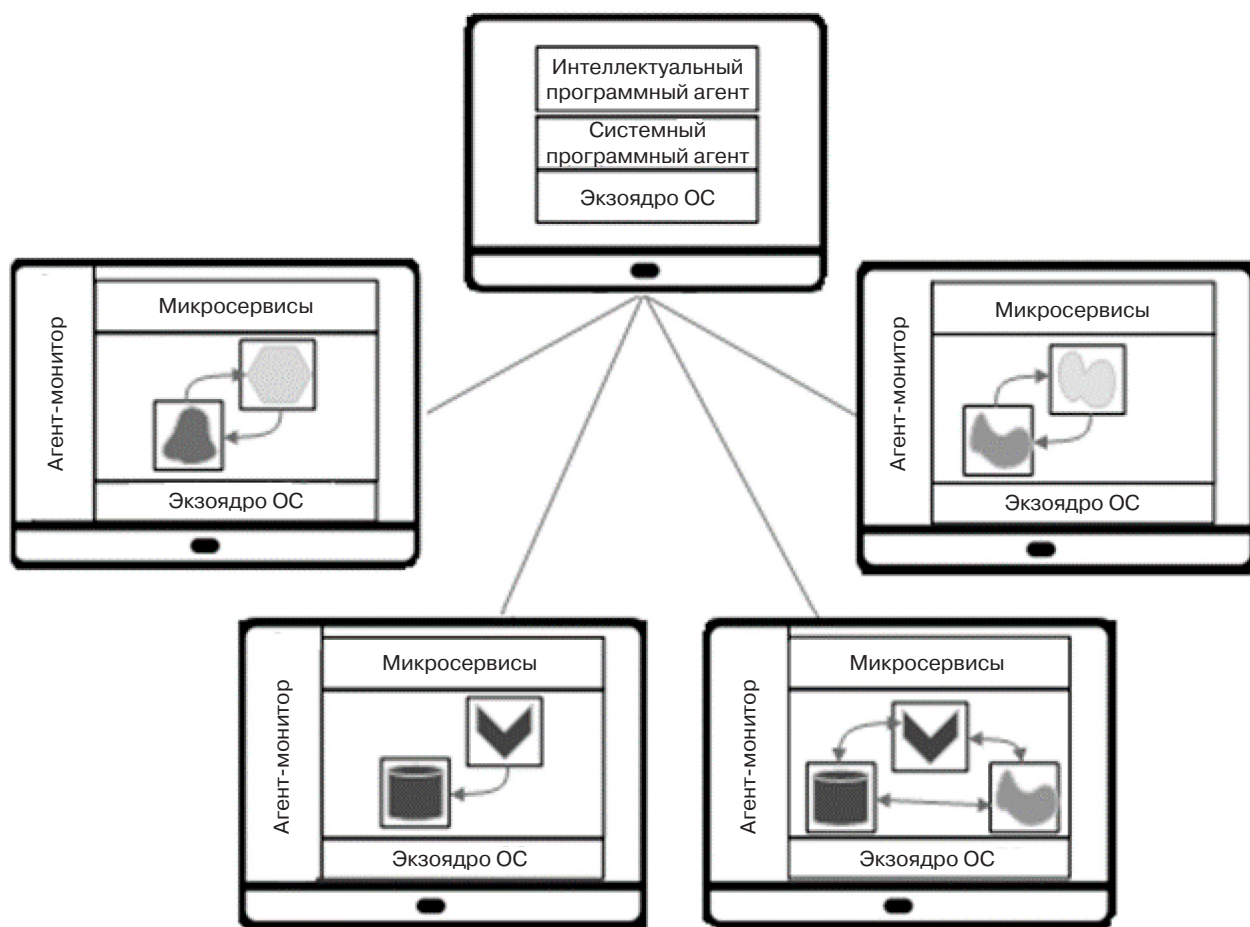


Рис. 5. Микросервисы и системные агенты-мониторы

включена в k -й запрос, и нулю – в противном случае. Учитывая по логике связи, что группы данных входят в логические записи, этот параметр усилен переменной x_{ij} , которая по аналогии равна единице, если i -я группа включена в j -ю логическую запись; $x_{ij} = 0$ – в противном случае.

Имеем следующий расчет параметра Z_{kj}^i :

$$Z_{kj}^i = \begin{cases} 1, & \text{при } \sum_{i=1}^I a_{ik} x_{ij} \geq 1, \\ 0, & \text{при } \sum_{i=1}^I a_{ik} x_{ij} = 0. \end{cases}$$

Полную задачу синтеза распределенной структуры данных для МСПОЗ будем решать с учетом критерия минимального суммарного времени исполнения запросов пользователей при таких ограничениях, как уникальность включения наборов данных в запись, длина логической записи, общее число типов логических записей в структуре, время поиска данных по запросам, уникальность входных узлов в структуре и их общее число.

В данной работе приведем приближенный алгоритм решения задачи синтеза оптимальной распределенной структуры данных по критерию минимального общего времени выполнения запроса. Итак, определяем распределение групп по критерию минимума суммарного трафика, для чего используем приближенный алгоритм распределения кластеров данных между сервером и клиентами локальной сети. И далее сводим канонический граф структуры данных к несвязному графу с вычислением веса каждой группы данных.

Вес группы данных включает веса самой группы и веса дуг с учетом требований пользователей МСПОЗ:

$$V_i = V_i^{\text{ГР}} + V_{ii'}^{\text{СВ}},$$

где $V_i^{\text{ГР}}$ – общий вес группы данных; $V_{ii'}^{\text{СВ}}$ – вес дуг связанного графа канонической структуры данных; i' – индекс группы, смежной с i -й группой данных.

$$V_i^{\text{ГР}} = \sum_{k=1}^{k_0} \sum_{p=1}^{p_0} \gamma_{kp}^3 \delta_{kp}^3 \vartheta_{pi},$$

$$V_{ii'}^{CB} = \sum_{k=1}^{k_0} \sum_{p=1}^{p_0} \gamma_{kp}^3 \delta_{kp}^3 \vartheta_{pi} \sum_{i' \neq i}^I \vartheta_{pi'} a_{ii'}^\Gamma,$$

где γ_{kp}^3 – частота использования запросов пользователями; δ_{kp}^3 – элементы матрицы запросов пользователей; $\vartheta_{pi'}$ – матрица групп данных при исполнении запросов; $a_{ii'}^\Gamma$ – матрица семантической смежности i -й группы данных с группой данных, имеющей индекс i' .

Для конкретной i -й группы вес будет составлять:

$$V_i = \sum_{k=1}^{k_0} \sum_{p=1}^{p_0} \gamma_{kp}^3 \delta_{kp}^3 \vartheta_{pi} \left(1 + \sum_{i' \neq i}^I \vartheta_{pi'} a_{ii'}^\Gamma \right).$$

Далее граф компьютерной сети преобразуем в несвязный граф с вычислением веса узла:

$$V_r = t_r + \sum_{r' \neq r}^{R_0} t_{rr'},$$

где t_r – общая средняя продолжительность обработки данных в r -м узле, состоящая из времени декомпозиции запроса на подзапросы, выбора маршрута, установления соединения; $t_{rr'}$ – средняя продолжительность передачи данных между узлами, определяемая на основе матрицы логических расстояний между серверами узлов компьютерной сети.

Далее формируется матрица $\mathbf{V} = \|v_{ir}\|$, элементы которой определены как декартово произведение веса каждого узла на вес каждой группы данных:

$$v_{ir} = V_i \times V_r \text{ для } i = \overline{1, I}, \quad r = \overline{1, R_0}.$$

После этого проблему

$$\min_{\{x_{ir}\}} \sum_{i=1}^I \sum_{r=1}^{R_0} v_{ir} x_{ir}$$

решаем со следующими ограничениями:

- количество групп данных, локализация которых возможна на одном узле:

$$\sum_{i=1}^I x_{ir} \leq N_r, \quad r = \overline{1, R_0};$$

- допустимая избыточность групп по узлам сети:

$$\sum_{r=1}^{R_0} x_{ir} \leq M_i, \quad i = \overline{1, I};$$

- объем доступной внешней памяти системы хранения данных:

$$\sum_{i=1}^I x_{ir} \rho_i \pi_i \leq \eta_r^{B3Y},$$

где ρ_i – значения длин групп данных; π_i – значения числа экземпляров в группах; η_r^{B3Y} – объем доступной памяти внешнего запоминающего устройства (ВЗУ) на узле; $x_{ir} = 1$, если i -я группа данных включена в r -й сетевой узел, иначе $x_{ir} = 0$.

Решение линейной задачи целочисленного программирования позволяет определить оптимальную локализацию групп данных по узлам сети, и, в дополнение, позволяет определить оптимальную структуру групп данных, размещенных по узлам сети.

На следующем этапе в каждом узле сети решаем задачу оптимального распределения групп данных узла по типам логических записей по критерию минимума общего времени локальной обработки данных. Количество задач синтеза здесь определяется количеством узлов сети. Исходными данными являются подграфы графа канонической структуры данных, а также временные и объемные характеристики подграфов их канонической структуры, набор запросов пользователей и узлов сети [14, 15]. Задача синтеза решается с помощью приближенных алгоритмов с ограничениями на количество групп в записи, на уникальность включения групп в запись, на стоимость хранения информации, на общее время обслуживания запроса. В результате определяются логические структуры базы данных для каждого узла сети, формируются матрицы распределения множества групп данных по типам логических записей, а затем – матрицы распределения множества групп записей по узлам сети.

Результаты решения задачи синтеза распределенных когнитивных структур данных имеют важное практическое значение для проектирования оптимальной структуры РБЗ и возможности формирования спецификаций для запросов и корректировки распределенных данных.

ЗАКЛЮЧЕНИЕ

В работе представлен методологический подход к разработке МСПОЗ. Описаны функциональная и структурная организации мультиагентного решателя, модели прикладных и системных программных агентов, методика проектирования и управления РБЗ. Рассмотрены методы распределения программных агентов по узлам МСПОЗ, определения доступности микросервисов, обеспечения надежной и согласованной работы вычислительных узлов. Приведены примеры диаграмм взаимодействия, состояний и переходов когнитивных и реактивных программных агентов МСПОЗ. Рассмотрены алгоритмы оптимизации логической структуры РБЗ. Оптимизация РБЗ позволяет повысить эффективность МСПОЗ по временным,

объемным и стоимостным характеристикам, делает ее более производительной, гибкой и функциональной. Полученные результаты подтверждают эффективность представленного подхода к разработке МСПОЗ.

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

1. Zaytsev E.I., Khalabiya R.F., Stepanova I.V., Bunina L.V. Multi-Agent System of Knowledge Representation and Processing. In: Kovalev S., Tarassov V., Snasel V., Sukhanov A. (Eds.). *Proceedings of the Fourth International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'19). Advances in Intelligent Systems and Computing*. Springer; 2020. V. I. P. 131–141. https://doi.org/10.1007/978-3-030-50097-9_14
2. Baranauskas R., Janaviciute A., Jasinevicius R., Jukavicius V. On Multi-Agent Systems Intellectics. *Inf. Technol. Control*. 2015;44(1):112–121. <https://doi.org/10.5755/j01.itc.44.1.8768>
3. Darweesh S., Shehata H. Performance Evaluation of a Multi-Agent System using Fuzzy Model. *2018 First International Workshop on Deep and Representation Learning (IWDRL)*. 2018. P. 7–12. <https://doi.org/10.1109/IWDRL.2018.8358208>
4. Рассел С., Норвиг П. *Искусственный интеллект: современный подход*. Т. 2. Знания и рассуждения в условиях неопределенности: пер. с англ. СПб.: Дialeктика; 2021. 480 с.
5. Рассел С., Норвиг П. *Искусственный интеллект: современный подход*. Т. 3. Обучение, восприятие и действие: пер. с англ. СПб.: Дialeктика; 2022. 640 с.
6. Саттон Р.С., Барто Э.Дж. *Обучение с подкреплением*: пер. с англ. М.: ДМК Пресс; 2020. 552 с.
7. Грессер Л., Кенг В.Л. *Глубокое обучение с подкреплением: теория и практика на языке Python*: пер. с англ. СПб.: Питер; 2022. 416 с.
[Graesser L., Keng W.L. *Foundations of Deep Reinforcement Learning: Theory and Practice in Python*. Addison-Wesley Professional; 2020. 416 p.]
8. Памперла М., Фергюсон К. *Глубокое обучение и игра в го*: пер. с англ. М.: ДМК Пресс; 2020. 372 с.
[Pumperla M., Ferguson K. *Deep Learning and the Game of Go*. Manning; 2019. ISBN 978-1-6172-9532-4. 384 p.]
9. Редько В.Г. *Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики*. М.: ЛИБРОКОМ; 2013. 224 с.
10. Winder P. *Reinforcement Learning. Industrial Applications of Intelligent Agents*. O'Reilly Media, Inc.; 2021. 382 p.
11. Houhamdi Z., Athamena B., Abuzaineddin R., Muhairat M. A Multi-Agent System for Course Timetable Generation. *TEM Journal*. 2019;8(1):211–221. <https://doi.org/10.18421/TEM81-30>
12. Aly S., Badoor H. Performance Evaluation of a Multi-Agent System using Fuzzy Model. *2018 First International Workshop on Deep and Representation Learning (IWDRL)*. 2018. P. 175–189. <https://doi.org/10.1109/IWDRL.2018.8358208>

REFERENCES

1. Zaytsev E.I., Khalabiya R.F., Stepanova I.V., Bunina L.V. Multi-Agent System of Knowledge Representation and Processing. In: Kovalev S., Tarassov V., Snasel V., Sukhanov A. (Eds.). *Proceedings of the Fourth International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'19). Advances in Intelligent Systems and Computing*. Springer; 2020. V. I. P. 131–141. https://doi.org/10.1007/978-3-030-50097-9_14
2. Baranauskas R., Janaviciute A., Jasinevicius R., Jukavicius V. On Multi-Agent Systems Intellectics. *Inf. Technol. Control*. 2015;44(1):112–121. <https://doi.org/10.5755/j01.itc.44.1.8768>
3. Darweesh S., Shehata H. Performance Evaluation of a Multi-Agent System using Fuzzy Model. *2018 First International Workshop on Deep and Representation Learning (IWDRL)*. 2018. P. 7–12. <https://doi.org/10.1109/IWDRL.2018.8358208>
4. Russel S., Norvig P. *Iskusstvennyi intellekt: sovremenniy podkhod*. T. 2. Znaniya i rassuzhdeniya v usloviyakh neopredelennosti (*Artificial Intelligence: A Modern Approach*. V. 2. Knowledge and Reasoning under Uncertainty); transl. from Engl. St. Petersburg: Dialektika; 2021. 480 p. (in Russ.).
5. Russel S., Norvig P. *Iskusstvennyi intellekt: sovremenniy podkhod*. T. 3. Obuchenie, vospriyatie i deistvie (*Artificial Intelligence: A Modern Approach*. V. 3. Learning, Perception, and Action); transl. from Engl. St. Petersburg: Dialektika; 2022. 640 p. (in Russ.).
6. Sutton R.S., Barto E.J. *Obuchenie s podkrepleniem (Reinforcement Learning)*: transl. from Engl. Moscow: DMK Press; 2020. 552 p. (in Russ.).
7. Graesser L., Keng W.L. *Glubokoe obuchenie s podkrepleniem: teoriya i praktika na yazyke Python (Reinforcement Learning: Theory and Practice in Python)*: transl. from Engl. St. Petersburg: Piter; 2022. 416 p. (in Russ.).
[Graesser L., Keng W.L. *Foundations of Deep Reinforcement Learning: Theory and Practice in Python*. Addison-Wesley Professional; 2020. 416 p.]
8. Pumperla M., Ferguson K. *Glubokoe obuchenie i igra v go (Deep Learning and the Game of Go)*: transl. from Engl. Moscow: DMK Press; 2020. 372 p. (in Russ.).
9. Red'ko V.G. *Evolutsiya, neironnye seti, intellekt: Modeli i kontseptsii evolyutsionnoi kibernetiki (Evolution, Neural Networks, Intelligence: Models and Concepts of Evolutionary Cybernetics)*. Moscow: LIBROKOM; 2013. 224 p. (in Russ.).
10. Winder P. *Reinforcement Learning. Industrial Applications of Intelligent Agents*. O'Reilly Media, Inc.; 2021. 382 p.

13. Зайцев Е.И. Методология представления и обработки знаний в распределенных интеллектуальных информационных системах. *Автоматизация и современные технологии*. 2008;1:29–34.
14. Batouma N., Sourrouille J.-L. Dynamic adaption of resource aware distributed applications. *Int. J. Grid Distrib. Comput.* 2011;4(2):25–42. URL: http://article.nadiapub.com/IJGDC/vol4_no2/3.pdf
15. Nurmatova E.V., Gusev V.V., Kotliar V.V. Analysis of the features of the optimal logical structure of distributed databases. In: *GRID Workshop Proceedings – GRID 2018 Selected Papers of the 8th International Conference “Distributed Computing and Grid-technologies in Science and Education.”* Dubna: 2018. V. 2267. P. 579–584. URL: <https://ceur-ws.org/Vol-2267/579-584-paper-111.pdf>
11. Houhamdi Z., Athamena B., Abuzaineddin R., Muhairat M. A Multi-Agent System for Course Timetable Generation. *TEM Journal*. 2019;8(1):211–221. <https://doi.org/10.18421/TEM81-30>
12. Aly S., Badoor H. Performance Evaluation of a Multi-Agent System using Fuzzy Model. *2018 First International Workshop on Deep and Representation Learning (IWDRL)*. 2018. P. 175–189. <https://doi.org/10.1109/IWDRL.2018.8358208>
13. Zaytsev E.I. Method of date representation and processing in the distributed intelligence information systems. *Avtomatizatsiya i sovremennye tekhnologii = Automation. Modern Technologies*. 2008;1:29–34 (in Russ.).
14. Batouma N., Sourrouille J.-L. Dynamic adaption of resource aware distributed applications. *Int. J. Grid Distrib. Comput.* 2011;4(2):25–42. Available from URL: http://article.nadiapub.com/IJGDC/vol4_no2/3.pdf
15. Nurmatova E.V., Gusev V.V., Kotliar V.V. Analysis of the features of the optimal logical structure of distributed databases. In: *GRID Workshop Proceedings - GRID 2018 Selected Papers of the 8th International Conference “Distributed Computing and Grid-technologies in Science and Education”*. Dubna; 2018. V. 2267. P. 579–584. Available from URL: <https://ceur-ws.org/Vol-2267/579-584-paper-111.pdf>

Об авторах

Зайцев Евгений Игоревич, к.т.н., доцент, кафедра «Аппаратное, программное и математическое обеспечение вычислительных систем» Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (107996, Россия, Москва, ул. Стромынка, д. 20). E-mail: zajcev@mirea.ru. Scopus Author ID 57218190023, ResearcherID ABA-4823-2020, SPIN-код РИНЦ 9662-7658, <https://orcid.org/0000-0002-1979-5611>

Нурматова Елена Вячеславовна, к.т.н., доцент, кафедра «Аппаратное, программное и математическое обеспечение вычислительных систем» Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (107996, Россия, Москва, ул. Стромынка, д. 20). E-mail: nurmatova@mirea.ru. Scopus Author ID 57205460003, ResearcherID GQI-3212-2022, SPIN-код РИНЦ 7036-3661, <https://orcid.org/0000-0001-8511-0978>

About the authors

Evgeniy I. Zaytsev, Cand. Sci. (Eng.), Associate Professor, Department of Hardware Software and Mathematical Support of Computing System, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (20, Stromynka ul., Moscow, 107996 Russia). E-mail: zajcev@mirea.ru. Scopus Author ID 57218190023, ResearcherID ABA-4823-2020, RSCI SPIN-code 9662-7658, <https://orcid.org/0000-0002-1979-5611>

Elena V. Nurmatova, Cand. Sci. (Eng.), Associate Professor, Department of Hardware Software and Mathematical Support of Computing System, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (20, Stromynka ul., Moscow, 107996 Russia). E-mail: nurmatova@mirea.ru. Scopus Author ID 57205460003, ResearcherID GQI-3212-2022, RSCI SPIN-code 7036-3661, <https://orcid.org/0000-0001-8511-0978>

Роботизированные комплексы и системы.
Технологии дистанционного зондирования неразрушающего контроля
Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing

УДК 004.89; 004.93
<https://doi.org/10.32362/2500-316X-2023-11-4-26-35>



НАУЧНАЯ СТАТЬЯ

Алгоритмы визуального анализа внешней среды автономным мобильным роботом в задаче уборки территории

М.Э. Беляков[®],
С.А.К. Диане

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[®] Автор для переписки, e-mail: beliakow.m@gmail.com

Резюме

Цели. В настоящее время опасной глобальной тенденцией становятся нарастающие темпы загрязнения огромных по площади территорий различными типами бытовых отходов. В связи с этим актуальной потребностью является создание робототехнических комплексов, способных в автономном режиме осуществлять сбор такого мусора. Одной из ключевых составляющих подобных комплексов должна стать система технического зрения для детекции и взаимодействия с целевыми объектами. Цель работы – разработка алгоритмического обеспечения системы технического зрения робототехнических комплексов в задаче уборки территории.

Методы. В рамках предложенной структуры системы визуального анализа внешней среды были оптимизированы под задачу распознавания мусора алгоритмы детекции и классификации объектов различного внешнего вида с применением технологии сверточных нейронных сетей. Настройка нейросетевого детектора производилась методом градиентного спуска на открытой базе обучающих примеров TACO. Для определения геометрических параметров плоского участка местности в поле зрения робота и оценки координат объектов на местности использована матрица гомографии, формируемая с учетом информации о характеристиках и расположении видеокамеры в пространстве.

Результаты. Разработанное программно-алгоритмическое обеспечение системы технического зрения для мобильного робота, оснащаемого монокулярной видеокамерой, реализует функции нейросетевой детекции и классификации объектов в кадре, а также проекции найденных объектов на карту местности для их последующего сбора.

Выводы. Проведенные экспериментальные исследования показали, что разработанная система визуального анализа внешней среды автономного мобильного робота обладает достаточной эффективностью для решения поставленных задач, в т.ч. для обнаружения мусора в поле зрения автономного мобильного робота.

Ключевые слова: нейросетевая детекция, техническое зрение, гомография, мобильные роботы, уборка территории

• Поступила: 10.03.2022 • Доработана: 28.02.2023 • Принята к опубликованию: 02.05.2023

Для цитирования: Беляков М.Э., Диане С.А.К. Алгоритмы визуального анализа внешней среды автономным мобильным роботом в задаче уборки территории. *Russ. Technol. J.* 2023;11(4):26–35. <https://doi.org/10.32362/2500-316X-2023-11-4-26-35>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Algorithms for the visual analysis of an environment by an autonomous mobile robot for area cleanup

Maksim E. Beliakov[@],
Sekou Abdel Kader Diane

MIREA – Russian Technological University, Moscow, 119454 Russia

[@] Corresponding author, e-mail: beliakow.m@gmail.com

Abstract

Objectives. At present, increasing rates of pollution of vast areas by various types of household waste are becoming an increasingly serious problem. In this connection, the creation of a robotic complex capable of performing autonomous litter collection functions becomes an urgent need. One of the key components of such a complex comprises a vision system for detecting and interacting with target objects. The purpose of this work is to develop the underlying algorithmics for the vision system of robots executing area cleaning functions.

Methods. Within the framework of the proposed structure of the system for visual analysis of the external environment, algorithms for detecting and classifying objects of various appearance have been developed using convolutional neural networks. The neural network detector was set up by gradient descent on the open dataset of TACO training samples. To determine the geometric parameters of a surface in the field of view of the robot and estimate the coordinates of objects on the ground, a homography matrix was formed to take into account information about the characteristics and location of the video camera.

Results. The developed software and algorithms for a mobile robot equipped with a monocular video camera are capable of implementing the functions of neural network detection and classification of litter objects in the frame, as well as projection of found objects on a terrain map for their subsequent collection.

Conclusions. Experimental studies have shown that the developed system of visual analysis of the external environment of an autonomous mobile robot has sufficient efficiency to solve the tasks of detecting litter in the field of view of an autonomous mobile robot.

Keywords: neural detection, computer vision, homography, mobile robots, territory cleaning

• Submitted: 10.03.2022 • Revised: 28.02.2023 • Accepted: 02.05.2023

For citation: Beliakov M.E., Diane S.A.K. Algorithms for the visual analysis of an environment by an autonomous mobile robot for area cleanup. *Russ. Technol. J.* 2023;11(4):26–35. <https://doi.org/10.32362/2500-316X-2023-11-4-26-35>

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 большинстве стран мира, включая Россию, произведенные в ходе повседневной деятельности человека отходы подвергаются одному из двух способов уничтожения: либо сжигаются на мусоросжигательных заводах, либо закапываются на свалках и полигонах. Оба способа уничтожения негативно сказываются на экологии. Сжигание отходов сопровождается выбросом токсичных газов и пыли в атмосферу, что в свою очередь приводит к глобальному потеплению и загрязнению водоемов, лесов и городов на другом краю земли от места, где все это было выпущено в атмосферу. Большинство мусоросжигательных заводов в постиндустриальных странах избавляются от токсичной и загрязняющей части отходов иначе, таким образом, сохраняя чистый воздух, однако большая часть мусора при этом попросту закапывается на полигонах.

Вышеперечисленные факты обуславливают важность сортировки отходов по категориям, т.к. сортировка позволяет снизить объем загрязнений в результате уничтожения мусора. Отходы, производимые человеком, можно разделить на опасные и безопасные [1]. К безопасным отходам относятся пищевые остатки, картон и бумага, целлофан, органический мусор (листва, древесина). Такие отходы при разложении не отравляют почву, источники воды и не оказывают сильного негативного влияния на окружающую среду, кроме того, такие материалы легко подвергаются переработке в новую продукцию. К опасным можно отнести батарейки и аккумуляторы, лакокрасочные изделия, полиэтилен и пр. Такие отходы отравляют почву и водоемы в радиусе нескольких километров. Таким образом, сортировка отходов помогает снизить объем загрязнений за счет сокращения опасных отходов, попадающих на полигоны вместо надлежащих мест утилизации; уменьшения количества перерабатываемых отходов, попадающих на полигоны вместо станций переработки; снижения расходов на вывоз мусора, т.к. больше предметов пойдет на переработку.

Не в меньшей степени заслуживает внимания проблема оставления мусора в общественных местах. Эта проблема представляет непреднамеренное или умышленное засорение общественных мест, таких как улицы, парки, места для пикников и пляжи, мелким мусором. Согласно одному из исследований¹, наиболее часто встречающиеся объекты мусора в общественных пространствах – это сигаретные окурки, пищевые обертки, пластиковые бутылки и пр. Учитывая разнообразие и огромные площади

территорий, подверженных загрязнению, а также тенденцию повсеместной автоматизации, потребностью сегодняшнего дня является создание автоматизированных робототехнических систем, способных автономно собирать мусор.

В статье будет рассмотрена система визуального анализа внешней среды для автономного мобильного робота, позволяющая искать, распознавать категорию отходов и локализовать мусор на заданной местности для последующей сборки. Для работы системы достаточно наличия одной камеры на борту робота.

АНАЛИЗ РАЗРАБОТОК В ОБЛАСТИ РОБОТИЗИРОВАННОЙ УБОРКИ ТЕРРИТОРИЙ

Целый ряд исследовательских и научно-производственных организаций по всему миру работает над вопросом автоматизации уборки местности от мусора.

Так, в работе [2] описана робототехническая система (РТС) для распознавания объектов мусора и дальнейшего их сбора, оснащенная захватным устройством, камерой и системой визуального оучувствления. Данная РТС представляет простой в реализации аппаратно-программный комплекс, который может являться базой и отправной точкой при создании автономного мобильного робота в задаче уборки территорий.

Поскольку в системе визуального анализа данной РТС используется предобученная модель машинного обучения MobileNet [3], то система способна распознавать как мусор только бутылки. Также в работе используется упрощенная система наведения и оценки расстояния до объекта. После обнаружения объекта на кадре с камеры используется разница между центром кадра и центром габаритной рамки объекта, полученной на выходе детектора MobileNet, для поворота и наведения робота на объект.

При управлении движением робота используется ПИД-регулятор (пропорционально-интегрально-дифференцирующий регулятор) [4]. Расстояние до объекта высчитывается по параметрам расположения камеры на базе робота. Учитываются угол наклона, высота от пола и угол раскрытия камеры. Недостаток такого подхода в том, что при распознавании нескольких объектов нет возможности построения оптимального маршрута сбора объектов. Кроме того, отсутствует функция сортировки мусора.

В работе [5] также описана робототехническая система для обработки мусора. Однако в данном случае представлена не мобильная, а конвейерная система, оснащенная механическим манипулятором и камерой для визуальной классификации и последующей сегрегационной сборки поступающих по конвейерной ленте отходов. Система визуального анализа способна

¹ https://www.legambienteverona.it/wp-content/uploads/2021/01/Report-Park-Litter_English-final.pdf. Дата обращения 25.01.2022 / Accessed January 25, 2022.

различать 4 класса объектов: бумага, металл, стекло, пластик. Данная система не использует нейросетевые алгоритмы. Ее работа основана на классических алгоритмах обработки изображения и состоит из следующих этапов: распознавание границ объектов на основе соответствующего алгоритма Кэнни [6]; определение порога для отделения объектов от фона и устранения шума; гауссовская фильтрация для размытия деталей объекта; преобразование фото в черно-белый формат с целью использовать яркость канала для поиска границ; контурное распознавание объектов на бинарном изображении. Дальнейшие действия обработки составляют использование дескрипторов Hu moments [7] и Fourier [8] для описания формы объектов и метода Hyperplane K-Nearest Neighbors [9] для классификации по дескрипторам объектов.

МОДЕЛИ И АЛГОРИТМЫ РАСПОЗНАВАНИЯ ТИПОВЫХ ОБЪЕКТОВ В ПОЛЕ ЗРЕНИЯ РОБОТА

Распознавание и локализация типовых объектов осуществляется путем обработки растровых изображений методами компьютерного зрения.

Существуют различные подходы и методы компьютерного зрения. К классическим методам относят алгоритмы для поиска ключевых точек, выделения границ объектов, геометрических преобразований. Такие алгоритмы хорошо подходят для более простых и детерминированных задач с точки зрения внешних факторов – освещение и расстояние до объекта съемки, небольшая вариация формы объекта. Выступающие в качестве альтернативы алгоритмы распознавания на основе глубокого обучения более устойчивы к ложным срабатываниям при изменении освещенности, размера или ракурса объекта. Такие алгоритмы строятся, как правило, с использованием сверточных нейронных сетей и позволяют создавать более совершенные и отказоустойчивые системы компьютерного зрения.

Свое название архитектура сверточной сети получила из-за наличия операции свертки, суть которой заключается в том, что каждый фрагмент изображения умножается на матрицу (ядро) свертки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения. Наилучшие результаты в обобщающей способности сети и ее вычислительной эффективности достигаются при чередовании сверточных слоев и субдискретизирующих слоев. Это приводит к последовательному понижению размерности входных данных и делает сеть устойчивой к незначительным трансформациям анализируемого изображения.

Среди задач компьютерного зрения, решаемых методами сверточных нейронных сетей, можно выделить задачу детекции объектов на изображении [10],

как подходящую для решения задачи, поставленной в данной статье. Задача детекции заключается в том, что обрабатывающий изображение алгоритм должен определить класс и координаты целевых объектов на изображении.

Нейронную сеть можно представить как многомерную функцию, поэтому процесс обучения заключается в оптимизации внутренних многочисленных параметров методом градиентного спуска [11].

Для обучения нейросетевой модели при решении задачи детекции мусора в поле зрения робота был выбран открытый датасет TACO² [12], содержащий 1500 изображений с 4784 размеченными объектами на 28 классов, глобально разделенных на бумагу, стекло, пластик, металл (рис. 1).

В качестве архитектуры нейросетевого детектора было решено использовать алгоритм YOLOv4 [13]. Данная архитектура предоставляет компромисс между точностью детектирования объектов и скоростью работы (до нескольких сотен кадров в секунду при работе на графическом ускорителе). В результате обучения модель была оптимизирована до 0.13 mAP (mean average precision) по 4 классам.



Рис. 1. Примеры размеченных изображений в базе обучающих примеров TACO

Отметим, что метрика mAP [14] является комплексной, учитывающей как погрешности классификации (FP, FN), так и погрешности локализации

² TACO – это растущий набор изображений отходов в дикой природе. Он содержит изображения мусора, сделанные в различных условиях: в лесах, на дорогах и пляжах. [TACO is a growing image dataset of waste in the wild. It contains images of litter taken under diverse environments: woods, roads and beaches.]

объектов на кадре. Поэтому полученное значение точности по метрике mAP является достаточным для практического применения. Результат работы детектора представлен на рис. 2.



Рис. 2. Результат работы детектора YOLOv4

МОДЕЛИ И АЛГОРИТМЫ ЛОКАЛИЗАЦИИ ТИПОВЫХ ОБЪЕКТОВ НА КАРТЕ МЕСТНОСТИ

Поскольку предполагается, что координаты робота на заданной местности известны, то для определения координат объектов мусора необходимо вычислить их координаты относительно робота (рис. 3). Сделать это возможно с помощью матрицы гомографии. Стоит заметить, что данное преобразование справедливо только для объектов на плоской поверхности. Таким образом, предложенная модель определения координат справедлива только для объектов, лежащих на ровной поверхности.

После определения координат объектов мусора на кадре с бортовой камеры мобильного робота, зная угол наклона, высоту над поверхностью, фокальное расстояние камеры, возможно устранить искажение перспективы, преобразовав их в координаты на виде сверху пространства перед роботом.

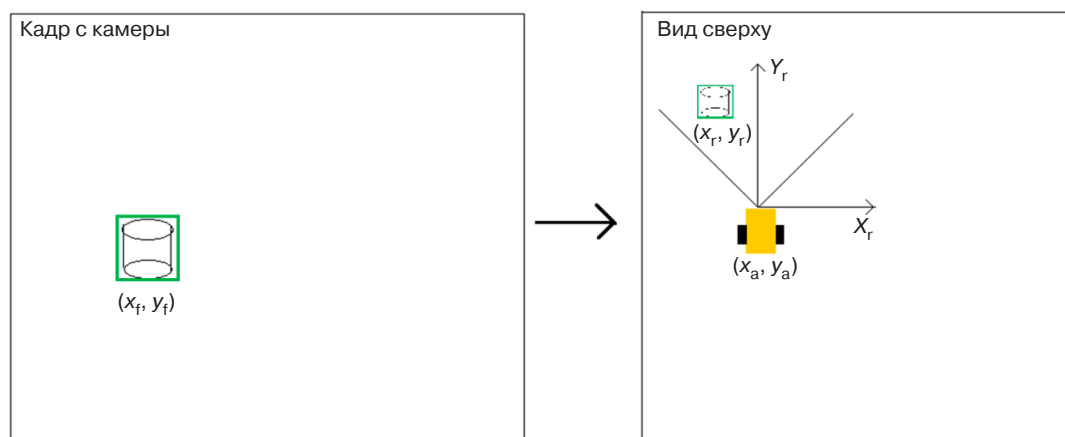


Рис. 3. Вычисление координат объекта относительно робота. (x_r, y_r) – координаты объекта в системе координат относительно робота, (X_r, Y_r) – система координат относительно робота, (x_a, y_a) – координаты робота в глобальной системе координат; (x_f, y_f) – координаты объекта на кадре

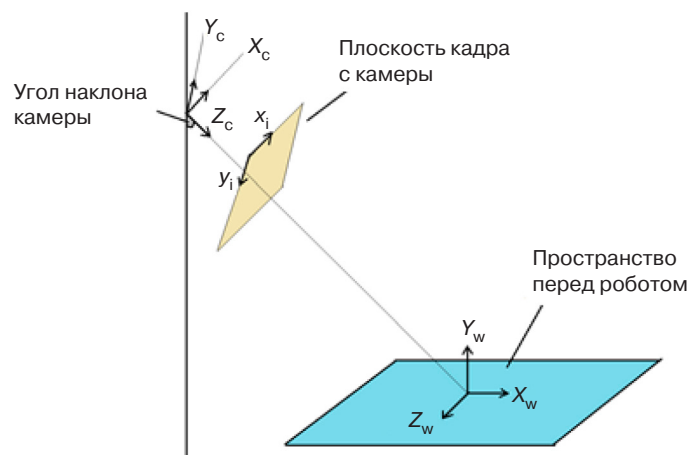


Рис. 4. Геометрическое пояснение матрицы гомографии. (X_c, Y_c, Z_c) – система координат относительно камеры, (X_w, Y_w, Z_w) – система координат в плоскости пространства перед роботом, (x_i, y_i) – координаты точек прямоугольника, обрамляющего объект

Координаты на плоскости кадра и на плоскости пространства определяются следующим соотношением:

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{H}_0 \times \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}. \quad (1)$$

В соответствии с геометрическим пояснением на рис. 4 матрица гомографии $\mathbf{H}_0$ может быть описана следующим образом [15]:

$$\mathbf{H}_0 = \begin{bmatrix} f & x_0 \cos \alpha & x_0 h \cos \alpha \\ 0 & f \sin \alpha + y_0 \cos \alpha & y_0 h \cos \alpha - f h \sin \alpha \\ 0 & \cos \alpha & h \cos \alpha \end{bmatrix}. \quad (2)$$

где f – фокусное расстояние камеры, h – высота камеры, α – угол наклона камеры, (x_0, y_0) – координаты точки пересечения осей изображения.

При условии, что угол наклона камеры отличен от 0° и 90° , к данному преобразованию существует обратное, поэтому вид сверху может быть получен из изображения с искаженной перспективой и наоборот.

Рис. 5 демонстрирует изображение с искаженной перспективой, снятое под углом наклона $\alpha = 45^\circ$, при высоте $h = 0.7$ м, и соответствующее ему преобразование к виду сверху.

Кадр с камеры робота анализируется нейросетевым алгоритмом YOLOv4, который определяет пиксельные координаты объекта на кадре в виде четырех координат точек $p_i = (x_i, y_i)$ обрамляющего прямоугольника (рис. 2), после чего за однозначную координату этого объекта принимается центр прямоугольника $p_c = \left\{ \frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right\}$.

Данная координата претерпевает преобразование (3), таким образом, определяя относительные пиксельные координаты объекта в рамках видимого пространства перед роботом:

$$\begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} = \mathbf{H}_0^{-1} \times \begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix}, \quad (3)$$

где x_m, y_m – координаты объекта на пространстве перед роботом.

Для преобразования пиксельных координат объектов на видимом пространстве перед роботом в метровые координаты относительно робота необходимо вычислить размеры видимого пространства. В соответствии с пояснениями на рис. 6, геометрические параметры пространства перед роботом однозначно задаются горизонтальным углом раскрытия, вертикальным углом раскрытия, высотой и углом наклона камеры.

Расстояние от камеры до ближнего края видимого пространства d_f , расстояние от камеры до дальнего края видимого пространства d_r и длина видимого пространства Y определяются соотношениями:

$$d_f = h \cdot \operatorname{tg}(\alpha - \beta), \quad (4)$$

$$d_r = h \cdot \operatorname{tg}(\alpha + \beta), \quad (5)$$

$$Y = d_r - d_f \quad (6)$$

Полуширина ближнего края видимого пространства w_f , полуширина дальнего края видимого пространства w_r , ширина видимого пространства X определяются соотношениями:

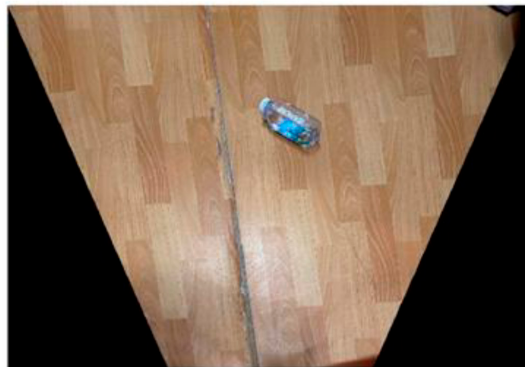
$$w_f = \operatorname{tg} \gamma \sqrt{d_f^2 + h^2}, \quad (7)$$

$$w_r = \operatorname{tg} \gamma \sqrt{d_r^2 + h^2}, \quad (8)$$

$$X = 2w_r. \quad (9)$$



(а)



(б)

Рис. 5. Вид сверху: (а) изображение, снятое на камеру; (б) изображение после устранения искажения перспективы

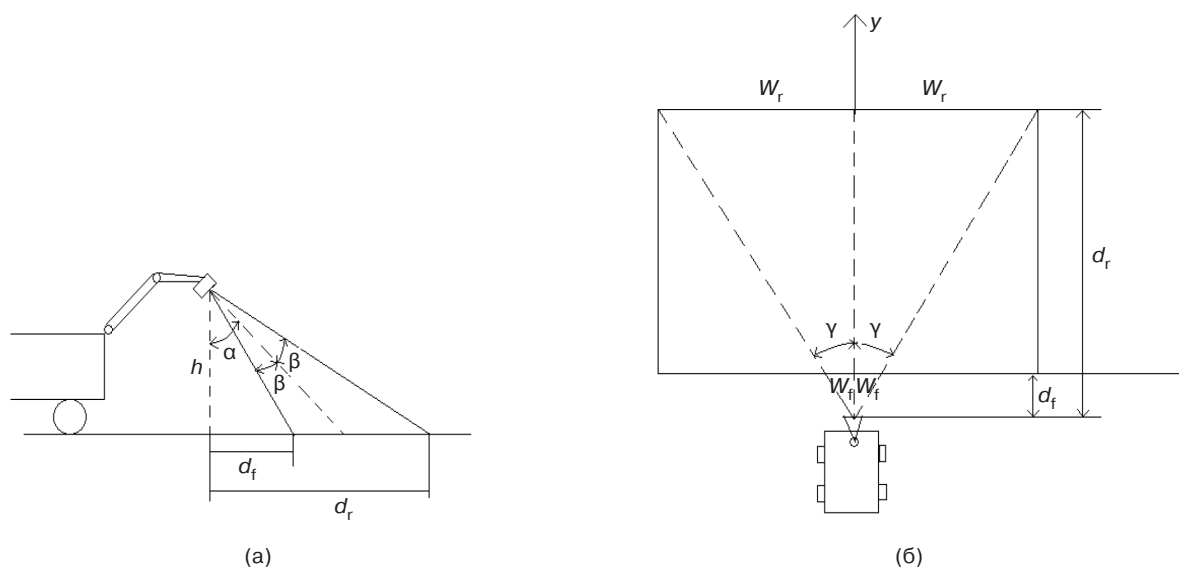


Рис. 6. Геометрия наблюдаемого пространства перед роботом: (а) поле зрения камеры в вертикальной проекции; (б) поле зрения камеры в горизонтальной проекции

В соответствии с вышеизложенным координаты объектов относительно робота определяются соотношениями:

$$x_0 = \frac{X \cdot x_{pxl}}{w_{pxl}} - w_r, \quad (10)$$

$$y_0 = Y \left(1 - \frac{X \cdot x_{pxl}}{h_{pxl}} \right) + d_f, \quad (11)$$

где x_{pxl} , y_{pxl} – экранные координаты объекта на видимом пространстве; w_{pxl} , h_{pxl} – экранная ширина и высота видимого пространства, соответственно.

СТРУКТУРА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Проблематика и задачи, поставленные ранее, определяют структуру программно-алгоритмического обеспечения (рис. 7). Программное обеспечение включает подсистему пользовательского интерфейса, библиотеки приема и предобработки изображения, а также модули распознавания целевых объектов и вычисления его координат относительно мобильного робота.

Отметим, что в состав модуля распознавания объекта входят не только процедуры, непосредственно отвечающие за процесс детекции и классификации, но и файл с выбранной архитектурой нейронной

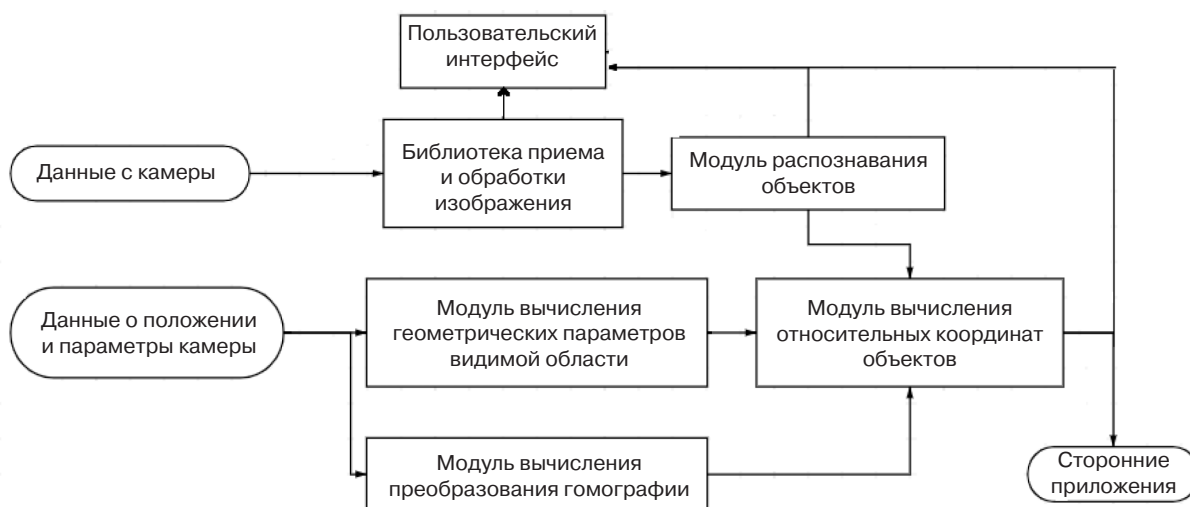


Рис. 7. Структура программно-алгоритмического обеспечения для визуального анализа среды мобильного робота

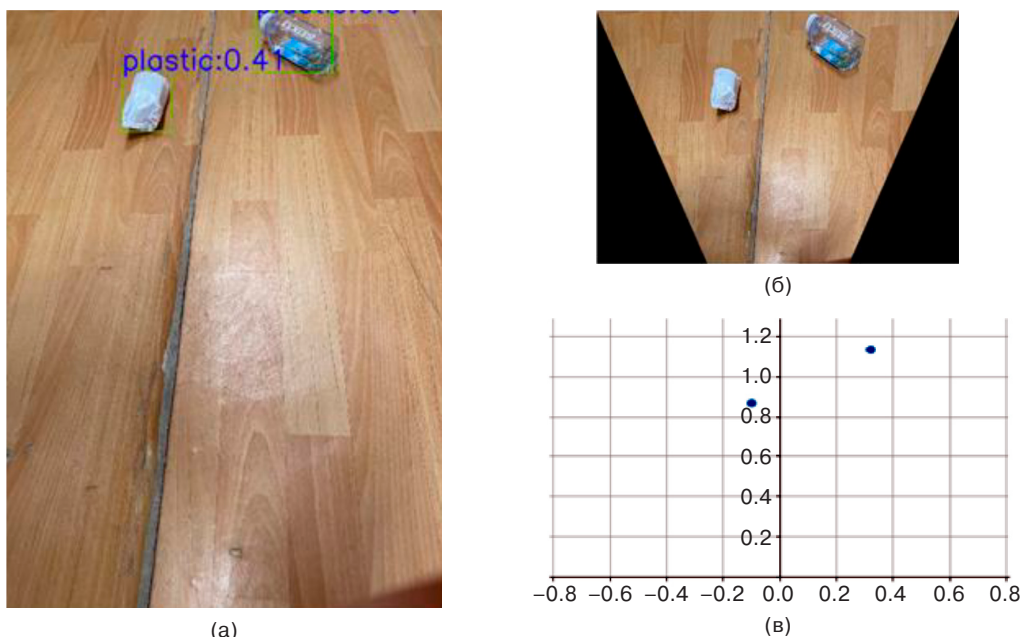


Рис. 8. Распознавание объектов: (а) вид с камеры; (б) вид сверху; (в) карта объектов в поле зрения робота с координатами в метрической системе

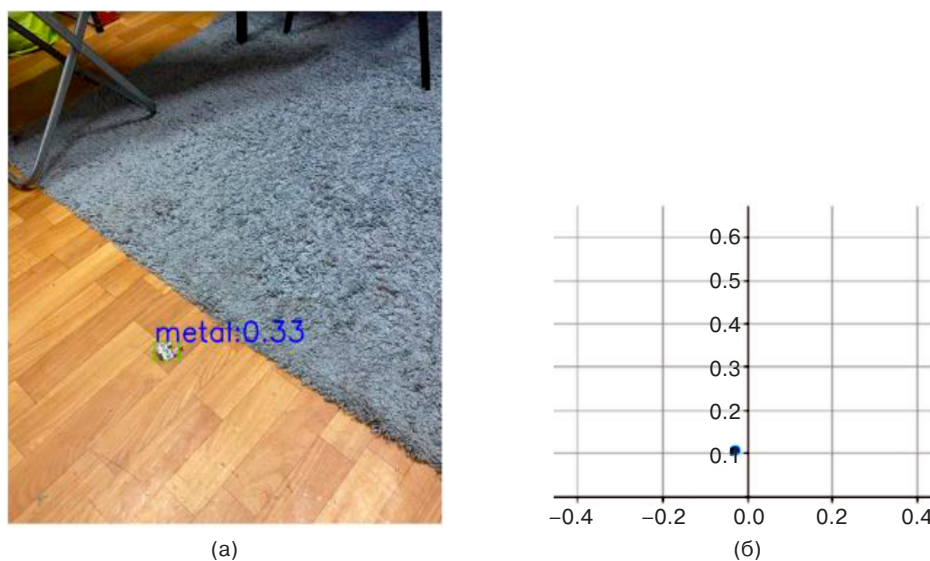


Рис. 9. Распознавание объектов: (а) вид с камеры; (б) карта объектов в поле зрения робота с координатами в метрической системе

сети и ее предварительно настроенными весовыми коэффициентами.

Реализация программного обеспечения на языке программирования Python позволяет упростить интеграцию различных технологий компьютерного зрения и уменьшить время отладки.

ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Проведенные эксперименты подтверждают работоспособность программно-алгоритмического обеспечения. Так, на рис. 8 продемонстрированы процедуры

распознавания и локализации нескольких объектов мусора в поле зрения робота, устранение искажения перспективы и вычисление координат объектов относительно робота в соответствии с высотой камеры $h = 0.5$ м, углом наклона $\alpha = 45^\circ$, вертикальным углом раскрытия камеры $\beta = 23.75^\circ$ и горизонтальным углом раскрытия камеры $\gamma = 30.41^\circ$.

В следующем экспериментальном исследовании (рис. 9) камера находится в положении с высотой $h = 0.5$ м, углом наклона $\alpha = 45^\circ$, другим фоном и объектом другой категории бытовых отходов. Этот случай демонстрирует, что программно-алгоритмическое обеспечение инвариантно к текстуре подстилающей

поверхности, количеству и категориям объектов, а также к их смещению относительно камеры.

ЗАКЛЮЧЕНИЕ

В рамках проведенного исследования показана актуальность автоматизированного раздельного сбора мусора. Существующие исследования и подходы в данной области не объединяют мобильность платформы и интеллектуальный раздельный сбор отходов. При этом большое значение имеет возможность именно гибкого интеллектуального сбора с автономным поиском объектов на заданной территории и возможность раздельного сбора отходов. В рамках данного исследования были достигнуты следующие результаты:

1. Предложена структура (рис. 7) программно-алгоритмического обеспечения для визуального анализа среды мобильного робота в задаче уборки территорий с функцией сортировки отходов.
2. Оптимизированы алгоритмы распознавания объектов разного внешнего вида и различных категорий отходов на изображениях с камеры робота с применением нейросетевых алгоритмов.
3. Разработана модель определения геометрических параметров плоского пространства в поле зрения робота.

4. Разработана модель определения координат объектов в поле зрения робота относительно него в соответствии с положением камеры в пространстве.

Проведенные экспериментальные исследования подтвердили работоспособность и универсальность разработанных алгоритмов. Получены корректные данные при распознавании различных объектов и дальнейшем определении их относительных координат. Кроме того, система показала свою работоспособность при обнаружении объектов разных категорий и различном положении камеры в пространстве с учетом неопределенностей в фоне подстилающей поверхности и количестве искомых объектов.

Дальнейшие перспективы развития представленных исследований связаны с разработкой алгоритмов планирования пути на местности и управления движением мобильного робота в процессе сбора и транспортировки обнаруженных объектов мусора.

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. Черняева Т.К. Актуальные проблемы влияния отходов производства и потребления на объекты окружающей среды и состояние здоровья населения (обзор). *Гигиена и санитария*. 2013;3:32–35.
[Chernyaeva T.K. Actual problems of the impact of production and consumption waste on environment and public health (review of literature). *Gigiena i sanitariya = Hygiene and Sanitation*. 2013;3:32–35 (in Russ.).]
2. Bansal S., Patel S., Shah I., Patel A., Makwana J., Thakker R. *AGDC: Automatic Garbage Detection and Collection*. arXiv preprint arXiv:1908.05849. 2019. <https://doi.org/10.48550/arXiv.1908.05849>
3. Howard A.G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. arXiv preprint arXiv:1704.04861. 2017. <https://doi.org/10.48550/arXiv.1704.04861>
4. Карпов В.Э. *ПИД-управление в нестрогом изложении*. М.: НИИ информационных технологий; 2012. 34 с. URL: http://radioservice.at.ua/_fr/0/Karpov_mobline1.pdf
[Karpov V.E. *PID-upravlenie v nestrogom izlozhenii (PID-Control in a Loose Presentation)*. Moscow; 2012. 34 p. (in Russ.). URL: http://radioservice.at.ua/_fr/0/Karpov_mobline1.pdf]
5. Salmador A., Cid J.P., Novelle I.R. Intelligent Garbage Classifier. *Int. J. Interact. Multimedia Artif. Intell.* 2008;1(1):31–36.
6. Xu Z., Ji X., Wang M., Sun X. Edge detection algorithm of medical image based on Canny operator. *J. Phys.: Conf. Ser.* 2021;1955(1):012080. <https://doi.org/10.1088/1742-6596/1955/1/012080>
7. Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Inform. Theory*. 1962;8(2): 179–187. <https://doi.org/10.1109/TIT.1962.1057692>
8. El-ghazal A., Basir O., Belkasim S. A novel curvature-based shape Fourier Descriptor. *2008 15th IEEE International Conference on Image Processing (ICIP)*. 2008. P. 953–956. <https://doi.org/10.1109/ICIP.2008.4711914>
9. Pascal V., Yoshua B. K-Local Hyperplane and Convex Distance Nearest Neighbor Algorithms. In: Dietterich T.G., Becker S., Ghahramani Z. (Eds.). *Advances in Neural Information Processing Systems*. The MIT Press; 2001. V. 14. P. 985–992.
10. Deng J., Xuan X., Wang W., Li Z., Yao H., Wang Z. A review of research on object detection based on deep learning. *J. Phys.: Conf. Ser.* 2020;1684(1):012028. <https://doi.org/10.1088/1742-6596/1684/1/012028>
11. Бредихин А.И. Алгоритмы обучения сверточных нейронных сетей. *Вестник Югорского государственного университета*. 2019;1(52):41–54. <https://doi.org/10.17816/byusu20190141-54>
[Bredikhin A.I. Training algorithms for convolutional neural networks. *Vestnik Yugorskogo gosudarstvennogo universiteta = Yugra State University Bulletin*. 2019;1(52):41–54 (in Russ.). <https://doi.org/10.17816/byusu20190141-54>]

12. Proença P.F., Simões P. *TACO: Trash Annotations in Context for Litter Detection*. arXiv preprint arXiv: 2003.06975. 2020. <https://doi.org/10.48550/arXiv.2003.06975>
13. Bochkovskiy A., Wang C.Y., Liao H.Y.M. *YOLOv4: Optimal Speed and Accuracy of Object Detection*. arXiv preprint arXiv: 2004.10934. 2020. <https://doi.org/10.48550/arXiv.2004.10934>
14. Liu L., Özsu M.T. (Eds.). *Encyclopedia of Database Systems*. New York, NY, USA: Springer; 2009. V. 6. <https://doi.org/10.1007/978-0-387-39940-9>
15. Yasuda T. (Ed.) *Multi-Robot Systems, Trends and Development*. London, United Kingdom: IntechOpen; 2011. 598 p. <https://doi.org/10.5772/544>

Об авторах

Беляков Максим Эдуардович, бакалавр, кафедра проблем управления Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: beliakow.m@gmail.com. <https://orcid.org/0000-0001-7193-049X>

Диане Секу Абдель Кадер, к.т.н., доцент, кафедра проблем управления Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: sekoudiane1990@gmail.com. ResearcherID T-5560-2017, Scopus Author ID 57188548666, SPIN-код РИНЦ 8691-0290, <https://orcid.org/0000-0002-8690-6422>

About the authors

Maksim E. Beliakov, Bachelor, Department of Control Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). beliakow.m@gmail.com. <https://orcid.org/0000-0001-7193-049X>

Sekou Abdel Kader Diane, Cand. Sci. (Eng.), Associate Professor, Department of Control Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: sekoudiane1990@gmail.com. ResearcherID T-5560-2017, Scopus Author ID 57188548666, RSCI SPIN-code 8691-0290, <https://orcid.org/0000-0002-8690-6422>

УДК 621.314:681.586.7:004.942
<https://doi.org/10.32362/2500-316X-2023-11-4-36-48>



НАУЧНАЯ СТАТЬЯ

Анализ пульсаций DC/DC-преобразователя, построенного по Zeta-топологии, с использованием его предельной непрерывной математической модели

В.К. Битюков,
А.И. Лавренов[@],
Д.А. Малицкий

МИРЭА – Российский технологический университет, Москва, 119454 Россия

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

Резюме

Цели. DC/DC-преобразователь, построенный по Zeta-топологии, является униполярным электронным устройством, которое обеспечивает преобразование входного положительного напряжения в стабилизированное выходное напряжение той же полярности с возможностью его регулирования как ниже входного напряжения, так и выше. Цель работы – проанализировать схемотехнику Zeta-преобразователя. Для этого необходимо решить следующие задачи: при помощи правил Кирхгофа получить системы уравнений, описывавшие работу преобразователя в режимах накопления и передачи энергии; по методике, предложенной А.И. Коршуновым, объединить системы уравнений в предельную непрерывную математическую модель преобразователя; при помощи выражений, описывающих постоянные составляющие токов и напряжений в Zeta-преобразователе, провести анализ их пульсаций и получить уравнения для их расчета; провести сравнение полученных при помощи предельной непрерывной математической модели значений токов и напряжений с результатами моделирования Zeta-преобразователя.

Методы. Задача решена при помощи правил Кирхгофа и методики получения предельной непрерывной математической модели, предложенной А.И. Коршуновым. Результаты проанализированы с использованием схемотехнического моделирования в среде *Multisim*.

Результаты. Показано, что фазовые координаты математической модели стремятся к значениям реальных токов и напряжений преобразователя при частоте коммутации силового ключа более 200 кГц. Установлено высокое соответствие расчетных значений пульсаций и их значений, полученных при моделировании (при изменении коэффициента заполнения).

Выводы. Математические модели являются основой унифицированных методик расчета любых радиоэлектронных схем. Полученная предельная непрерывная математическая модель Zeta-преобразователя позволяет оценить диапазон изменения токов, протекающих через обмотки дросселей, и напряжений на обкладках конденсаторов, их максимальные и минимальные значения при различных параметрах преобразователя, таких как частота коммутации силового ключа, коэффициент заполнения, номиналы элементов и т.д. Эта модель позволяет выполнить рациональный подбор электронной компонентной базы преобразователя.

Ключевые слова: DC/DC-преобразователь, понижающе-повышающий преобразователь, эквивалентная схема, топология Zeta, преобразователь, предельная модель, непрерывная модель, математическая модель, правила Кирхгофа, размах пульсаций

• Поступила: 28.12.2022 • Доработана: 10.02.2023 • Принята к опубликованию: 05.05.2023

Для цитирования: Битюков В.К., Лавренов А.И., Малицкий Д.А. Анализ пульсаций DC/DC-преобразователя, построенного по Zeta-топологии, с использованием его предельной непрерывной математической модели. *Russ. Technol. J.* 2023;11(4):36–48. <https://doi.org/10.32362/2500-316X-2023-11-4-36-48>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Analysis of the DC/DC Zeta topology converter ripples by applying its limiting continuous mathematical model

Vladimir K. Bityukov,
Alexey I. Lavrenov[@],
Daniil A. Malitskiy

MIREA – Russian Technological University, Moscow, 119454 Russia

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

Abstract

Objectives. A DC/DC Zeta topology converter represents a unipolar electronic device for converting an input positive voltage into a stabilized output voltage of the same polarity, which can be set at voltages both below and above the input voltage. The aim of this work is to analyze Zeta converter circuitry, which requires the following tasks to be solved: using Kirchhoff's Circuit Laws, obtain systems of equations describing converter operation in the phase of energy accumulation and in the phase of energy transfer; using a method proposed by A.I. Korshunov, combine the resulting systems of equations into a marginal continuous mathematical model; using expressions describing constant components of currents and voltages in Zeta converter, analyze their ripples and obtain equations for their calculation; compare the current and voltage values obtained from the continuous limiting mathematical model with the Zeta simulation results.

Methods. The tasks are solved using Kirchhoff's rules and the method for obtaining the limiting continuous mathematical model proposed by A.I. Korshunov. The results are analyzed using a circuit modelling in NI Multisim.

Results. It is shown that the phase coordinates of the mathematical model tend to real values of converter currents and voltages at a switching frequency of the power switch of more than 200 kHz. A strong correspondence was established between the calculated ripple values and their values obtained in the simulation (when changing the duty factor).

Conclusions. Mathematical models comprise the basis of unified calculation methods for any radio electronic circuit. The developed limiting continuous mathematical model allows a range of changes in current flowing through the choke windings and voltages on capacitor plates to be evaluated, including their maximum and minimum values for various converter parameters, such as power switch switching frequency, duty factor, element ratings, etc. Obtaining this information in turn enables the rational selection of the electronic component base of the converter.

Keywords: DC/DC converter, step-up and step-down converter, equivalent circuit, Zeta topology, converter, limiting continuous mathematical model, Kirchhoff's rules, ripple spreading

• Submitted: 28.12.2022 • Revised: 10.02.2023 • Accepted: 05.05.2023

For citation: Bitukov V.K., Lavrenov A.I., Malitskiy D.A. Analysis of the DC/DC Zeta topology converter ripples by applying its limiting continuous mathematical model. *Russ. Technol. J.* 2023;11(4):36–48. <https://doi.org/10.32362/2500-316X-2023-11-4-36-48>

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

DC/DC-преобразователи широко используются в автономных устройствах с батарейным питанием, таких как беспилотные летательные аппараты, пирометры, кардиостимуляторы, автомобильная электроника, роботы и т.д. [1–8]. Наличие DC/DC-преобразователя в устройстве электропитания определяет как массогабаритные параметры и КПД, так и энергоэффективность всего радиоэлектронного средства (РЭС), а значит, и время его автономной работы без подзарядки. Сложность построения таких устройств электропитания обусловлена также тенденцией использования автономными устройствами стабилизированных низковольтных напряжений. Все это предопределяет необходимость в разработке специализированных методов, алгоритмов и средств проектирования DC/DC-преобразователей.

Достаточно подробно разработаны и исследованы DC/DC-преобразователи, которые реализуют базовые топологии повышающего, понижающего и поллярно-инвертирующего типов [9, 10], а комплексные DC/DC-преобразователи, т.е. преобразователи понижающе-повышающего типа, требуют разработки методов проектирования и дополнительного исследования [2, 4, 11–13].

Создание DC/DC-преобразователей, как и других РЭС, базируется на соответствующих математических моделях, которые являются основой для единого методического подхода к разработке, проектированию и исследованию устройств. Предельные непрерывные математические модели базовых DC/DC-преобразователей предложены в [9, 14–17]. А для понижающе-повышающих DC/DC-преобразователей имеется лишь математическая модель одного преобразователя, построенного по топологии Cuck [18, 19].

Математическая модель униполярных DC/DC-преобразователей, построенных по топологии Zeta, либо по топологии SEPIC отсутствует. Надо отметить, что преобразователи SEPIC, Zeta и Cuck идентичны с точки зрения электронной компонентной базы, но со схемотехнической точки зрения они существенно различаются.

1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Для каждой непрерывной математической модели, отображающей процессы реального устройства,

можно построить фазовую плоскость, которая представляет собой совокупность точек возможных ее состояний. Точка на фазовой плоскости отражает текущее состояние модели, а при ее движении – изменение состояния. След от движения изображающей точки называется фазовой траекторией, а сама точка – фазовой координатой. Непрерывность системы означает, что в любой момент времени можно установить состояние системы, т.е. значения фазовых координат.

Математическая модель устройства позволяет получить соотношение между фазовыми координатами, которые соответствуют реальным токам и напряжениям DC/DC-преобразователя. Известно, что совпадение фазовых координат и реальных значений токов и напряжений математической модели достигается при периоде T коммутации электронного ключа, стремящегося к нулю. Такие математические модели ключевых устройств принято называть предельными [14, 15].

Униполярный Zeta-преобразователь (рис. 1), впервые предложенный в [12], в составе своей схемы имеет два дросселя $L1$ и $L2$, два конденсатора $C1$ и $C2$, электронный ключ $VT1$, обычно реализуемый полевым транзистором, и блок управления, определяющий режим работы транзистора. Дроссели $L1$ и $L2$ выполняют функцию накопления и передачи энергии посредством электромагнитной индукции, а конденсатор $C1$ присутствует в схеме для разделения входа преобразователя от его выхода. Разделительный конденсатор $C1$ также иногда называют «летающим», т.к. он выполняет функции не только разделения, но также накопления и передачи энергии между секциями преобразователя [11, 12]. Остальные электрорадиоэлементы выполняют традиционные функции.

Для DC/DC-преобразователя, как и для большинства ключевых устройств, выделяют две фазы работы. Первая фаза определяет режим накопления энергии, а вторая фаза – режим передачи энергии. Для построения математической модели необходимо составить эквивалентные схемы каждой фазы преобразователя. На основе эквивалентных схем были записаны системы уравнений, описывающие две фазы работы.

В [20] представлена только математическая модель преобразователя с одинаковыми дросселями $L1$ и $L2$. Поэтому целесообразно построить

математическую модель DC/DC-преобразователя понижающе-повышающего типа, построенного по топологии Zeta, в общем виде.

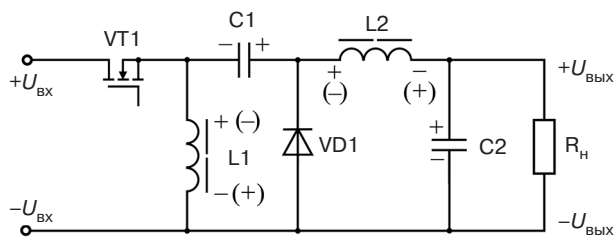


Рис. 1. Принципиальная электрическая схема преобразователя понижающе-повышающего типа, построенного по топологии Zeta. $U_{ВХ}$, $U_{ВЫХ}$ – входное и выходное напряжения, R_H – нагрузочный резистор. Здесь и далее на рисунках используются обозначения, принятые в ГОСТ 2.710-81¹

По правилам Кирхгофа были составлены системы уравнений, описывающие обе фазы его работы. DC/DC-преобразователь в режиме накопления энергии (первая фаза) описывается пятью уравнениями алгебраических сумм токов и напряжений:

$$i_{ВХ} - i_{L1} - i_{L2} = 0, \quad (1)$$

$$\left\{ \begin{array}{l} \frac{du_{C2}}{dt} = \frac{1}{C_2} i_{L2} - \frac{1}{R_H C_2} u_{C2}, \\ \frac{du_{C1}}{dt} = \frac{1}{C_1} i_{L2}, \\ \frac{di_{L1}}{dt} = \frac{1}{L_1} U_{ВХ} - \frac{r_1}{L_1} i_{L1}, \\ \frac{di_{L2}}{dt} = -\frac{1}{L_2} u_{C1} - \frac{1}{L_2} u_{C2} - \frac{r_2}{L_2} i_{L2} - \frac{1}{L_2} U_{ВХ}, \end{array} \right. \quad (2)$$

где L_1 – индуктивность дросселя L1; L_2 – индуктивность дросселя L2; C_1 – емкость конденсатора C1; C_2 – емкость конденсатора C2; R_H – сопротивление нагрузочного резистора R_H ; i_{L1} , i_{L2} – мгновенные токи, протекающие через обмотку дросселей L1 и L2; r_1 , r_2 – активные сопротивления обмоток дросселей L1 и L2; u_{C1} , u_{C2} – мгновенные напряжения на обкладках конденсаторов C1 и C2.

Уравнение (1), содержащее в себе входной ток $i_{ВХ}$, по правилам Кирхгофа необходимо для полного описания первой фазы работы преобразователя. Но учитывая,

что уравнения системы (2) не зависят от входного тока, уравнение (1) можно исключить из системы.

Система уравнений, определяющая работу преобразователя в режиме передачи энергии (вторая фаза), выглядит следующим образом:

$$\left\{ \begin{array}{l} \frac{du_{C2}}{dt} = \frac{1}{C_2} i_{L2} - \frac{1}{R_H C_2} u_{C2}, \\ \frac{du_{C1}}{dt} = \frac{1}{C_1} i_{L1}, \\ \frac{di_{L1}}{dt} = -\frac{r_1}{L_1} i_{L1} - \frac{1}{L_1} u_{C1}, \\ \frac{di_{L2}}{dt} = -\frac{r_2}{L_2} i_{L2} - \frac{1}{L_2} u_{C2}. \end{array} \right. \quad (3)$$

Для построения математической модели необходимо объединить системы уравнений (2)–(3) по методике [14, 15]. Математическая модель может быть записана в виде матричной системы уравнений

$$\mathbf{X} = \begin{bmatrix} i_{L1} \\ i_{L2} \\ u_{C1} \\ u_{C2} \end{bmatrix}, \quad (4)$$

$$\mathbf{A} = \begin{bmatrix} -\frac{r_1}{L_1} & 0 & -(1-D)\frac{1}{L_1} & 0 \\ 0 & -\frac{r_2}{L_2} & -D\frac{1}{L_2} & -\frac{1}{L_2} \\ (1-D)\frac{1}{C_1} & D\frac{1}{C_1} & 0 & 0 \\ 0 & \frac{1}{C_2} & 0 & -\frac{1}{RC_2} \end{bmatrix}, \quad (5)$$

$$\mathbf{B} = \begin{bmatrix} D\frac{1}{L_1} \\ -D\frac{1}{L_2} \\ 0 \\ 0 \end{bmatrix}, \quad (6)$$

где $\mathbf{X}$ – матрица фазовых координат системы; $\mathbf{A}$ – коэффициентная матрица фазовых координат; $\mathbf{B}$ – коэффициентная матрица внешнего источника, например входного напряжения; D – коэффициент заполнения широтно-импульсно модулированного сигнала, управляющего силовым ключом VT1.

¹ ГОСТ 2.710-81. Межгосударственный стандарт. Единая система конструкторской документации. Обозначения буквенно-цифровые в электрических схемах. М.: Стандартинформ; 2008. [GOST 2.710-81. Unified system for design documentation. Alpha-numerical designations in electrical diagrams. Moscow: Standartinform; 2008 (in Russ.).]

Таким образом, система уравнений (4)–(6) является предельной непрерывной математической моделью DC/DC-преобразователя, построенного по топологии Zeta.

Мгновенные токи и напряжения содержат постоянную и переменную (называемую пульсацией) составляющие. Анализ математической модели показал, что определение постоянных составляющих токов I_{L1} , I_{L2} , напряжений U_{C1} , U_{C2} и их пульсаций целесообразно выполнять раздельно.

Решив систему уравнений (4)–(6) математической модели для постоянных составляющих токов I_{L1} , I_{L2} и напряжений U_{C1} , U_{C2} , получим уравнения:

$$I_{L1} = \left| \frac{U_{BX} D^2}{R_H D^2 - (r_1 + 2R_H)D + R_H} \right|, \quad (7)$$

$$I_{L2} = \left| \frac{U_{BX} D(D-1)}{R_H D^2 - (r_1 + 2R_H)D + R_H} \right|, \quad (8)$$

$$U_{C1} = \left| \frac{R_H U_{BX} D(D-1)}{R_H D^2 - (r_1 + 2R_H)D + R_H} \right|, \quad (9)$$

$$U_{C2} = \left| -\frac{U_{BX} D((r_2 + r_1 + R_H)D - (r_2 + R_H))}{R_H D^2 - (r_1 + 2R_H)D + R_H} \right|. \quad (10)$$

Используя уравнения (1), (7) и (8), можно записать уравнение для постоянной составляющей входного тока I_{BX}

$$I_{BX} = \left(\left| \frac{U_{BX} D^2}{R_H D^2 - (r_1 + 2R_H)D + R_H} \right| + \left| \frac{U_{BX} D(D-1)}{R_H D^2 - (r_1 + 2R_H)D + R_H} \right| \right) D. \quad (11)$$

Уравнения (7)–(11) являются основой для предварительного расчета преобразователя, а значит и выбора его электронной компонентной базы.

2. ПУЛЬСАЦИИ ТОКОВ И НАПРЯЖЕНИЙ

Мгновенные токи i_{L1} и i_{L2} , протекающие через обмотки дросселей L1 и L2, и мгновенные напряжения u_{C1} и u_{C2} на конденсаторах C1 и C2 содержат постоянную и переменную составляющие

$$\begin{aligned} i_{L1} &= I_{L1} + \delta i_{L1}, & u_{C1} &= U_{C1} + \delta u_{C1}, \\ i_{L2} &= I_{L2} + \delta i_{L2}, & u_{C2} &= U_{C2} + \delta u_{C2}, \end{aligned} \quad (12)$$

где δi_{L1} – переменная составляющая тока i_{L1} ; δi_{L2} – переменная составляющая тока i_{L2} ; δu_{C1} – переменная составляющая напряжения u_{C1} ; δu_{C2} – переменная составляющая напряжения u_{C2} .

Подставив (12) в системы уравнений (2) и (3), описывающие обе фазы работы преобразователя, получим:

$$\left\{ \begin{aligned} \frac{d(I_{L1} + \delta i_{L1})}{dt} &= \frac{1}{L_1} U_{BX} - \frac{r_1}{L_1} (I_{L1} + \delta i_{L1}), & (13.1) \\ \frac{d(I_{L2} + \delta i_{L2})}{dt} &= -\frac{1}{L_2} (U_{C1} + \delta u_{C1}) - \frac{1}{L_2} (U_{C2} + \delta u_{C2}) - \frac{r_2}{L_2} (I_{L2} + \delta i_{L2}) - \frac{1}{L_2} U_{BX}, & (13.2) \\ \frac{d(U_{C1} + \delta u_{C1})}{dt} &= \frac{1}{C_1} (I_{L2} + \delta i_{L2}), & (13.3) \\ \frac{d(U_{C2} + \delta u_{C2})}{dt} &= \frac{1}{C_2} (I_{L2} + \delta i_{L2}) - \frac{1}{R_H C_2} (U_{C2} + \delta u_{C2}). & (13.4) \end{aligned} \right. \quad (13)$$

$$\left\{ \begin{aligned} \frac{d(I_{L1} + \delta i_{L1})}{dt} &= -\frac{r_1}{L_1} (I_{L1} + \delta i_{L1}) - \frac{1}{L_1} (U_{C1} + \delta u_{C1}), & (14.1) \\ \frac{d(I_{L2} + \delta i_{L2})}{dt} &= -\frac{r_2}{L_2} (I_{L2} + \delta i_{L2}) - \frac{1}{L_2} (U_{C2} + \delta u_{C2}), & (14.2) \\ \frac{d(U_{C1} + \delta u_{C1})}{dt} &= \frac{1}{C_1} (I_{L1} + \delta i_{L1}), & (14.3) \\ \frac{d(U_{C2} + \delta u_{C2})}{dt} &= \frac{1}{C_2} (I_{L2} + \delta i_{L2}) - \frac{1}{R_H C_2} (U_{C2} + \delta u_{C2}). & (14.4) \end{aligned} \right. \quad (14)$$

Учитывая, что $\frac{U_{C2}}{R_H} \approx I_{L2}$, уравнения (13.4) и (14.4) можно упростить:

$$\frac{d(U_{C2} + \delta u_{C2})}{dt} \approx \frac{1}{C_2} \delta i_{L2} - \frac{1}{R_H C_2} \delta u_{C2}.$$

Постоянные составляющие токов и напряжений, как правило, много больше пульсаций, поэтому в правых частях уравнений (13.1)–(13.3) и (14.1)–(14.3) пульсациями можно пренебречь по сравнению с соответствующими постоянными составляющими:

$$\left\{ \begin{array}{l} \frac{d\delta i_{L1}}{dt} = \frac{1}{L_1} U_{BX} - \frac{r_1}{L_1} I_{L1}, \quad (15.1) \\ \frac{d\delta i_{L2}}{dt} = -\frac{1}{L_2} U_{C1} - \frac{1}{L_2} U_{C2} - \frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{BX}, \quad (15.2) \\ \frac{d\delta u_{C1}}{dt} = \frac{1}{C_1} I_{L2}, \quad (15.3) \\ \frac{d\delta u_{C2}}{dt} = \frac{1}{C_2} \delta i_{L2} - \frac{1}{R_H C_2} \delta u_{C2}. \quad (15.4) \end{array} \right. \quad (15)$$

$$\left\{ \begin{array}{l} \frac{d\delta i_{L1}}{dt} = -\frac{r_1}{L_1} I_{L1} - \frac{1}{L_1} U_{C1}, \quad (16.1) \\ \frac{d\delta i_{L2}}{dt} = -\frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{C2}, \quad (16.2) \\ \frac{d\delta u_{C1}}{dt} = \frac{1}{C_1} I_{L1}, \quad (16.3) \\ \frac{d\delta u_{C2}}{dt} = \frac{1}{C_2} \delta i_{L2} - \frac{1}{R_H C_2} \delta u_{C2}. \quad (16.4) \end{array} \right. \quad (16)$$

Используя уравнения (15.1)–(15.3), получаем формулы для переменной составляющей (пульсаций) δi_{L1} , δi_{L2} и δu_{C1} первой фазы:

$$\delta i_{L1} = \int \left(\frac{1}{L_1} U_{BX} - \frac{r_1}{L_1} I_{L1} \right) dt = \left(\frac{1}{L_1} U_{BX} - \frac{r_1}{L_1} I_{L1} \right) t, \quad (17)$$

$$\delta i_{L2} = \int \left(-\frac{1}{L_2} U_{C1} - \frac{1}{L_2} U_{C2} - \frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{BX} \right) dt = \left(-\frac{1}{L_2} U_{C1} - \frac{1}{L_2} U_{C2} - \frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{BX} \right) t, \quad (18)$$

$$\delta u_{C1} = \int \left(\frac{1}{C_1} I_{L2} \right) dt = \left(\frac{1}{C_1} I_{L2} \right) t. \quad (19)$$

Используя уравнения (16.1)–(16.3), получаем формулы для переменных составляющих δi_{L1} , δi_{L2} и δu_{C1} второй фазы:

$$\delta i_{L1} = \int \left(-\frac{r_1}{L_1} I_{L1} - \frac{1}{L_1} U_{C1} \right) dt = \left(-\frac{r_1}{L_1} I_{L1} - \frac{1}{L_1} U_{C1} \right) t, \quad (20)$$

$$\delta i_{L2} = \int \left(-\frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{C2} \right) dt = \left(-\frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{C2} \right) t, \quad (21)$$

$$\delta u_{C1} = \int \left(\frac{1}{C_1} I_{L1} \right) dt = \left(\frac{1}{C_1} I_{L1} \right) t. \quad (22)$$

Выражения (15.4) и (16.4) имеют одинаковый вид, поэтому представим их следующим образом:

$$\frac{d\delta u_{C2}}{dt} = y' = px - qy, \quad (23)$$

где $q = \frac{1}{R_H C_2}$, коэффициент p для первой фазы выражается формулой

$$p_1 = \frac{1}{C_2} \left(-\frac{1}{L_2} U_{C1} - \frac{1}{L_2} U_{C2} - \frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{BX} \right),$$

а для второй фазы:

$$p_2 = \frac{1}{C_2} \left(-\frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{C2} \right).$$

Решив дифференциальное уравнение (23), получаем выражение:

$$y = \left(\frac{p(qx - 1)e^{qx}}{q^2} + K \right) e^{-qx}$$

$$\text{или } \delta u_{C2} = -\frac{p}{q^2} + \frac{pt}{q} + Ke^{-qt}. \quad (24)$$

Для упрощения расчетов можно принять граничное условие $y(0) = 0$. Выражения констант для первой и второй фаз:

$$K_1 = \frac{p_1}{q^2} \text{ и } K_2 = \frac{p_2}{q^2}. \quad (25)$$

Используя выражения (17)–(19), получаем формулы для определения размаха пульсаций Δi_{L1} , Δi_{L2} и Δu_{C1} первой фазы:

$$\Delta i_{L1} = \delta i_{L1}(TD) - \delta i_{L1}(0) = \left(\frac{1}{L_1} U_{BX} - \frac{r_1}{L_1} I_{L1} \right) TD =$$

$$= - \frac{U_{BX} DT \left((r_1 - R_H) D^2 + (r_1 + 2R_H) D - R_H \right)}{L_1 (R_H D^2 - (r_1 + 2R_H) D + R_H)}, \quad (26)$$

$$\Delta i_{L2} = \delta i_{L2}(TD) - \delta i_{L2}(0) =$$

$$= \left(-\frac{1}{L_2} U_{C1} - \frac{1}{L_2} U_{C2} - \frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{BX} \right) TD =$$

$$= \frac{U_{BX} DT \left((r_1 - R_H) D^2 + (r_1 + 2R_H) D - R_H \right)}{L_2 (R_H D^2 - (r_1 + 2R_H) D + R_H)}, \quad (27)$$

$$\Delta u_{C1} = \delta u_{C1}(TD) - \delta u_{C1}(0) = \left(\frac{1}{C_1} I_{L2} \right) TD =$$

$$= \frac{U_{BX} D^2 T (D - 1)}{C_1 (R_H D^2 - (r_1 + 2R_H) D + R_H)}. \quad (28)$$

Используя выражения (20)–(22), получаем формулы для определения размаха пульсаций Δi_{L1} , Δi_{L2} и Δu_{C1} второй фазы:

$$\Delta i_{L1} = \delta i_{L1}(TD) - \delta i_{L1}(T) =$$

$$= \left(-\frac{r_1}{L_1} I_{L1} - \frac{1}{L_1} U_{C1} \right) T(1 - D) =$$

$$= - \frac{U_{BX} DT (1 - D) \left((r_1 + R_H) D - R_H \right)}{L_1 (R_H D^2 - (r_1 + 2R_H) D + R_H)}, \quad (29)$$

$$\Delta i_{L2} = \delta i_{L2}(TD) - \delta i_{L2}(T) =$$

$$= \left(-\frac{r_2}{L_2} I_{L2} - \frac{1}{L_2} U_{C2} \right) T(1 - D) =$$

$$= \frac{U_{BX} DT (1 - D) \left((r_1 + R_H) D - R_H \right)}{L_2 (R_H D^2 - (r_1 + 2R_H) D + R_H)}, \quad (30)$$

$$\Delta u_{C1} = \delta u_{C1}(TD) - \delta u_{C1}(T) =$$

$$= \left(\frac{1}{C_1} I_{L1} \right) T(D - 1) =$$

$$= \frac{U_{BX} D^2 T (D - 1)}{C_1 (R_H D^2 - (r_1 + 2R_H) D + R_H)}. \quad (31)$$

Учитывая, что формулы для определения размаха пульсаций токов Δi_{L1} и Δi_{L2} отличаются между собой в двух фазах и при этом являются равнозначными,

целесообразно определить средние размахи пульсации Δi_{L1cp} и Δi_{L2cp} токов, протекающих через обмотки дросселей:

$$\Delta i_{L1cp} =$$

$$= \left| \frac{U_{BX} DT \left((r_1 - R_H) D^2 + (r_1 + 2R_H) D - R_H \right)}{L_1 (R_H D^2 - (r_1 + 2R_H) D + R_H)} \right| +$$

$$+ \left| \frac{U_{BX} DT (1 - D) \left((r_1 + R_H) D - R_H \right)}{L_1 (R_H D^2 - (r_1 + 2R_H) D + R_H)} \right|, \quad (32)$$

$$\Delta i_{L2cp} =$$

$$= \left| \frac{U_{BX} DT \left((r_1 - R_H) D^2 + (r_1 + 2R_H) D - R_H \right)}{L_2 (R_H D^2 - (r_1 + 2R_H) D + R_H)} \right| +$$

$$+ \left| \frac{U_{BX} DT (1 - D) \left((r_1 + R_H) D - R_H \right)}{L_2 (R_H D^2 - (r_1 + 2R_H) D + R_H)} \right|. \quad (33)$$

Используя выражения (24), получаем формулу для определения размаха пульсаций Δu_{C2}

$$\Delta u_{C2} = \left| \delta u_{C21} \left(\frac{TD}{2} \right) \right| + \left| \delta u_{C22} \left(T \frac{1-D}{2} \right) \right| =$$

$$= -\frac{p_1}{q^2} + \frac{p_1 \frac{TD}{2}}{q^2} + K_1 e^{-q \frac{TD}{2}} -$$

$$- \frac{p_2}{q^2} + \frac{p_2 T \frac{1-D}{2}}{q^2} + K_2 e^{-q T \frac{1-D}{2}} =$$

$$= \frac{1}{2} \left(\left| \frac{R_H U_{BX} D \left((r_1 + R_H) D - R_H \right)}{L_2 (R_H D^2 - (r_1 + 2R_H) D + R_H)} \times \right. \right. \quad (34)$$

$$\times \left. \left(2R_H C_2 e^{\frac{T(D-1)}{2R_H C_2}} + ((1-D)T - 2R_H C_2) \right) \right| +$$

$$+ \left| \frac{R_H U_{BX} \left((r_1 - R_H) D^2 + (r_1 + 2R_H) D - R_H \right)}{L_2 (R_H D^2 - (r_1 + 2R_H) D + R_H)} \times \right.$$

$$\times \left. \left(2R_H C_2 e^{\frac{-TD}{2R_H C_2}} + (DT - 2R_H C_2) \right) \right| \Bigg|,$$

где δu_{C21} – функция переменной составляющей пульсации напряжения на конденсаторе C2,

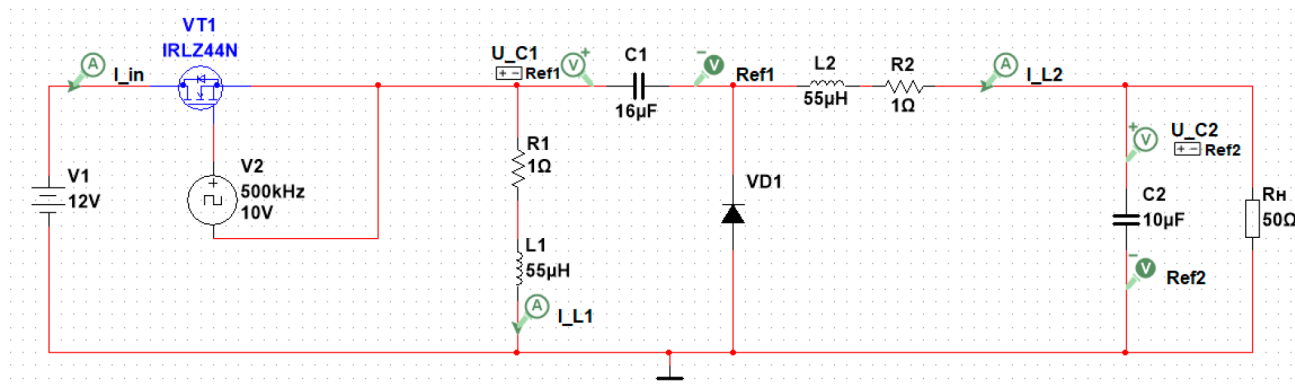


Рис. 2. Схема моделирования DC/DC-преобразователя

полученная с помощью решения дифференциального уравнения в первой фазе; δu_{C22} – функция переменной составляющей пульсации напряжения на конденсаторе C2, полученная с помощью решения дифференциального уравнения во второй фазе.

Используя уравнения (1), (32) и (33), можно записать формулу определения размаха пульсации входного тока $\Delta i_{\text{вх}}$

$$\begin{aligned} \Delta i_{\text{вх}} &= \frac{\Delta i_{L1\text{cp}} + \Delta i_{L2\text{cp}}}{2} + \frac{1}{D} I_{\text{вх}} = \\ &= \frac{1}{4} \left(\left| -\frac{U_{\text{вх}} D T ((r_1 - R_{\text{H}}) D^2 + (r_1 + 2R_{\text{H}}) D - R_{\text{H}})}{L_1 (R_{\text{H}} D^2 - (r_1 + 2R_{\text{H}}) D + R_{\text{H}})} \right| + \right. \\ &\quad \left. + \left| -\frac{U_{\text{вх}} D T (1-D) ((r_1 + R_{\text{H}}) D - R_{\text{H}})}{L_1 (R_{\text{H}} D^2 - (r_1 + 2R_{\text{H}}) D + R_{\text{H}})} \right| + \right. \\ &\quad \left. + \left| \frac{U_{\text{вх}} D T ((r_1 - R_{\text{H}}) D^2 + (r_1 + 2R_{\text{H}}) D - R_{\text{H}})}{L_2 (R_{\text{H}} D^2 - (r_1 + 2R_{\text{H}}) D + R_{\text{H}})} \right| + \right. \\ &\quad \left. + \left| \frac{U_{\text{вх}} D T (1-D) ((r_1 + R_{\text{H}}) D - R_{\text{H}})}{L_2 (R_{\text{H}} D^2 - (r_1 + 2R_{\text{H}}) D + R_{\text{H}})} \right| \right) + \\ &\quad + \left(\left| \frac{U_{\text{вх}} \cdot D^2}{R_{\text{H}} \cdot D^2 - (r_1 + 2 \cdot R_{\text{H}}) \cdot D + R_{\text{H}}} \right| + \right. \\ &\quad \left. + \left| \frac{U_{\text{вх}} \cdot D \cdot (D-1)}{R_{\text{H}} \cdot D^2 - (r_1 + 2 \cdot R_{\text{H}}) \cdot D + R_{\text{H}}} \right| \right). \end{aligned} \quad (35)$$

Таким образом, формулы (28) или (31) и (32)–(35) позволяют рассчитать размах пульсации токов i_{L1} , i_{L2} , протекающих через обмотки дросселей L1 и L2, входного тока $i_{\text{вх}}$ и напряжений u_{C1} , u_{C2} на обкладках конденсаторов C1 и C2, зная номиналы выбранных элементов электронной компонентной базы и режим работы преобразователя (коэффициент заполнения D и период T).

3. МОДЕЛИРОВАНИЕ В СРЕДЕ MULTISIM

Для проверки достоверности полученных выражений для определения пульсаций токов и напряжений DC/DC-преобразователя, построенного по Zeta-топологии, использовалось схемотехническое моделирование в среде *Multisim*². Вывод аналитических формул и результаты моделирования постоянных составляющих токов и напряжений рассматриваемого преобразователя представлены в [20–22].

Моделированию предшествовали исследования силовых ключей MOSFET, как рекомендовано в [23, 24], в статическом и динамическом режиме. На этом основании был выбран транзистор IRLZ44N, характеристики модели которого соответствуют данным Datasheet³.

Схема моделирования представлена на рис. 2. Ключ VT1 коммутировал ток входного источника питания V1 с частотой тактовых импульсов, задаваемых генератором V2. В качестве элементов были выбраны компоненты из базы данных *Multisim*. Дроссели представлены эквивалентными схемами. Активное сопротивление дросселя, имеющего индуктивность 55 мкГн, не превышало 1 Ом.

Для исследования влияния коэффициента заполнения D на пульсации было выполнено моделирование схемы в режиме анализа переходных процессов. При этом пульсации токов и напряжений регистрировались в установившемся режиме через 5–12 мс после начала моделирования. Результаты исследования влияния коэффициента заполнения D как основного параметра, определяющего режим работы преобразователя, представлены на рис. 3–5. На рис. 6–8 показано влияние частоты коммутации f на величину размаха пульсаций.

² <https://www.ni.com/ru-ru.html>. Дата обращения 27.03.2023. / Accessed March 27, 2023 (in Russ.).

³ International rectifier, IRFZ44N HEXFET Power MOSFET, Data Sheet. <https://static.chipdip.ru/lib/158/DOC000158617.pdf>. Дата обращения 27.03.2023. / Accessed March 27, 2023.

Исследования влияния коэффициента заполнения D на величину размаха пульсаций показывают хорошее совпадение результатов расчета по математической модели и моделированию. Однако при коэффициентах заполнения D меньше 0.3 и больше 0.7 наблюдается существенное различие. Это расхождение обусловлено неидеальностью математической модели и влиянием паразитных параметров электрорадиоэлементов на работу преобразователя. Совпадение рассчитанных размахов пульсаций Δi_{BX} , Δi_{Lcp} , Δu_{C} и величин Δi_{BXM} , Δi_{LM} , Δu_{CM} , полученных при моделировании, наблюдалось при коэффициенте заполнения D равном 0.5. Здесь Δi_{BXM} – размах пульсаций входного тока, полученный при помощи моделирования; Δi_{LM} – размах пульсаций тока, протекающего через обмотку дросселей L1 и L2, полученный при помощи моделирования; Δu_{CM} – размах пульсации напряжения на конденсаторах C1 и C2, полученный при помощи моделирования. Разница между расчетными значениями и результатами моделирования для пульсаций входного тока составляла 6 мА при $\Delta i_{\text{BX}} = 718$ мА. Разница между расчетными значениями и результатами моделирования для пульсаций токов i_{L1} , i_{L2} составляла 7 мА при $\Delta i_{\text{Lcp}} = 218$ мА. Для размаха пульсаций напряжения Δu_{C1} разница между расчетными значениями и результатами моделирования составила 0.3 мВ при $\Delta u_{\text{C1}} = 15$ мВ и для пульсаций Δu_{C2} составила 0.05 мВ при $\Delta u_{\text{C2}} = 5.45$ мВ.

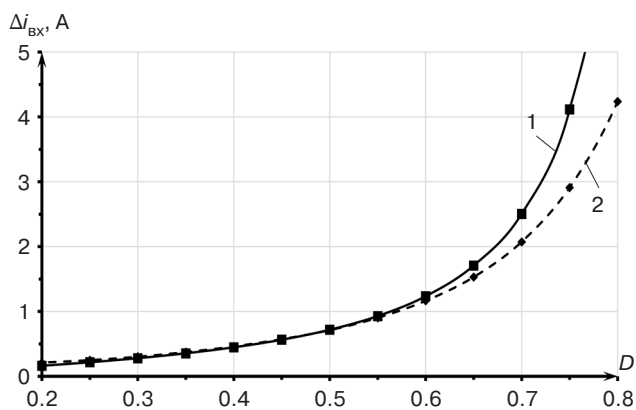


Рис. 3. Влияние коэффициента заполнения на пульсации входного тока: 1 – расчетное значение Δi_{BX} ; 2 – результат моделирования Δi_{BXM}

Из рис. 3 видно, что расчетные значения пульсаций входного тока соответствуют результатам моделирования почти во всем диапазоне изменения коэффициента заполнения D . Однако результаты расчета и моделирования различаются при коэффициенте заполнения $D < 0.25$ и $D > 0.75$. Особенно это проявляется при коэффициенте заполнения D , превышающем 0.75. Например, при коэффициенте заполнения $D = 0.8$ разница расчетных и пульсаций, полученных при моделировании, составляла ~4 А

при $\Delta i_{\text{BX}} = 4.3$ А, а при коэффициенте заполнения $D = 0.2$ эта разница составляла ~48 мА при $\Delta i_{\text{BX}} = 163$ мА.

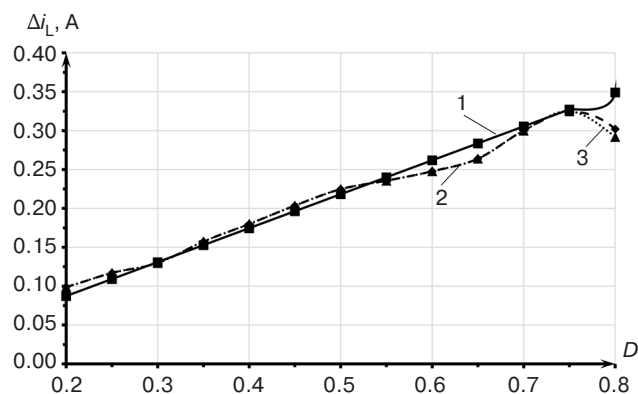


Рис. 4. Влияние коэффициента заполнения на пульсации токов, протекающих через обмотки дросселей L1 и L2: 1 – расчетное значение Δi_{L} ; 2 – результат моделирования Δi_{L1M} ; 3 – результат моделирования Δi_{L2M}

Из рис. 4 видно, что пульсации токов, протекающих через обмотки дросселей, практически совпадают. Результаты расчета и моделирования, так же как и в случае входного тока, систематически различаются при коэффициенте заполнения $D < 0.25$ и $D > 0.75$. При коэффициенте заполнения $D = 0.8$ разница расчетных и значений пульсаций, полученных при помощи моделирования, составляет ~47 мА при $\Delta i_{\text{L}} = 349$ мА, а при коэффициенте заполнения $D = 0.2$ эта разница составляет ~13 мА при $\Delta i_{\text{L}} = 87$ мА.

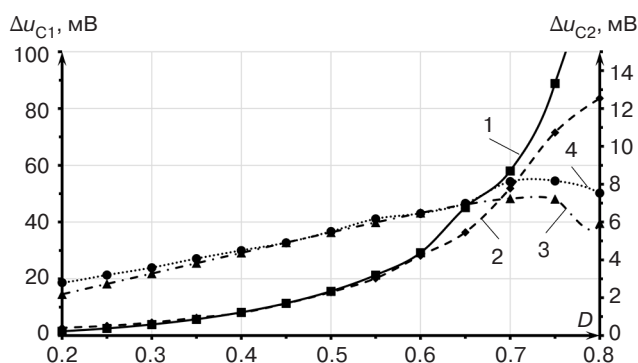


Рис. 5. Влияние коэффициента заполнения на пульсации напряжения на конденсаторах C1 и C2: 1 – расчетное значение Δu_{C1} ; 2 – результат моделирования Δu_{C1M} ; 3 – расчетное значение Δu_{C2} ; 4 – результат моделирования Δu_{C2M}

Из рис. 5 следует, что при коэффициенте заполнения 0.5 расчетные пульсации и результаты моделирования совпадают. Однако пульсации напряжения на конденсаторе C2 различаются на 1.59 мВ при 6 мВ, если $D = 0.8$, и на 0.6 мВ при $\Delta u_{\text{C2}} = 2.2$ мВ, если $D = 0.2$.

Разница расчетного значения напряжения Δu_{C1} на конденсаторе C1 и результата моделирования Δu_{C1M} достигает 76 мВ при $\Delta u_{C1} = 160$ мВ, если $D \sim 0.8$, однако при $D \sim 0.2$ различие результатов расчета и моделирования незначительное.

При повышении частоты коммутации силового транзистора пульсации токов и напряжений значительно уменьшаются. Как показано на рис. 6, расчетные значения пульсации Δi_{BX} и результаты моделирования Δi_{BXM} совпадают при частотах коммутации от 50 до 800 кГц. Максимальная разница расчетных пульсаций и моделирования в рабочей полосе частот наблюдается при частоте 50 кГц (в увеличенном масштабе различие показано на вставке рис. 7) и равно 0.35 А при $\Delta i_{BX} = 4.2$ А.

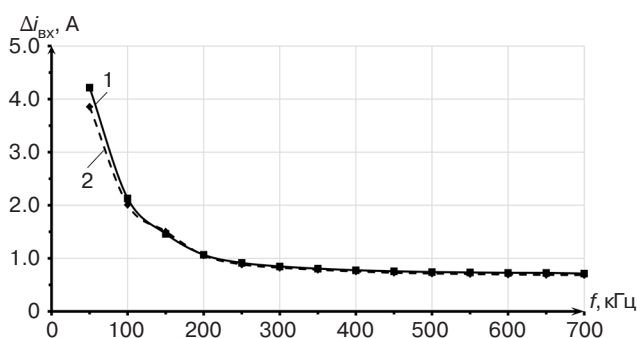


Рис. 6. Влияние частоты коммутации на пульсацию входного тока при коэффициенте заполнения, равном 0.5: 1 – расчетное значение Δi_{BX} ; 2 – результат моделирования Δi_{BXM}

На рис. 7 показано, что расчетные пульсации Δi_L и пульсации при моделировании Δi_{LM} так же, как и входной ток, совпадают при частотах коммутации от 50 до 800 кГц. Максимальная разница расчетных пульсаций и пульсаций при моделировании в рабочей полосе частот наблюдается при частоте 50 кГц (в увеличенном масштабе различие показано на вставке рис. 7) и равно 0.26 А для Δi_{L1} и 0.2 А для Δi_{L2} при $\Delta i_L = 2.18$ А.

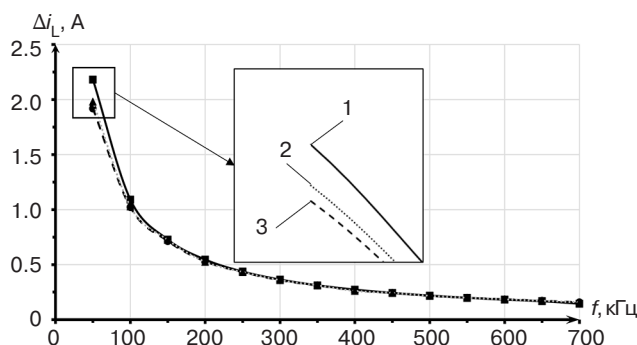


Рис. 7. Влияние частоты коммутации на пульсации токов, протекающих через обмотку дросселей L1 и L2 при коэффициенте заполнения, равном 0.5: 1 – расчетное значение Δi_L ; 2 – результат моделирования Δi_{LM} ; 3 – результат моделирования Δi_{L1M}

На рис. 8, в отличие от рис. 7, явно видна предельность непрерывной математической модели. Например, при частоте 100 кГц расчетное значение Δu_{C1} равно 78 мВ, а моделирование Δu_{C1} равно 136 мВ. Значения фазовых координат предельной непрерывной математической модели стремятся к значениям токов и напряжений преобразователя при частотах более 200 кГц. Из рис. 8 видно, что расчетные пульсации и пульсации при моделировании совпадают на частотах более 200 кГц. Максимальное различие наблюдается при частоте ~ 230 кГц и равно ~ 3 мВ для Δu_{C1} при $\Delta u_{C1} = 31$ мВ. Для пульсаций Δu_{C2} различие составляет ~ 0.9 мВ при $\Delta u_{C2} = 21.8$ мВ.

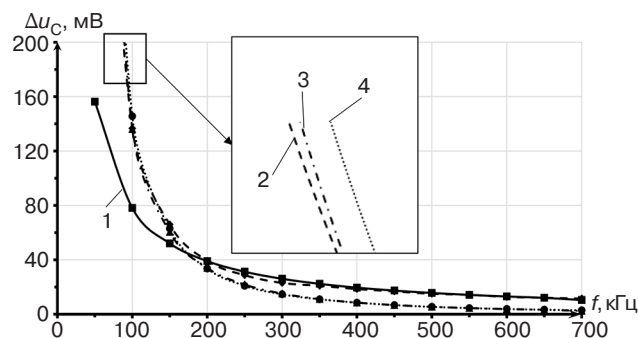


Рис. 8. Влияние частоты коммутации на напряжение на конденсаторах C1, C2 при коэффициенте заполнения, равном 0.5: 1 – расчетное значение Δu_{C1} ; 2 – результат моделирования Δu_{C1M} ; 3 – расчетное значение Δu_{C2} ; 4 – результат моделирования Δu_{C2M}

ЗАКЛЮЧЕНИЕ

Составлены эквивалентные схемы DC/DC-преобразователя, построенного по топологии Zeta, в режимах накопления и передачи энергии. С использованием правил Кирхгофа составлены системы уравнений, описывающие каждую фазу работы устройства. Для формулировки математической модели в матричном виде системы уравнений были преобразованы в матрицу фазовых координат, коэффициентную матрицу фазовых координат и коэффициентную матрицу внешнего источника. Для полного анализа предельной непрерывной математической модели записаны уравнения для постоянных составляющих.

С помощью представления токов и напряжений в качестве суммы постоянной и переменной составляющих записаны системы уравнений, описывающие преобразователь в режимах накопления и передачи энергии. Были использованы полученные выражения постоянных составляющих токов и напряжений, проинтегрированы системы уравнений, описывающие каждую фазу, сформулированы выражения для определения пульсаций токов,

протекающих через обмотку дросселей, и напряжений на конденсаторах.

Выполнено сопоставление результатов расчетов пульсаций при использовании предельной непрерывной математической модели и моделировании преобразователя. Получены зависимости пульсаций от коэффициента заполнения D и частоты переключения силового ключа f . При коэффициенте заполнения $D = 0.5$ пульсации, найденные с помощью математической модели, совпадают с полученными при моделировании. Разница составляет ~ 7 мА при $\Delta i_L = 220$ мА для пульсаций токов Δi_{L1} , Δi_{L2} . Разница между пульсациями напряжений равна ~ 0.3 мВ при $\Delta u_{C1} = 15$ мВ для пульсаций Δu_{C1} и ~ 0.05 мВ при $\Delta u_{C2} = 5.5$ мВ для пульсаций Δu_{C2} . Максимальные отклонения расчетных значений от результатов моделирования при коэффициенте заполнения $D = 0.8$

составляют: 47 мА для Δi_{L1} и Δi_{L2} , 76.4 мВ для Δu_{C1} и 1.59 мВ для Δu_{C2} .

Созданная предельная непрерывная математическая модель позволяет оценить диапазон изменения токов, протекающих через обмотки дросселей, и напряжений на обкладках конденсаторов, их максимальные и минимальные значения при различных параметрах преобразователя, таких как частота коммутации силового ключа, коэффициент заполнения, номиналы элементов и т.д. Полученные зависимости позволяют выполнить рациональный подбор электронной компонентной базы.

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

1. Битюков В.К., Иванов А.А., Миронов А.В., Михневич Н.Г., Перфильев В.С., Петров В.А. Стенд для исследования характеристик микросхем источников вторичного электропитания с накачкой заряда. *Russ. Technol. J.* 2016;4(3):37–52. <https://doi.org/10.32362/2500-316X-2016-4-3-37-52>
2. Захаров Л.Ф. Импульсный стабилизатор напряжения с широкими пределами изменения входного напряжения. *Электросвязь*. 2018;11:81–82.
3. Гарелина С.А., Латышенко К.П., Фрунзе А.В., Горбунов Р.А. Практическая реализация пирометров для измерения температуры пламени и объектов сквозь пламя. *Научные и образовательные проблемы гражданской защиты*. 2020;2(45):110–115.
4. Одинокоев А.О., Кремзуков Ю.А. Выбор топологии преобразователя постоянного напряжения. SEPIC или Zeta. *Практическая силовая электроника*. 2022;4(88):44–47.
5. Васюков И.В., Павленко А.В., Батищев Д.В. Обзор и анализ топологий преобразователей систем электропитания на водородных топливных элементах для беспилотных летательных аппаратов киловаттного класса мощности. *Известия ВУЗов. Электромеханика*. 2022;65(2):19–26. <https://doi.org/10.17213/0136-3360-2022-2-19-26>
6. Чоудхари В. Решения устойчивого к переходным процессам первичного DC/DC-преобразователя системы питания автомобильной электроники. *Силовая электроника*. 2017;3(66):30–34. URL: <https://power-e.ru/wp-content/uploads/6630.pdf>
7. Казанцев Д.П., Щербак В.Ф., Закамалдин Д.А., Казанцев Ю.Е., Пронин А.В., Молодых С.В. *Опτικο-электронный аппарат*: пат. 2531695 РФ. Заявка № 2012156979/14; заявл. 24.12.2012; опубл. 27.10.2014. Бюл. № 30.

REFERENCES

1. Bitukov V.K., Ivanov A.A., Mironov A.V., Mikhnevich N.G., Perfiliev V.S., Petrov V.A. Test bench for studying characteristics integrated circuit chips of secondary regulated charge pump power supply. *Russ. Technol. J.* 2016;4(3):37–52 (in Russ.). <https://doi.org/10.32362/2500-316X-2016-4-3-37-52>
2. Zakharov L.F. Pulsed voltage regulator with wide input voltage change limits. *Elektrosvyaz = The Elektrosvyaz Magazine*. 2018;11:81–82 (in Russ.).
3. Garelina S.A., Latyshenko K.P., Frunze A.V., Gorbunov R.A. Practical implementation of pyrometers for measuring flame temperature and objects through a flame. *Nauchnye i obrazovatel'nye problemy grazhdanskoi zashchity = Scientific and Educational Tasks of Civil Defence*. 2020;2(45):110–115 (in Russ.).
4. Odinokov A.O., Kremzucov Yu.A. Selecting DC converter topology. SEPIC or Zeta. *Prakticheskaya silovaya elektronika = Practical Power Electronics*. 2022;4(88):44–47 (in Russ.).
5. Vasyukov I.V., Pavlenko A.V., Batishchev D.V. Review and analysis of topologies of converters of power supply systems on hydrogen fuel cells for unmanned aerial vehicles of kilowatt power class. *Izvestiya vysshikh uchebnykh zavedenii. Elektromekhanika = Russian Electromechanics*. 2022;65(2):19–26 (in Russ.). <https://doi.org/10.17213/0136-3360-2022-2-19-26>
6. Choudhary V. Transient resistant primary DC/DC converter solutions for automotive electronics power systems. *Silovaya elektronika = Power Electronics*. 2017;3(66):30–34 (in Russ.). Available from URL: <https://power-e.ru/wp-content/uploads/6630.pdf>
7. Kazantsev D.P., Shcherbak V.F., Zakamaldin D.A., Kazantsev Yu.E., Pronin A.V., Molodykh S.V. *Electric Cardiac Pacemaker*: RF Pat. 2531695. Publ. 27.10.2014 (in Russ.).

8. Кастров М.Ю., Макаров В.В. Основы разработки преобразователя Zeta на базе контроллеров SIPEX SP16125/6/7. *Практическая силовая электроника*. 2010;2(38):15–18.
9. Коршунов А.И. Методика построения непрерывных моделей импульсных преобразователей напряжения постоянного тока. *Компоненты и технологии*. 2006;8(61):124–130. URL: <https://kit-e.ru/powerel/metodika-postroeniya-nepreryvnyh-modelej-impulsnyh-preobrazovatelej-napryazheniya-postoyannogo-toka/>
10. Битюков В.К., Симачков Д.С., Бабенко В.П. *Источники вторичного электропитания*. М.: Инфра-Инженерия; 2020. 376 с.
11. Erickson R.W., Maksimović D. *Fundamentals of Power Electronics*. NY: Springer; 2020. 1084 p. <https://doi.org/10.1007/978-3-030-43881-4>
12. Jozwik J.J., Kazimierczuk M.K. Dual Sepic PWM Switching-Mode DC/DC Power Converter. *IEEE Transactions on Industrial Electronics*. 1989;36(1): 64–70. <https://doi.org/10.1109/41.20346>
13. Пономарев Ю.Г., Присмотров Н.И., Шураков И.А. Вентильный электропривод с высоким коэффициентом мощности. *Интеллектуальная электротехника*. 2022;1(17):27–41. https://doi.org/10.46960/2658-6754_2022_1_27
14. Коршунов А.И. Предельная непрерывная модель системы с высокочастотным периодическим изменением структуры. *Известия ВУЗов. Приборостроение*. 2009;52(9):42–48. URL: <https://pribor.ifmo.ru/file/article/4902.pdf>
15. Коршунов А.И. Предельная непрерывная модель системы с периодическим высокочастотным изменением структуры. *Силовая электроника*. 2021;5(92):48–51.
16. Коршунов А.И. Повышение качества стабилизации выходного напряжения импульсного преобразователя постоянного тока. *Известия ВУЗов. Приборостроение*. 2013;56(3):48–57. URL: <https://pribor.ifmo.ru/file/article/6176.pdf>
17. Коршунов А.И. Два подхода к анализу устойчивости стабилизаторов напряжения постоянного тока с переменной структурой силовой части. *Практическая силовая электроника*. 2017;66:12–19.
18. Коршунов А.И. Импульсный преобразователь напряжения постоянного тока по схеме Чука. *Силовая электроника*. 2017;4(67):60–66.
19. Коршунов А.И. Особенности стабилизации напряжения постоянного тока с помощью преобразователя Чука. *Практическая силовая электроника*. 2017;4(68):2–9.
20. Битюков В.К., Лавренов А.И. Математическая модель DC/DC преобразователя напряжения, построенного по Zeta топологии. *Фундаментальные, поисковые, прикладные исследования и инновационные проекты: сборник трудов Национальной научно-практической конференции*; под ред. С.У. Увайсова. М.: РТУ МИРЭА; 2022. С. 209–215.
21. Битюков В.К., Лавренов А.И., Малицкий Д.А. Математическая модель DC/DC преобразователя, построенного по Zeta топологии (часть 1). *Проектирование и технология электронных средств*. 2022;4:53–57.
8. Kastrov M.Yu., Makarov V.V. Zeta converter based on SIPEX SP16125/6/7 controllers design baseline. *Prakticheskaya silovaya elektronika = Practical Power Electronics*. 2010;2(38):15–18 (in Russ.).
9. Korshunov A.I. Methodology for constructing continuous models of pulsed DC voltage converters. *Komponenty i tekhnologii = Components & Technologies*. 2006;8(61):124–130 (in Russ.). Available from URL: <https://kit-e.ru/powerel/metodika-postroeniya-nepreryvnyh-modelej-impulsnyh-preobrazovatelej-napryazheniya-postoyannogo-toka/>
10. Bityukov V.K., Simachkov D.S., Babenko V.P. *Istochniki vtorichnogo elektropitaniya (Secondary Power Sources)*. Moscow: Infra-Inzheneriya; 2020. 376 p. (in Russ.).
11. Erickson R.W., Maksimović D. *Fundamentals of Power Electronics*. NY: Springer; 2020. 1084 p. <https://doi.org/10.1007/978-3-030-43881-4>
12. Jozwik J.J., Kazimierczuk M.K. Dual Sepic PWM Switching-Mode DC/DC Power Converter. *IEEE Transactions on Industrial Electronics*. 1989;36(1): 64–70. <https://doi.org/10.1109/41.20346>
13. Ponomarev Yu.G., Prismotrov N.I., Shurakov I.A. BLDC and PMBLAC motor drives with high power factor. *Intellectual'naya elektrotehnika = Smart Electrical Engineering*. 2022;1(17):27–41 (in Russ.). https://doi.org/10.46960/2658-6754_2022_1_27
14. Korshunov A.I. Limiting continuous model of a system with high-frequency structure variation. *Izvestiya vysshikh uchebnykh zavedenii. Priborostroenie = J. Instrument Eng.* 2009;52(9):42–48 (in Russ.). Available from URL: <https://pribor.ifmo.ru/file/article/4902.pdf>
15. Korshunov A.I. Limiting continuous model of a system with periodic high-frequency structure variation. *Silovaya elektronika = Power Electronics*. 2021;5(92):48–51 (in Russ.).
16. Korshunov A.I. Improving the quality of output voltage stabilisation of a pulsed DC converter. *Izvestiya vysshikh uchebnykh zavedenii. Priborostroenie = J. Instrument Eng.* 2013;56(3):48–57 (in Russ.). Available from URL: <https://pribor.ifmo.ru/file/article/6176.pdf>
17. Korshunov A.I. Two approaches to stability analysis of DC-DC converters with variable structure of power stage. *Prakticheskaya silovaya elektronika = Practical Power Electronics*. 2017;2(66):12–19 (in Russ.).
18. Korshunov A.I. Pulsed DC voltage converter based on Chuk scheme. *Silovaya elektronika = Power Electronics*. 2017;4(67):60–66 (in Russ.).
19. Korshunov A.I. Specifics of DC voltage stabilization by Chuk converter. *Prakticheskaya silovaya elektronika = Practical Power Electronics*. 2017;4(68):2–9 (in Russ.).
20. Bityukov V.K., Lavrenov A.I. Mathematical model of a DC/DC voltage converter based on Zeta topology. In: *Fundamental, Prospecting, Applied Research and Innovative Projects: Proceedings of the National Scientific and Practical Conference*. Moscow: RTU MIREA; 2022. P. 209–215 (in Russ.).
21. Bityukov V.K., Lavrenov A.I., Malitskiy D.A. Mathematical model of DC/DC converter based on Zeta topology (Part 1). *Proektirovanie i tekhnologiya elektronnykh sredstv = Design and Technology of Electronic Means*. 2022;4:53–57 (in Russ.).

22. Битюков В.К., Лавренов А.И., Малицкий Д.А. Математическая модель DC/DC преобразователя, построенного по Zeta топологии (часть 2). *Проектирование и технология электронных средств*. 2023;1:48–53.
23. Бабенко В.П., Битюков В.К. Имитационное моделирование процессов переключения силовых полевых транзисторов в программе Electronics Workbench. *Радиотехника и электроника*. 2019;64(2):199–205. <https://doi.org/10.1134/S0033849419020025>
24. Бабенко В.П., Битюков В.К., Кузнецов В.В., Симачков Д.С. Моделирование статических и динамических потерь в MOSFET ключах. *Russ. Technol. J.* 2018;6(1):20–39. <https://doi.org/10.32362/2500-316X-2018-6-1-20-39>
22. Bitukov V.K., Lavrenov A.I., Malitskiy D.A. Mathematical model of DC/DC converter based on Zeta topology (Part 2). *Proektirovanie i tekhnologiya elektronnykh sredstv = Design and Technology of Electronic Means*. 2023;1:48–53 (in Russ.).
23. Babenko V.P., Bitukov V.K. Simulation of Switching of High-Power FETs Using the Electronics Workbench Software. *J. Commun. Technol. Electron.* 2019;64(2): 176–181. <https://doi.org/10.1134/S1064226919020025> [Original Russian Text: Babenko V.P., Bitukov V.K. Simulation of Switching of High-Power FETs Using the Electronics Workbench Software. *Radiotekhnika i elektronika = J. Commun. Technol. Electron.* 2019;64(2):199–205 (in Russ.). <https://doi.org/10.1134/S0033849419020025>]
24. Babenko V.P., Bitukov V.K., Kuznetsov V.V., Simachkov D.S. Simulation of static and dynamic losses in MOSFET keys. *Russ. Technol. J.* 2018;6(1):20–39 (in Russ.). <https://doi.org/10.32362/2500-316X-2018-6-1-20-39>

Об авторах

Битюков Владимир Ксенофонович, д.т.н., профессор, кафедра радиоволновых процессов и технологий Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: bitukov@mirea.ru. ResearcherID Y-8325-2018, Scopus Author ID 6603797260, SPIN-код РИНЦ 3834-5360, <https://orcid.org/0000-0001-6448-8509>

Лавренов Алексей Игоревич, ассистент, кафедра радиоволновых процессов и технологий Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: lavrenov@mirea.ru. SPIN-код РИНЦ 6048-5027, <https://orcid.org/0000-0001-5722-541X>

Малицкий Даниил Александрович, ассистент, кафедра радиоволновых процессов и технологий Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: malickij@mirea.ru. SPIN-код РИНЦ 4912-3018, <https://orcid.org/0000-0003-4558-9085>

About the authors

Vladimir K. Bitukov, Dr. Sci. (Eng.), Professor, Department of Radio Wave Processes and Technology, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: bitukov@mirea.ru. ResearcherID Y-8325-2018, Scopus Author ID 6603797260, RSCI SPIN-code 3834-5360, <https://orcid.org/0000-0001-6448-8509>

Alexey I. Lavrenov, Assistant, Department of Radio Wave Processes and Technology, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: lavrenov@mirea.ru. SPIN-код РИНЦ 6048-5027, <https://orcid.org/0000-0001-5722-541X>

Daniil A. Malitskiy, Assistant, Department of Radio Wave Processes and Technology, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: malickij@mirea.ru. SPIN-код РИНЦ 4912-3018, <https://orcid.org/0000-0003-4558-9085>

УДК 621.391

<https://doi.org/10.32362/2500-316X-2023-11-4-49-58>



НАУЧНАЯ СТАТЬЯ

Многозадачная нейронная сеть в задаче распознавания вида QAM- и PSK-модуляции в условиях параметрической априорной неопределенности

А.А. Парамонов[@], В.М. Нгуен, М.Т. Нгуен

МИРЭА – Российский технологический университет, Москва, 119454 Россия

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

Резюме

Цели. Автоматическое распознавание видов модуляции неизвестных сигналов является важной задачей для различных областей техники: радиоконтроля и радиомониторинга, идентификации помех и источников радиоизлучения. Основная цель работы – разработка метода распознавания видов модуляции сигналов в условиях параметрической априорной неопределенности, в т.ч. неопределенности значений несущей частоты и начальной фазы сигнала. Дополнительной задачей является оценка значений отстроек от несущей частоты или фазы сигнала на начальном этапе процесса распознавания.

Методы. Использована многозадачная искусственная нейронная сеть, теория кумулянтов случайных величин.

Результаты. Для сигналов со сдвигом несущей частоты и начальной фазы вычислены кумулянты для модуляции QAM-8, APSK-16, QAM-64 и PSK-8. Представлена использующая кумулянтные признаки и алгоритм стандартизации данных многозадачная нейронная сеть. Результаты эксперимента показали, что использование многозадачной нейронной сети обеспечивает высокую точность распознавания модуляции QAM-8 и APSK-16, QAM-64 и PSK-8 в случае небольших отстроек несущей частоты или начальной фазы. Точность определения значений отстройки несущей частоты или начальной фазы сигнала для модуляции QAM-8, APSK-16, QAM-64 и PSK-8 оказывается высокой.

Выводы. Многозадачная нейронная сеть, использующая кумулянты сигналов высокого порядка, позволяет не только распознавать с высокой точностью виды модуляции в условиях априорной неопределенности параметров сигналов, но определять при этом значения отстроек несущей частоты или начальной фазы сигнала от ожидаемых значений.

Ключевые слова: распознавание, нейронная сеть, несущая частота, начальная фаза, кумулянтный признак

• Поступила: 24.03.2023 • Доработана: 11.04.2023 • Принята к опубликованию: 02.05.2023

Для цитирования: Парамонов А.А., Нгуен В.М., Нгуен М.Т. Многозадачная нейронная сеть в задаче распознавания вида QAM- и PSK-модуляции в условиях параметрической априорной неопределенности. *Russ. Technol. J.* 2023;11(4):49–58. <https://doi.org/10.32362/2500-316X-2023-11-4-49-58>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Multi-task neural network for solving the problem of recognizing the type of QAM and PSK modulation under parametric a priori uncertainty

Aleksei A. Paramonov[@], Van Minh Nguyen, Minh Tuong Nguyen

MIREA – Russian Technological University, Moscow, 119454 Russia

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

Abstract

Objectives. Automatic modulation recognition of unknown signals is an important task for various fields of technology such as radio control, radio monitoring, and identification of interference and sources of radio emission. The paper aims to develop a method for recognizing the types of signal modulation under conditions of parametric a priori uncertainty, including the uncertainty of carrier frequency- and initial signal phase values. An additional task consists in estimating the offset values of the carrier frequency or signal phase at the initial stage of the recognition process.

Methods. A multi-task learning with artificial neural network and the theory of cumulants of random variables are used.

Results. For signals with a carrier frequency and initial phase shift, cumulant approaches for QAM-8, APSK-16, QAM-64, and PSK-8 modulations are calculated. A multi-task learning with artificial neural network using cumulant features and a data standardization algorithm is presented. The results of the experiment show that using multi-task learning with an artificial neural network provides high accuracy of recognizing QAM-8 and APSK-16, QAM-64 and PSK-8 modulations with small mismatches of the carrier frequency or initial phase. The accuracy of determining the offset values from the carrier frequency or the initial phase for QAM-8, APSK-16, QAM-64, and PSK-8 modulation is high.

Conclusions. The multi-task learning with neural network using high-order signal cumulants makes it possible not only to recognize modulation types with high accuracy under conditions of a priori uncertainty of signal parameters, but also to determine the offset values of carrier frequency or initial signal phase from expected values.

Keywords: recognition, neural network, carrier frequency, initial phase, cumulant feature

• Submitted: 24.03.2023 • Revised: 11.04.2023 • Accepted: 02.05.2023

For citation: Paramonov A.A., Nguyen V.M., Nguyen M.T. Multi-task neural network for solving the problem of recognizing the type of QAM and PSK modulation under parametric a priori uncertainty. *Russ. Technol. J.* 2023;11(4):49–58. <https://doi.org/10.32362/2500-316X-2023-11-4-49-58>

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

Одним из перспективных направлений развития современных телекоммуникационных систем является внедрение в них интеллектуальных технологий, основная особенность которых заключается в способности к распознаванию и анализу информации, передаваемой через канал связи. Эта способность позволяет интеллектуальным системам обрабатывать данные в режиме

реального времени, делать прогнозы и принимать решения на основе полученных результатов. Одна из подобных задач – распознавание различных видов цифровой модуляции, которые могут быть использованы при передаче данных [1–8]. Знание параметров принимаемых сигналов позволяет идентифицировать передающее устройство, восстановить передаваемую информацию, а также дает возможность введения помеховых воздействий в радиоканалы локации и связи.

Однако в реальных условиях передачи информации часто возникают параметрические неопределенности, которые могут существенно повлиять на точность распознавания.

В настоящее время наиболее эффективным методом автоматического распознавания видов модуляции сигналов является использование многозадачной нейронной сети. В работах [1, 4] представлен алгоритм распознавания видов модуляции с использованием многозадачной нейронной сети в условиях априорной определенности значений несущей частоты и начальной фазы сигналов. В [1] для распознавания десяти видов модуляции (GMSK, 8-QAM, 16-QAM, 64-QAM, 16-APSK, 32-APSK, BPSK, QPSK, 8-PSK и 2-FSK)¹ в качестве информационных признаков использованы кумулянты до 9-го порядка. Результаты компьютерного моделирования показали, что многозадачная нейронная сеть с использованием кумулянтов высокого порядка позволяет выполнить не только распознавание видов модуляции, но и определение значения отношения сигнал/шум (ОСШ) принимаемого сигнала с высокой точностью. При ОСШ = 0 дБ для модуляции GMSK, QAM-8, APSK-16, APSK-32, BPSK и QPSK точность распознавания составляет 0.98, однако для модуляции QAM-16, QAM-64, PSK-8 и FSK-2 эта точность оказывается более низкой. Здесь и далее под точностью распознавания понимается вероятность правильного определения того или иного вида модуляции сигнала среди всех рассматриваемых видов модуляции. Целью настоящей статьи является анализ распознавания четырех видов модуляции QAM-8 и APSK-16, QAM-64 и PSK-8 в условиях параметрической априорной неопределенности и определение отстройки несущей частоты $\Delta\omega$ и начальной фазы $\Delta\varphi_0$ принимаемого сигнала от ожидаемых значений при ОСШ = 3 дБ, которое, согласно результатам моделирования, достаточно для распознавания сигналов с приемлемой точностью. Предполагалось, что принимаемый сигнал подвергнут предварительной

обработке, которая обеспечивает перенос на нулевую частоту, фильтрацию и дискретизацию сигнала с выхода фильтра нижних частот². Получаемый в результате предварительной обработки сигнал описывается следующим выражением:

$$r_k(t) = A(t) \{ \cos[\Delta\omega t + \varphi(t) + \Delta\varphi_0] + i \sin[\Delta\omega t + \varphi(t) + \Delta\varphi_0] \} = I_k(t) + iQ_k(t),$$

где $A(t)$ и $\varphi(t)$ – огибающая и фаза сигнала, $\Delta\omega$ – отстройка несущей частоты, $\Delta\varphi_0$ – отстройка начальной фазы, $I_k(t)$ и $Q_k(t)$ – синфазная и квадратурная составляющие сигнала.

Полученный комплексный сигнал $r_k(t)$ и комплексно-сопряженный с ним сигнал $\bar{r}_k(t) = I_k(t) - iQ_k(t)$ являются исходными данными для вычисления моментов и кумулянтов. Формулы вычисления моментов и кумулянтов высокого порядка подробно описаны в [1, 9]. В табл. 1 представлены примеры значений кумулянтов до 9-го порядка для модуляции QAM-64 и PSK-8 с различными значениями отстройки несущей частоты $\Delta\omega$ и начальной фазы $\Delta\varphi_0$. Анализ полученных значений кумулянтов позволяет утверждать, что информативность того или иного кумулянта о виде модуляции сигнала существенно зависит от отстроек $\Delta\omega$ и $\Delta\varphi_0$. Так, например, при отсутствии отстроек первый в таблице кумулянт $C_{2,0}$ для обоих различаемых видов модуляции QAM-64 и PSK-8 имеет один и тот же отрицательный знак, при $\Delta\omega = 900$ Гц знаки кумулянтов разные, при $\Delta\varphi_0 = 0.04$ рад знаки кумулянтов положительные.

ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

Нейросетевые методы распознавания видов модуляции основаны на выборе определенных информационных признаков, которые могут помочь в определении типа модуляции, и на построении базы знаний на основе анализа этих признаков. Каждая нейронная сеть распознавания видов модуляции имеет свой набор используемых информационных признаков и параметров, включая вид функции обработки данных и активации. Эти различия могут влиять на эффективность распознавания и на общую точность работы системы. На рис. 1 представлен процесс обучения нейронной сети.

¹ GMSK – гауссовская частотная модуляция с минимальным сдвигом (Gaussian minimum shift keying).

8-QAM, 16-QAM, 64-QAM – квадратурная амплитудная модуляция (quadrature amplitude modulation) с количеством уровней 8, 16 или 64.

16-APSK, 32-APSK – амплитудно-фазовая модуляция (amplitude and phase-shift keying) с количеством уровней 16 или 32.

BPSK – двоичная фазовая модуляция (binary phase-shift keying).

QPSK – четырехуровневая фазовая модуляция (quaternary phase-shift keying).

8-PSK – восьмиуровневая фазовая модуляция (8-phase-shift keying).

2-FSK – двоичная частотная модуляция (binary frequency-shift keying).

² Караван О.В. *Различение созвездий сигналов с квадратурной амплитудной модуляцией в условиях параметрической априорной неопределенности*: дис. ... канд. физ.-мат. наук. Воронеж; 2010. 120 с. [Caravan O.V. *Distinguishing constellations of signals with quadrature amplitude modulation under parametric a priori uncertainty*. Cand. Sci. Thesis (Phys.-Math.). Voronezh; 2010. 120 p. (in Russ.).]

Таблица 1. Значения кумулянтов различных порядков для модуляции QAM-64 и PSK-8

Кумулянт \ Модуляция	$\Delta\omega = 0$ и $\Delta\varphi_0 = 0$		$\Delta\omega = 900$ Гц		$\Delta\varphi_0 = 0.04$ рад	
	QAM-64	PSK-8	QAM-64	PSK-8	QAM-64	PSK-8
$C_{2,0}$	-0.00913	-0.0063	0.01209	-0.01225	1.08181	1.12291
$C_{3,0}$	-0.03605	0.04699	-0.04641	0.04051	-0.01704	0.00593
$C_{2,1}$	-0.00800	0.00439	0.01381	0.01362	-0.01714	0.00621
$C_{4,0}$	-0.56435	-0.01765	3.12438	3.20105	-0.36449	-0.27393
$C_{2,2}$	-0.61285	-0.97795	0.92006	0.85290	-0.36504	-0.27938
$C_{5,0}$	-0.20728	0.092786	0.04405	0.05804	-0.20601	0.20011
$C_{3,2}$	-0.00930	-0.00627	-0.05664	0.01384	-0.21641	0.19121
$C_{6,0}$	0.05008	-0.06516	0.97633	1.33265	1.17776	1.29652
$C_{3,3}$	1.62832	3.84222	-0.99868	-0.566	1.14810	1.33444
$C_{7,0}$	-1.56151	-0.47311	-1.01121	0.51321	1.33945	-0.51879
$C_{6,1}$	0.59933	-0.50570	-1.20013	-1.02734	1.40554	-0.47957
$C_{4,3}$	-0.12922	0.01030	-0.33111	-0.76300	3.20184	-1.62158
$C_{8,0}$	-12.8750	-1.42191	-191.842	-201.990	-4.24775	-11.0519
$C_{6,2}$	-11.6686	0.01107	-28.2597	-22.6214	-4.12208	-11.3511
$C_{4,4}$	1356.218	1413.430	1412.013	1359.459	1041.279	1197.02
$C_{9,0}$	-8.11293	4.10262	-28.8307	-15.8577	-37.4355	2.39996
$C_{8,1}$	6.42682	5.20209	27.0642	-7.68852	-244.414	108.016
$C_{6,3}$	-121.649	114.854	534.622	368.480	3453.083	4508.33
$C_{5,4}$	2467.71	2217.473	5219.517	4876.258	6296.198	8169.887

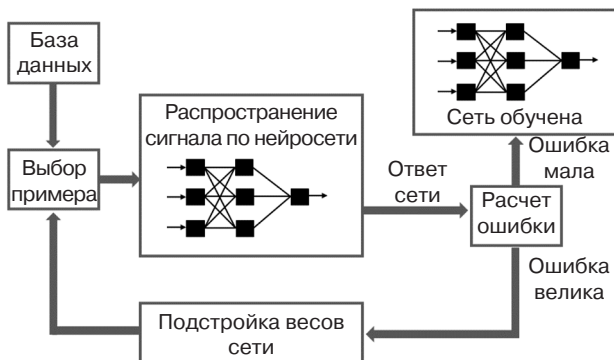


Рис. 1. Процесс обучения нейронной сети

Сама нейронная сеть представляет собой систему из множества нейронов, которая получает информацию, производит над ней простые вычисления и передает ее дальше. На каждый вход нейрона подается взвешенная сумма входных сигналов $net^{(j,1)}$, затем нейрон пропускает эту величину активации $\sigma^{(j,1)}$ через передаточную функцию, таким образом получаются выходные значения [10–15]:

$$net^{(j,1)} = \mathbf{w}^{(j,1)} \mathbf{x}^T = w_0^{(j,1)} + \sum_{i=1}^n w_i^{(j,1)} x_i,$$

$$o^{(j,1)} = f\left(net^{(j,1)}\right),$$

где $\mathbf{w}^{(j,1)} = (w_0^{(j,1)}, w_1^{(j,1)}, \dots, w_n^{(j,1)})$, $j = \overline{1, N_1}$ – вектор-строка синаптических связей на входе j -го нейрона, x_i – вектор-строка i -го входа, N_1 – количество нейронов в первом скрытом слое.

Функция активации $f(net^{(j,1)})$ в нейронной сети играет важную роль в определении выходного сигнала каждого нейрона на основе его входного сигнала. Эта функция позволяет нейрону принимать решения о том, активироваться ли и передавать информацию дальше или оставаться неактивным. Существует множество различных функций активации, которые можно использовать в нейронных сетях, каждая со своими преимуществами и ограничениями. В данной работе использована функция активации ReLU (Rectified Linear Unit), которая представляет собой простую нелинейную функцию, преобразующую входной сигнал, обнуляя все отрицательные значения и сохраняя положительные значения без изменений. Формально функция ReLU определяется следующим образом:

$$o^{(j,1)} = \text{ReLU}(x_i) = \max(0, x_i).$$

Обучение этой нейронной сети заключается в минимизации функции ошибки $E(\bar{\mathbf{w}})$, которая определяется следующим выражением:

$$E(\bar{w}) = \frac{1}{2} \sum_{l=1}^{N_{\text{вых}}} (u_l - o_l)^2,$$

где u_l и o_l – желаемое и фактическое состояния выходов нейронной сети l -го нейрона выходного слоя, $N_{\text{вых}}$ – количество нейронов в выходном слое.

В настоящее время существует несколько методов минимизации функции ошибки [2, 12, 13], но наибольшая эффективность метода обнаруживается при использовании функции стандартизации базы данных. Метод стандартизации (StandardScaler) является одним из способов предобработки данных в машинном обучении. Он используется для приведения всех исходных значений набора данных к набору значений из распределения с нулевым средним и стандартным отклонением, равным 1. Процесс стандартизации состоит из двух шагов. На первом шаге вычисляется среднее значение и стандартное отклонение каждого признака в наборе данных. На втором шаге каждое значение признака преобразуется по формуле [14, 15]:

$$x_i = \frac{z_i - \bar{Z}}{\sigma_z},$$

где z_i – исходное значение данных, $\bar{Z}$ и σ_z – среднее значение и стандартное отклонение признака.

В результате применения метода стандартизации формируется стандартизированная шкала, которая определяет место каждого значения в наборе данных, измеряя его отклонение от среднего в единицах стандартного отклонения. Это позволяет сделать данные сравнимыми и использовать их для машинного обучения. В качестве примера полученные в результате стандартизации значения кумулянтов для модуляции QAM-64 и PSK-8 и определенных расстройок частоты и начальной фазы сигнала представлены в табл. 2.

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Моделирование многозадачной нейронной сети проводилось в среде Python с помощью блокнота Google Colab [14]. Для распознавания двух групп модуляции QAM-8 и APSK-16, QAM-64 и PSK-8 сформированы четыре базы данных. Из них первые две базы сформированы для распознавания этих видов модуляции в условиях отстройки несущей частоты, каждая база состоит из 12800 сигналов (800 сигналов по каждому значению отстройки от несущей частоты). В условиях отстройки начальной фазы сформированы также две базы данных, каждая база состоит из 16000 сигналов (800 сигналов по каждому значению отстройки начальной фазы). Результаты моделирования распознавания вида модуляции

Таблица 2. Полученные в результате стандартизации значения кумулянтов

Модуляция Кумулянт	$\Delta\omega = 0$ и $\Delta\phi_0 = 0$		$\Delta\omega = 900$ Гц		$\Delta\phi_0 = 0.04$ рад	
	QAM-64	PSK-8	QAM-64	PSK-8	QAM-64	PSK-8
$C_{2,0}$	-0.04020	-0.03519	0.09287	-0.07583	-0.51078	-0.50800
$C_{3,0}$	0.001785	-0.04103	-0.03306	0.03589	-0.00352	0.000098
$C_{2,1}$	0.02363	0.00961	0.01718	0.01703	-0.00368	-0.000024
$C_{4,0}$	-0.43586	-0.43379	-0.42396	-0.42372	-0.00605	0.00768
$C_{2,2}$	-0.43481	-0.43849	-0.42072	-0.42135	0.00736	0.00891
$C_{5,0}$	-0.00613	-0.00639	-0.00661	-0.00655	0.005990	0.006273
$C_{3,2}$	-0.00892	-0.00816	-0.00861	-0.00798	0.005992	0.006272
$C_{6,0}$	-0.00173	-0.00179	-0.00152	-0.00143	0.03657	0.036584
$C_{3,3}$	0.03518	0.03641	0.03381	0.03404	0.03717	0.03718
$C_{7,0}$	-0.00901	-0.00902	-0.00903	-0.009	-0.00052	-0.000533
$C_{6,1}$	-0.01062	-0.01063	-0.01065	-0.01065	-0.00049	-0.0005
$C_{4,3}$	-0.00693	-0.00696	-0.00698	-0.007	0.00054	0.000528
$C_{8,0}$	0.37488	0.37488	0.37485	0.37485	-0.03074	-0.03074
$C_{6,2}$	0.32239	0.32241	0.32237	0.32238	-0.03235	-0.03236
$C_{4,4}$	-0.38226	-0.38227	-0.38226	-0.38226	-0.34224	-0.34224
$C_{9,0}$	-0.00205	-0.00205	-0.00205	-0.00205	-0.00450	-0.004506
$C_{8,1}$	-0.00829	-0.00829	-0.00828	-0.00829	-0.00401	-0.00401
$C_{6,3}$	-0.02267	-0.02226	-0.02226	-0.02226	-0.33499	-0.33499
$C_{5,4}$	-0.37783	-0.37783	-0.37783	-0.37783	-0.33517	-0.33517

в условиях отстройки несущей частоты представлены на рис. 2 и 3. Рисунки имеют вид таблиц, строки и столбцы которых соответствуют виду модуляции сигнала и отстройке несущей частоты. В ячейках приведены результаты распознавания вида модуляции. Например, для рис. 2: при распознавании сигналов QAM-8 с нулевым сдвигом частоты (на рисунке первая строка – QAM-8 0) все 80 сигналов, участвовавших в компьютерном эксперименте, распознаны верно. При распознавании сигнала QAM-8 со сдвигом частоты 1800 Гц (QAM-8 1800) распознаны верно 75 сигналов, а для 5 сигналов принято ошибочное решение, что это APSK-16 1800.

Из рисунков хорошо видно, что использование многозадачной нейронной сети позволяет выполнить не только распознавание видов модуляции, но и определение значений отстройки несущей частоты. Выше под точностью распознавания некоторого вида модуляции понималась вероятность правильного определения этого вида модуляции сигнала среди всех рассматриваемых видов модуляции. При моделировании эта вероятность оценивалась как выборочное среднее, т.е. отношение количества правильно распознанных сигналов с данным видом модуляции к общему числу участвовавших

в компьютерном эксперименте реализаций разных сигналов. Точность распознавания модуляции QAM-8 и APSK-16 с различными значениями $\Delta\omega$ составляет 0.96. На рис. 4 и 5 представлены результаты распознавания в условиях отстройки начальной фазы.

Из рис. 3 и 5 можно заметить, что при больших расстройках как частоты, так и фазы, точность распознавания модуляции QAM-64 и PSK-8 падает. Причина этого заключается в том, что значения нескольких кумулянтов при больших расстройках имеют нестабильное поведение для разных реализаций сигналов, а сами значения этих нескольких кумулянтов для указанных видов модуляции различаются незначительно. Но для оценки значений $\Delta\omega$ и $\Delta\phi_0$ использование многозадачной нейронной сети обеспечивает высокую точность.

На рис. 6 представлены результаты эксперимента по распознаванию принимаемого сигнала с неизвестным значением $\Delta\omega$. Результат эксперимента показал, что точность распознавания составляет 0.53 для модуляции QAM-64 и 0.47 для модуляции PSK-8. Значение отстройки от несущей частоты $\Delta\omega$, равное 600 Гц, определяется с высокой достоверностью.

Действительный вид модуляции и отстройки несущей частоты	QAM-8 0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 300	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 600	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 900	0	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 1200	0	0	0	0	80	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 1500	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0	0
	QAM-8 1800	0	0	0	0	0	0	75	0	0	0	0	0	0	5	0	0
	QAM-8 2000	0	1	0	0	0	0	0	48	0	0	0	0	0	0	0	31
	APSK-16 0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0
	APSK-16 300	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0
	APSK-16 600	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0
	APSK-16 900	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0
	APSK-16 1200	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0
	APSK-16 1500	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0
	APSK-16 1800	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0
	APSK-16 2000	0	0	0	0	0	0	0	30	0	0	0	0	0	0	1	49
	QAM-8 0																
	QAM-8 300																
	QAM-8 600																
	QAM-8 900																
	QAM-8 1200																
	QAM-8 1500																
	QAM-8 1800																
	QAM-8 2000																
	APSK-16 0																
	APSK-16 300																
	APSK-16 600																
	APSK-16 900																
	APSK-16 1200																
	APSK-16 1500																
	APSK-16 1800																
	APSK-16 2000																
		Решения алгоритма распознавания															

Рис. 2. Результаты распознавания модуляции QAM-8 и APSK-16 при разных значениях $\Delta\omega$

Действительный вид модуляции и отстройки несущей частоты	QAM-64 0	79	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	QAM-64 300	0	30	0	0	0	0	0	0	0	50	0	0	0	0	0	0
	QAM-64 600	0	0	72	0	0	0	0	0	0	0	8	0	0	0	0	0
	QAM-64 900	0	0	0	65	0	0	0	0	0	0	0	15	0	0	0	0
	QAM-64 1200	0	0	0	0	68	0	0	0	0	0	0	0	12	0	0	0
	QAM-64 1500	0	0	0	0	0	60	0	0	0	0	0	0	0	20	0	0
	QAM-64 1800	0	0	0	0	0	0	57	0	0	0	0	0	0	0	23	0
	QAM-64 2000	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0	25
	PSK-8 0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0
	PSK-8 300	0	13	0	0	0	0	0	0	67	0	0	0	0	0	0	0
	PSK-8 600	0	0	62	0	0	0	0	0	0	18	0	0	0	0	0	0
	PSK-8 900	0	0	0	46	0	0	0	0	0	0	34	0	0	0	0	0
	PSK-8 1200	0	0	0	0	52	0	0	0	0	0	0	28	0	0	0	0
	PSK-9 1500	0	0	0	0	0	36	0	0	0	0	0	0	44	0	0	0
	PSK-9 1800	0	0	0	0	0	0	37	0	0	0	0	0	0	43	0	0
	PSK-9 2000	0	0	0	0	0	0	0	29	0	0	0	0	0	0	51	0
	QAM-64 0	QAM-64 300	QAM-64 600	QAM-64 900	QAM-64 1200	QAM-64 1500	QAM-64 1800	QAM-64 2000	PSK-8 0	PSK-8 300	PSK-8 600	PSK-8 900	PSK-8 1200	PSK-9 1500	PSK-9 1800	PSK-9 2000	
Решения алгоритма распознавания																	

Рис. 3. Результаты распознавания модуляции QAM-64 и PSK-8 при разных значениях $\Delta\omega$

Действительный вид модуляции и отстройки начальной фазы	QAM-8 0	79	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	QAM-8 0.01	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 0.02	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 0.03	0	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 0.04	0	0	0	0	76	0	0	0	0	0	0	0	0	0	0	0	4	0	0
	QAM-8 0.05	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 0.06	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0	0	0	0
	QAM-8 0.07	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	8
	QAM-8 0.08	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0	0
	QAM-8 0.09	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0
	APSK-16 0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0	0
	APSK-16 0.01	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0
	APSK-16 0.02	0	0	0	0	0	0	0	0	0	0	1	79	0	0	0	0	0	0	0
	APSK-16 0.03	1	0	0	0	0	0	0	0	0	0	0	0	79	0	0	0	0	0	0
	APSK-16 0.04	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0
	APSK-16 0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0
	APSK-16 0.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0
	APSK-16 0.07	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	76	0	0
	APSK-16 0.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0
	APSK-16 0.09	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	68
		QAM-8 0	QAM-8 0.01	QAM-8 0.02	QAM-8 0.03	QAM-8 0.04	QAM-8 0.05	QAM-8 0.06	QAM-8 0.07	QAM-8 0.08	QAM-8 0.09	APSK-16 0	APSK-16 0.01	APSK-16 0.02	APSK-16 0.03	APSK-16 0.04	APSK-16 0.05	APSK-16 0.06	APSK-16 0.07	APSK-16 0.08
	Решения алгоритма распознавания																			

Рис. 4. Результаты распознавания модуляции QAM-8 и APSK-16 при разных значениях $\Delta\varphi_0$

Действительный вид модуляции и отстройки начальной фазы	QAM-64 0	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	QAM-64 0.01	0	78	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	
	QAM-64 0.02	0	0	78	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	
	QAM-64 0.03	0	0	0	55	0	0	0	0	0	0	0	0	0	25	0	0	0	0	
	QAM-64 0.04	0	0	0	0	56	0	0	0	0	0	0	0	0	0	24	0	0	0	
	QAM-64 0.05	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	60	0	0	
	QAM-64 0.06	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	76	0	
	QAM-64 0.07	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0	0	0	25	
	QAM-64 0.08	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	20	
	QAM-64 0.09	0	0	0	0	0	0	0	0	0	51	0	0	0	0	0	0	0	29	
	PSK-8 0	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	
	PSK-8 0.01	0	0	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	
	PSK-8 0.02	0	0	8	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	
	PSK-8 0.03	0	0	0	7	0	0	0	0	0	0	0	0	0	73	0	0	0	0	
	PSK-8 0.04	0	0	0	0	29	0	0	0	0	0	0	0	0	0	51	0	0	0	
	PSK-8 0.05	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	69	0	0	
	PSK-8 0.06	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	76	0	
	PSK-8 0.07	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	20	
	PSK-8 0.08	0	0	0	0	0	0	0	0	69	0	0	0	0	0	0	0	0	11	
	PSK-8 0.09	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	20
	QAM-64 0 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0 -	-	-	-	-	-	-	-	
	QAM-64 0.01 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.01 -	-	-	-	-	-	-	-	
	QAM-64 0.02 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.02 -	-	-	-	-	-	-	-	
	QAM-64 0.03 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.03 -	-	-	-	-	-	-	-	
	QAM-64 0.04 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.04 -	-	-	-	-	-	-	-	
	QAM-64 0.05 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.05 -	-	-	-	-	-	-	-	
	QAM-64 0.06 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.06 -	-	-	-	-	-	-	-	
	QAM-64 0.07 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.07 -	-	-	-	-	-	-	-	
	QAM-64 0.08 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.08 -	-	-	-	-	-	-	-	
	QAM-64 0.09 -	-	-	-	-	-	-	-	-	-	-	PSK-8 0.09 -	-	-	-	-	-	-	-	
	Решения алгоритма распознавания																			

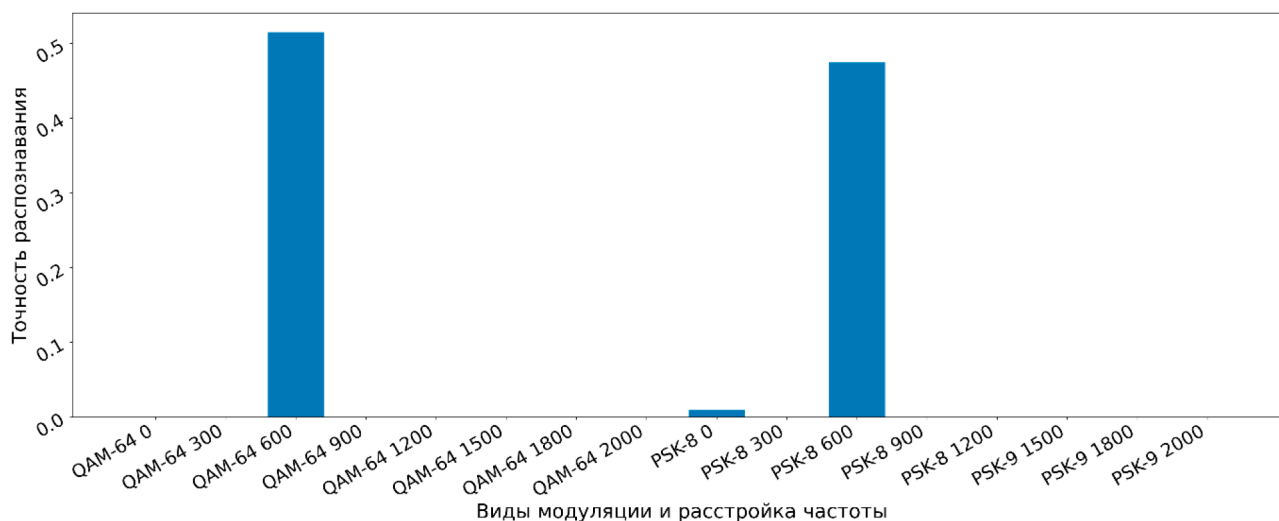
Рис. 5. Результаты распознавания модуляции QAM-64 и PSK-8 при разных значениях Δf_0 

Рис. 6. Результат эксперимента по распознаванию вида модуляции принимаемого сигнала

ЗАКЛЮЧЕНИЕ

Рассмотрен метод распознавания видов цифровой модуляции (QAM-8, APSK-8, QAM-64 и PSK-8) в случае неточного знания параметров сигнала, в т.ч. несущей частоты и начальной фазы. Построена многозадачная нейронная сеть с помощью алгоритма

стандартизации (StandardScaler) данных. Результаты моделирования позволяют утверждать, что многозадачная нейронная сеть с использованием кумулянтов в качестве информационного признака позволяет выполнить не только распознавание видов цифровой модуляции с высокой достоверностью в случае неточного определения значений несущей частоты и начальной

фазы, но и оценивать эти значения. В дальнейшем предполагается рассмотреть распознавание видов модуляции при одновременной неопределенности значений несущей частоты и начальной фазы.

СПИСОК ЛИТЕРАТУРЫ

1. Парамонов А.А., Нгуен М.В. Распознавание видов цифровой модуляции радиосигналов с использованием многозадачной нейронной сети. *Вестник воздушно-космической обороны*. 2022;4(36):91–97. URL: <https://www.elibrary.ru/item.asp?id=49815162>
2. Парамонов А.А., Тихонова О.В., Нгуен В.М. Распознавание видов цифровой модуляции радиосигналов с использованием многослойной нейронной сети по кумулятивным признакам. *Системы компьютерной математики и их приложения: Материалы XXIII международной научной конференции*. Смоленск: Изд-во СмолГУ; 2022. Вып. 23. С. 23–28.
3. Набилков В.Д., Приоров А.Л., Дубов М.А. Использование сверточной нейронной сети CLDNN в задаче распознавания видов модуляции радиосигналов. *Цифровая обработка сигналов и ее применение (ЦОС – 2021): доклады XXIII Международной конференции*. М.: Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова; 2021. С. 228–231. URL: <https://www.elibrary.ru/item.asp?id=45841831>
4. Нгуен М.В., Милорадов Г.А., Парамонов А.А. Сверточная нейронная сеть в задаче распознавания цифровой модуляции радиосигналов. *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем (Радиоинфоком – 2022): сборник научных статей по материалам VI Международной научно-практической конференции*. М.: МИРЭА – Российский технологический университет; 2022. С. 181–185. URL: <https://www.elibrary.ru/item.asp?id=49447332&pf=1>
5. Аведьян Э.Д., Дам В.Н. К выбору кумулятивных признаков в задаче распознавания видов цифровой модуляции радиосигналов. *Информатизация и связь*. 2015;4:11–15. URL: <https://www.elibrary.ru/item.asp?id=24853422>
6. Аджемов С.С., Кленов Н.В., Терешонок М.В., Чиров Д.С. Методы распознавания видов цифровой модуляций сигналов в когнитивных радиосистемах. *Вестник Московского Университета. Серия 3. Физика. Астрономия*. 2015;6:19–27. URL: <https://www.elibrary.ru/item.asp?id=25580690>
7. Гришин П.С., Шабанов А.В., Чеглов А.В. Распознавание цифровых модуляций радиосигналов с использованием многозадачной сверточной нейронной сети. *Интеллектуальные информационные системы: Теория и практика: сборник научных статей по материалам I Всероссийской конференции*. Часть 1. Курск: Курский государственный университет; 2020. С. 22–31. URL: <https://elibrary.ru/KADKBX>
8. Архипенков Д.В. Анализ параметров радиосигналов для идентификации источника излучения. *Доклады Белорусского государственного университета информатики и радиоэлектроники*. 2020;18(1):52–58. URL: <https://doi.org/10.35596/1729-7648-2020-18-1-52-58>

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

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

REFERENCES

1. Paramonov A.A., Nguyen M.V. Recognition of types of digital modulation of radio signals with multi-task neural network. *Vestnik vozduшно-kosmicheskoi oborony = Aerospace Defense Herald*. 2022;4(36):91–97 (in Russ.). Available from URL: <https://www.elibrary.ru/item.asp?id=49815162>
2. Paramonov A.A., Tikhonova O.V., Nguyen V.M. Recognition of digital modulation of radio signals using a multilayer neural network based on cumulative features. In: *Systems of computer mathematics and their applications: Proceedings of the 23rd International Scientific Conference*. Smolensk: SmolGU; 2022. Issue 23. P. 23–28 (in Russ.).
3. Nabilkov V.D., Priorov A.L., Dubov M.A. Using of the convolutional neural network CLDNN for classification of modulation types. In: *Digital Signal Processing and its Application (DSPA 2021): Reports 23rd International Conference*. Moscow: A.S. Popov Russian Scientific and Technical Society of Radio Engineering, Electronics and Communications; 2021. P. 228–231 (in Russ.). Available from URL: <https://www.elibrary.ru/item.asp?id=45841831>
4. Nguyen M.V., Miloradov G.A., Paramonov A.A. Convolutional neural network in the problem of recognizing digital modulation of radio signals. In: *Actual Problems and Prospects for the Development of Radio Engineering and Infocommunication Systems (Radioinfocom 2022): Collection of Scientific Articles based on the Materials of the 6th International Scientific and Practical Conference*. Moscow: MIREA – Russian Technological University; 2022. P. 181–185 (in Russ.). Available from URL: <https://www.elibrary.ru/item.asp?id=49447332&pf=1>
5. Avedyan E.D., Nzhich D.V. To the selection of the best cumulants features in the recognition task of the digital modulation kind of the radio signals. *Informatizatsiya i svyaz' = Informatization and Communication*. 2015;4:11–15 (in Russ.). Available from URL: <https://www.elibrary.ru/item.asp?id=24853422>
6. Adjemov S.S., Klenov N.V., Tereshonok M.V., Chirov D.S. Methods for the automatic recognition of digital modulation of signals in cognitive radio systems. *Moscow Univ. Phys. Bull.* 2015;70(6):448–456. <https://doi.org/10.3103/S0027134915060028> [Original Russian Text: Adjemov S.S., Klenov N.V., Tereshonok M.V., Chirov D.S. Methods for the automatic recognition of digital modulation of signals in cognitive radio systems. *Vestnik Moskovskogo Universiteta. Ser. 3. Fizika. Astronomiya*. 2015;6:19–27 (in Russ.). Available from URL: <https://www.elibrary.ru/item.asp?id=25580690>]
7. Grishin P.S., Shabanov A.V., Shcheglov A.V. Recognition of digital radio signal modulations using a multitasking convolutional neural network. In: *Intelligent Information Systems: Theory and Practice: A collection of scientific articles based on the materials of the First All-Russian Conference. Part 1*. Kursk: Kursk State University; 2020. P. 22–31 (in Russ.). Available from URL: <https://elibrary.ru/KADKBX>

9. Кендалл М., Стюарт А. *Теория распределений*: пер. с англ. М.: Наука; 1966. 588 с.
10. Хайкин С. *Нейронные сети: полный курс*: пер. с англ. М.: Вильямс; 2006. 1104 с.
11. Elgendy M. *Deep Learning for Vision Systems*. Manning Publications Co; 2020. 480 p. ISBN 978-1-6172-9619-2
12. Geron A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc.; 2019. 856 p. ISBN 978-1-4920-3264-9
13. Finlay S. *Artificial Intelligence and Machine Learning for Business: A No-Nonsense Guide to Data Driven Technologies*. Relativistic; 2017. 150 p.
14. Raschka S., Mirjalili V. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*. Packt Publishing; 2019. 770 p. ISBN 978-1-7899-5575-0
15. Воронина В.В., Михеев А.В., Ярушкина Н.Г., Святков К.В. *Теория и практика машинного обучения*. Ульяновск: УлГТУ; 2017. 290 с. ISBN 978-5-9795-1712-4
8. Arkhipenkov D.V. Analysis of radio signal parameters for emission source identification. *Doklady Belorusskogo Gosudarstvennogo Universiteta Informatiki i Radioelektroniki (Doklady BGUIR)*. 2020;18(1):52–58 (in Russ.). <https://doi.org/10.35596/1729-7648-2020-18-1-52-58>
9. Kendall M., Stuart A. *Teoriya raspredelenii (Distribution Theory)*: transl. from Engl. Moscow: Nauka; 1966. 588 p. (in Russ.). [Kendall M.G., Stuart A. *The Advanced Theory of Statistics. V. I. Distribution Theory*. London; 1963. 433 p.]
10. Haykin S. *Neironnye seti: polnyi kurs (Neural networks: a complete course)*: transl. from Engl. Moscow: Vil'yams; 2006. 1104 p. (in Russ.). [Haykin S. *Neural Networks*. Upper Saddle River, NJ: Prentice Hall; 1999. 842 p.]
11. Elgendy M. *Deep Learning for Vision Systems*. Manning Publications Co; 2020. 480 p. ISBN 978-1-6172-9619-2
12. Geron A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc.; 2019. 856 p. ISBN 978-1-4920-3264-9
13. Finlay S. *Artificial Intelligence and Machine Learning for Business: A No-Nonsense Guide to Data Driven Technologies*. Relativistic; 2017. 150 p.
14. Raschka S., Mirjalili V. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*. Packt Publishing; 2019. 770 p. ISBN 978-1-7899-5575-0
15. Voronina V.V., Mikheev A.V., Yarushkina N.G., Svyatov K.V. *Teoriya i praktika mashinnogo obucheniya (Theory and Practice of Machine Learning)*. Ul'yanovsk: UISTU; 2017. 290 p. (in Russ.). ISBN 978-5-9795-1712-4

Об авторах

Парамонов Алексей Анатольевич, д.т.н., профессор, кафедра радиоэлектронных систем и комплексов Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: paramonov@mirea.ru. Scopus Author ID 57208923552, <http://orcid.org/0000-0002-4537-4626>

Нгуен Ван Минь, аспирант, кафедра радиоэлектронных систем и комплексов Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: nguyenvinhvq1009@gmail.com. <https://orcid.org/0009-0001-5062-7093>

Нгуен Минь Тьонг, к.т.н., доцент, кафедра информатики Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: nguen_m@mirea.ru. <https://orcid.org/0009-0002-7267-1121>

About the authors

Aleksei A. Paramonov, Dr. Sci. (Eng.), 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: paramonov@mirea.ru. Scopus Author ID 57208923552, <http://orcid.org/0000-0002-4537-4626>

Van Minh Nguyen, 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: nguyenvinhvq1009@gmail.com. <https://orcid.org/0009-0001-5062-7093>

Minh Tuong Nguyen, Cand. Sci. (Eng.), Associate Professor, Department of Informatics, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: nguen_m@mirea.ru. <https://orcid.org/0009-0002-7267-1121>

Математическое моделирование
Mathematical modeling

УДК 519.6
<https://doi.org/10.32362/2500-316X-2023-11-4-59-71>



НАУЧНАЯ СТАТЬЯ

Моделирование пространственного распространения волн пандемии COVID-19 в России на основе кинетико-переносного описания

В.В. Аристов^{1, 2, @},
А.В. Строганов¹,
А.Д. Ястребов¹

¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия

² Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, 119333 Россия

@ Автор для переписки, e-mail: aristowl@yandex.ru

Резюме

Цели. Пандемия COVID-19 обладает рядом важных особенностей по сравнению с прошлыми эпидемиями. Помимо высокой степени заражения, она имеет высокую скорость распространения за счет мобильности населения, связанной, в частности, с возросшей скоростью средств передвижения. Целью данной работы является построение математической модели распространения пандемии и выявление закономерностей в предположении, что основным источником вирусной инфекции в России является г. Москва. Для этого строится двухпараметрическая кинетическая модель, описывающая пространственное распространение эпидемии. Параметры находятся с помощью теоретических построений, оценок известных данных о статистике передвижения транспортных средств и плотности населения в различных странах, а также с учетом развития первой волны на примере России, Италии и Чили с проверкой значений для последующих эпидемических волн. Исследуется возможность предсказывать скорость пространственного распространения вируса по временному интервалу запаздывания достижения пика заражений в России по сравнению с Москвой. Это связано с географическими особенностями: в России, как и в некоторых других странах, можно выделить основной источник распространения инфекции. Таким источником в России выступает г. Москва – крупнейший в стране транспортный узел. Для реализации цели в настоящей работе изучается развитие эпидемических событий в России, начиная с 3-й, и вплоть до последних 5-й и 6-й волн.

Методы. Используются методы математического моделирования и методы обработки статистических данных.

Результаты. Подтверждено, что величина запаздывания достижения пика заражений составляет в среднем 2.5 недели. Выявлена сохраняемость параметров для различных волн, поэтому модель обладает предсказательными возможностями. Проверки проводились для последовательности волн, для которых делались соответствующие предсказания о развитии заражения для России в целом и о том, когда произойдет спад. Данные прогнозы подтверждались для всех волн, начиная с 3-й, и вплоть до последней 6-й волны, что подтверждает найденную закономерность, важную для прогнозирования будущих событий.

Выводы. Прогнозы о начале и скорости выздоровления подтвердились, что дало возможность уверенно прогнозировать, в частности, протекание 5-й и 6-й волн пандемии, связанной с новым вирусным штаммом «омикрон». Предсказания, которые делались заранее, были проверены и получили подтверждение.

Ключевые слова: кинетическое уравнение, COVID-19, распространение волн

• Поступила: 15.09.2022 • Доработана: 02.02.2023 • Принята к опубликованию: 15.05.2023

Для цитирования: Аристов В.В., Строганов А.В., Ястребов А.Д. Моделирование пространственного распространения волн пандемии COVID-19 в России на основе кинетико-переносного описания. *Russ. Technol. J.* 2023;11(4):59–71. <https://doi.org/10.32362/2500-316X-2023-11-4-59-71>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Modeling of spatial spread of COVID-19 pandemic waves in Russia using a kinetic-advection model

Vladimir V. Aristov ^{1, 2, @},
Andrey V. Stroganov ¹,
Andrey D. Yastrebov ¹

¹ MIREA – Russian Technological University, Moscow, 119454 Russia

² Federal Research Center “Computer Science and Control”, Russian Academy of Sciences, Moscow, 119333 Russia

@ Corresponding author, e-mail: aristowl@yandex.ru

Abstract

Objectives. COVID-19 has a number of specific characteristics that distinguish it from past pandemics. In addition to the high infection rate, the high spread rate is due to the increased mobility of contemporary populations. The aim of the present work is to construct a mathematical model for the spread of the pandemic and identify patterns under the assumption that Moscow comprises the main source of viral infection in Russia. For this purpose, a two-parameter kinetic model describing the spatial spread of the epidemic is developed. The parameters are determined using theoretical constructions alongside statistical vehicle movement and population density data from various countries, additionally taking into account the development of the first wave on the examples of Russia, Italy and Chile with verification of values obtained from subsequent epidemic waves. This paper studies the development of epidemic events in Russia, starting from the third and including the most recent fifth and sixth waves. Our two-parameter model is based on a kinetic equation. The investigated possibility of predicting the spatial spread of the virus according to the time lag of reaching the peak of infections in Russia as a whole as compared to Moscow is connected with geographical features: in Russia, as in some other countries, the main source of infection can be identified. Moscow represents such a source in Russia due to serving as the largest transport hub in the country.

Methods. Mathematical modeling and data analysis methods are used.

Results. A predicted time lag between peaks of daily infections in Russia and Moscow is confirmed. Identified invariant parameters for COVID-19 epidemic waves can be used to predict the spread of the disease. The checks were carried out for the wave sequence for which predictions were made about the development of infection for Russia and when the recession following peak would occur. These forecasts for all waves were confirmed from the third to the last sixth waves to confirm the found pattern, which can be important for predicting future events.

Conclusions. The confirmed forecasts for the timing and rate of the recession can be used to make good predictions about the fifth and sixth waves of infection of the Omicron variant of the COVID-19 virus. Earlier predictions were confirmed by the statistical data.

Keywords: kinetic equation, COVID-19, wave propagation

• Submitted: 15.09.2022 • Revised: 02.02.2023 • Accepted: 15.05.2023

For citation: Aristov V.V., Stroganov A.V., Yastrebov A.D. Modeling of spatial spread of COVID-19 pandemic waves in Russia using a kinetic-advection model. *Russ. Technol. J.* 2023;11(4):59–71. <https://doi.org/10.32362/2500-316X-2023-11-4-59-71>

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

Пандемия COVID-19 явилась вызовом не только в социальном и медицинском плане, но также потребовала развития известных и выработки новых научных подходов для изучения этого сложного явления. Все это обусловило интенсивные исследования, к которым подключились многочисленные лаборатории по всему миру. Современное развитие информационных технологий позволяет быстро обрабатывать большие объемы данных. Помимо сложных математических моделей с использованием мощных компьютерных ресурсов, для описания распространения вируса свою роль могут играть простые модели, передающие важные черты процесса. Развитие пандемии COVID-19 по сравнению с протеканием предыдущих известных эпидемий отличается рядом особенностей. Это связано, в частности, с тем, что существенно увеличились скорости транспортных средств (ТС). Цель данной работы – моделировать процессы пространственного распространения эпидемии с изучением сценариев быстрого переноса вируса (традиционные методы, как правило, основаны на уравнениях диффузионного типа, применяемых для медленных процессов, связанных с контактами [1–3]).

Отметим ряд работ, где используются методы (в т.ч. статистические) для изучения развития эпидемий [4–14]. По большей части здесь применяются модели SIR¹ (SEIR²), в которых, как правило, изучаются пространственно-локальные процессы, происходящие во времени. Сопряжение с пространственным развитием в таких работах проводится нечасто. Мы исследуем только пространственное распространение эпидемии.

В настоящей работе строится кинетическая модель, аналогичные модельные уравнения применяются для изучения различных физических процессов [15, 16].

Настоящая работа является продолжением, развитием и обобщением наших исследований в работе [17], где были определены общие черты метода с помощью рассмотрения и изучения 1-й и 2-й волн пандемии. Разработанная модель применяется для изучения всех

последующих волн пандемии COVID-19 в России. С помощью предложенного кинетического подхода можно судить о характере распространения современной пандемии для стран, к которым применима одномерная модель. Для 1-й волны пандемии исследовались процессы в Италии, Чили и России. Начиная со 2-й волны – в основном в России. Предполагалось (это подтверждается данными), что в указанных странах были основные центры, откуда происходило затем распространение заражения, и определялась величина запаздывания развития болезни в отдельных регионах и в стране в целом. Это позволяет давать прогнозы и для характера последующих волн пандемии, что демонстрируется на примере распространения таких волн в России.

Важное предположение заключается в следующем. Выделяются два механизма заражения: переносный (изучаемый в настоящей работе) и контактный, реализуемый в регионах. Наложение этих двух факторов и дает сумму количества заражений. Первая фаза, соответствующая переносному заражению, закладывает основы для следующей фазы, а именно, развития заражения, связанного с контактами. Мы изучаем распространение носителей вируса, которое осуществляется при использовании различных ТС. Использовалась осредненная скорость передвижения, в которой учитывались скорости самолета, поезда, автобуса и автомашины. Параметром модели является и величина «сопротивления» продвижению носителей вирусов, имеющая размерность частоты, что задается частотой высаживания пассажиров из ТС при их движении к местам проживания. Этот параметр зависит, главным образом, от плотности населения. Допускается одинаковый характер протекания болезни в дальнейшем в различных местах. Третья фаза связана с «распространением выздоровления», что определяется по достижению максимума новых заражений в день для каждого региона.

Географические особенности России позволяют применять простую одномерную модель также, как для Италии и Чили (для этих стран рассматривалась распространение пандемии для 1-й волны). При этом, с учетом характера заражения для России, основной источник во всех волнах связан с г. Москвой (изучались регионы к востоку от столицы). Это дает возможность судить о сдвиге во времени начала выздоровления.

¹ Susceptible infected recovered.

² Susceptible exposed infected recovered.

РАЗРАБОТАННАЯ ПЕРЕНОСНО-КИНЕТИЧЕСКАЯ МОДЕЛЬ И ЕЕ УТОЧНЕНИЕ

Приведем кратко использовавшуюся одномерную двухпараметрическую модель переноса с учетом кинетического члена:

$$\frac{\partial n}{\partial t} + U \frac{\partial n}{\partial x} = -\sigma n(t, x), \quad (1)$$

где t – время, x – расстояние, $n(t, x)$ – плотность перемещающихся носителей вируса в ТС, U – средняя скорость ТС, σ – отмеченный выше коэффициент «сопротивления» (имеющий размерность частоты) движению инфицированных элементов в основном за счет высаживания пассажиров из ТС в местах проживания. Начальное условие для задачи Коши принимается в виде:

$$n_0(x) = H(-x),$$

где $H(x)$ – функция Хевисайда. Такая постановка задачи означает, что фактически в исследуемую область $x > 0$ через границу $x = 0$ поступают носители вируса. Линейное уравнение (1) разрешается стандартными методами, аналитическое решение имеет вид:

$$n(t, x) = n_0(x - Ut) e^{-\frac{\sigma x}{U}}. \quad (2)$$

Обозначаем плотность высадившихся из ТС в данной точке носителей вируса через $n_M(t, x)$. Эта плотность растет также, как убывает $n(t, x)$, следовательно, можно записать

$$\begin{aligned} \frac{dn}{dt} + \frac{dn_M}{dt} &= 0, \\ \frac{dn}{dt} &= \frac{\partial n}{\partial t} + U \frac{\partial n}{\partial x}, \\ \frac{dn_M}{dt} &= \frac{\partial n_M}{\partial t}. \end{aligned}$$

Отметим, что полная и частная производная от n отличаются друг от друга, поскольку есть продвижение «элементов» (носителей вируса) со скоростью U . Для n_M эти производные совпадают, поскольку «элементы» с такой плотностью неподвижны.

С учетом (1) уравнение для $n_M(t, x)$ принимает следующий вид:

$$\frac{\partial n_M}{\partial t} = \sigma n. \quad (3)$$

С появлением штамма вируса «дельта», проникшего преимущественно из Индии через аэропорты г. Москвы, связана 3-я волна. Четвертая волна, в определенном смысле аналогичная второй, порождается в основном возвращавшимися из летних отпусков, но и здесь также главным источником развития новой волны был г. Москва. Другие города также вносили свой вклад, поэтому развитие процессов имело несколько «размытый» характер по сравнению с развитием процессов в 1-й и 3-й волнах. Пятая волна имеет свои особенности, поскольку «взрывной» характер заражения вирусом штамма «омикрон» допустимо может приводить к некоторым коррективам пространственного распространения, в чем-то ускоряя этот процесс.

Важно подчеркнуть, что общие закономерности, определенные по 1-й волне и подтвержденные во 2-й, остаются неизменными, поскольку принятый учет инфицирования, связанный с перемещением транспортом зараженных пассажиров, приводит к повторению характера распространения пандемических волн. Средняя скорость распространения зависит не от интенсивности заражения новыми вирусами, но от средней скорости ТС. Поэтому модель применяется для описания и последующих пандемических волн. Это относится и к 6-й волне, последней по времени.

В настоящей работе используется одномерная кинетическая модель для предсказания и затем – проверки характера распространения 3-й, 4-й и 5-й (и фактически 6-й) волн пандемии. Существенный общий вывод заключается в следующем: выявлено запаздывание максимума заражения в России по сравнению с максимумом заражения в г. Москве примерно на 2.5 недели (с учетом неизбежной статистической погрешности максимум находится в пределах от 2 до 3 недель). Таким образом, по развитию заболевания в г. Москве можно судить о сроках прохождения волны пандемии по России в целом. Количество заражений в день меняется от волны к волне, но можно выявить некоторые инварианты формы волн. Поэтому предпринимается попытка найти функции, описывающие повторяющиеся формы волн заражения, учитывая, конечно, различную интенсивность инфекции для разных штаммов и, соответственно, различную «амплитуду» колебаний количества заразившихся за день (основной изучаемый показатель).

Влияет ли величина инкубационного периода заражения на представленные результаты? Для 1-й волны период составлял 7–10 дней, для последующих волн он был таким же или меньшим. Но данная величина для г. Москвы и регионов России одна и та же, поэтому влияния этого параметра на величину запаздывания заражения практически не отмечено.

Решение данного уравнения записывается так

$$n_M(t, x) = n_{M_0}(x) + \int_0^t \sigma n(\tau, x) d\tau.$$

Подстановка выражения (2) и начальных условий $n_{M_0}(x) = 0$ и $n_0(x) = H(-x)$ позволяет получить выражение для $n_M(t, x)$

$$n_M(t, x) = \frac{\sigma}{U} (Ut - x) \cdot H(Ut - x) e^{-\frac{\sigma}{U} x}. \quad (4)$$

Вместо аналитического выражения в большой серии расчетов часто удобнее было использовать численный подход (применяется простая схема Куранта – Изааксона – Риса).

Вопрос уточнения параметров U и σ важен. Мы будем обсуждать оценку первого параметра, принимая значение второго равным полученному в наших предыдущих работах. С помощью значений первого из указанных параметров, полученных при изучении 1-й и 2-й волн, делаются предсказания о скорости распространения пандемии, а также соответственно о скорости волны выздоровления. Можно попытаться теоретически оценить это среднее значение скорости передвижения обобщенного ТС и сравнить с реальными данными распространения заболевания, причем допустимо вносить некоторые поправки в получающийся параметр скорости для последующих волн. Мы рассматриваем 4 основных вида ТС: автомобиль, автобус, поезд, самолет. Свою роль вносят и велосипеды, мотоциклы, самокаты, а также и пеший способ передвижения. Но для перемещения на достаточно ощутимые расстояния с большим количеством пассажиров именно указанные виды транспорта являются основными. Надо вводить некоторое весовое осреднение, учитывая относительную долю пассажиров, пользующихся тем или иным средством передвижения.

Были найдены также более точные количественные оценки для средневесовой скорости поездов. Согласно имеющимся данным³ в 2021 г. на поездах по России было перевезено поездами ближнего следования $9.614 \cdot 10^8$ пассажиров, которые в сумме преодолели $2.9 \cdot 10^{10}$ км. Поездами дальнего следования было перевезено $9.2 \cdot 10^7$ пассажиров, которые в сумме преодолели $7.44 \cdot 10^{10}$ км. Таким образом, в среднем поездами ближнего следования каждый пассажир проезжал 30.16 км, а поездами дальнего

следования – 808.7 км. По данным Министерства транспорта Российской Федерации⁴ средняя скорость поездов в России в 2021 г. находится на интервале от 57 до 65 км/ч, т.е. пассажиры поездов и ближнего, и дальнего следования в среднем доезжали до пункта назначения менее, чем за сутки. Можно считать, что найденные ранее средние расстояния преодолеваются за сутки, значит, средняя скорость поездов ближнего следования полагается 30.16 км/день, а поездов дальнего следования – 808.7 км/день. В итоге средняя скорость поездов в 2021 г. в России равна

$$U = \frac{9.614}{9.614 + 0.92} \cdot 30.16 + \frac{0.92}{9.614 + 0.92} \cdot 808.7 = 98.15 \text{ км/день}.$$

Можно ожидать, что реальные данные будут близки к этой величине. Так и оказалось для всех изучаемых волн пандемии. Значит средневзвешенная скорость ТС и скорость распространения эпидемии в первые несколько дней близки, что относится ко всем волнам. Отличия в 10–15% лежат в пределах статистической ошибки. Это подтверждает предположение о возможности использования параметра U в модели.

Заметим, что некоторые полумпирические выкладки, основанные на анализе величин пассажирооборота за год по другим ТС, подтверждают вышеприведенные оценки, но мы их не приводим, поскольку они требуют более тщательного рассмотрения, что не является целью настоящей статьи.

РАСПРОСТРАНЕНИЕ 3-Й ВОЛНЫ ПАНДЕМИИ COVID-19 В РОССИИ

Результаты, полученные в [17], используются для анализа развития последующих волн. Прежде всего рассмотрим характер распространения 3-й волны. Вначале в [17] приводятся данные развития пандемии по г. Москве. По достижении максимума заражений за день судят о начале выздоровления, т.е. по календарному числу, соответствующему максимуму. Делается прогноз о возможном дне достижения такого максимума и начале выздоровления по России в целом. Затем эта гипотеза проверяется.

Согласно обработке данных в [17], было получено значение параметра осредненной скорости ТС. Эта скорость оценивалась как $U = 75\text{--}90$ км/день.

³ В 2021 году на сети ОАО «РЖД» перевезено свыше 1 млрд пассажиров | Пресс-релизы | Компания (rzd.ru). [More than 1 bn passengers carried on the Russian Railways network in 2021 | Press releases | Company (rzd.ru)]. <https://company.rzd.ru/ru/9397/page/104069?id=269758>. Дата обращения 21.12.2022. / Accessed December 21, 2022 (in Russ.).

⁴ Средняя скорость движения пассажирских поездов в России может вырасти до 65 км/ч к 2031 году. Министерство транспорта Российской Федерации. [Average speed of passenger trains in Russia may grow to 65 km/h by 2031. Ministry of Transport of the Russian Federation.] <https://mintrans.gov.ru/press-center/branch-news/595>. Дата обращения 21.12.2022. / Accessed December 21, 2022 (in Russ.).

Это фактически близко к вышеприведенным оценкам средней скорости с точностью 10% для величины 90–100 км/день.

Можно подсчитать время продвижения волны от г. Москвы до «центра масс» населения России. По графику распределения плотности, приведенному в [11], эта точка примерно соответствует расстоянию 1000–1200 км от г. Москвы, так что ожидается, что существенное проявление заражения в г. Москве скажется в целом по России примерно через 2 или более недель. Оценка достаточно грубая, поэтому можно ожидать, что это запаздывание составит и до 3 недель. Тогда, например, максимум заражений в день в г. Москве появляется примерно на 2 недели ранее, чем максимум заражения в целом по России. Эта характерная точка трактуется как начало выздоровления, так что волна выздоровления имеет соответствующее запаздывание.

На основе данных о количестве заражений в г. Москве и России в целом построены графики (рис. 1), относящиеся к июню 2021 г. (данные

по 29 июня). Видно, что максимум заражения для г. Москвы пройден (согласно применявшемуся среднеквадратичному приближению) в конце месяца при том, что для России в целом происходило только нарастание числа заражений в день. Поэтому можно было ожидать, что максимум заражения по России в целом будет достигнут к середине июля.

На рис. 2 показаны данные по заражению для г. Москвы и России на 11 июля, что на 12 дней позже по сравнению с предыдущими графиками на рис. 1. Виден отчетливо сформировавшийся «горб» заражений по г. Москве, при том что по России максимум только начинает обозначаться.

На рис. 3 представлена ожидаемая парабола развития событий (числа заражений по России) на основе полученных данных вплоть до 15 июля 2021 г. В принципе замечен намечающийся максимум по реальным данным к середине июля. Так что, по-видимому, линию параболы надо было бы уточнить, поскольку дается только экстраполяция.

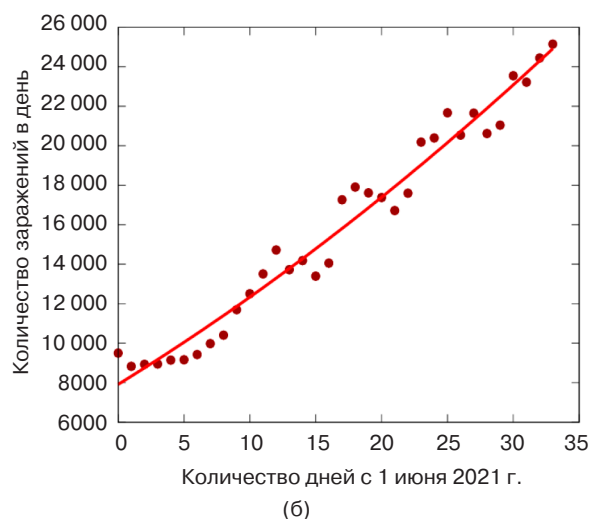
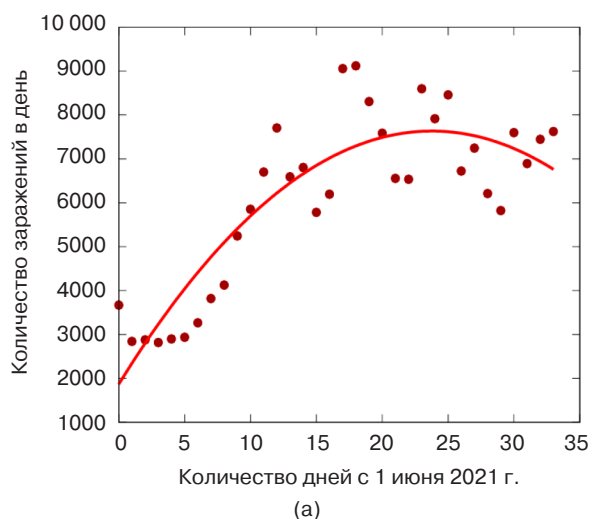


Рис. 1. Число заражений в день для г. Москвы (а) и для России в целом (б) в июне 2021 г.

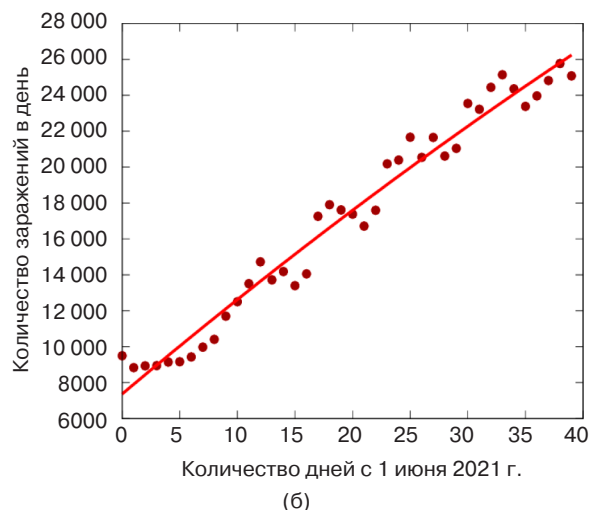
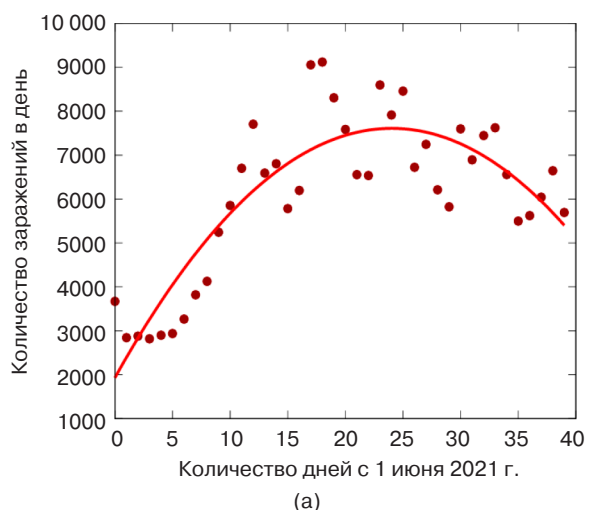


Рис. 2. Число заражений в день для г. Москвы (а) и для России в целом (б) в июне – июле 2021 г.

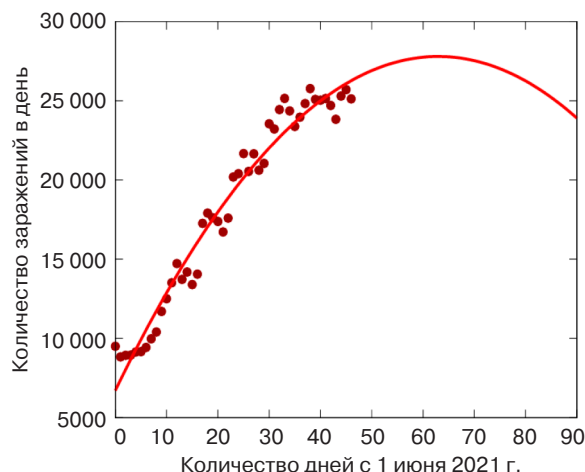
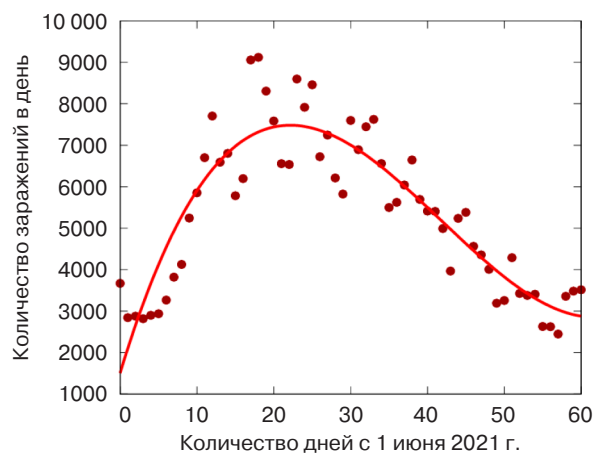


Рис. 3. Ожидаемое число заражений для 3-й волны для России в целом на основе полученных данных до середины июля 2021 г.

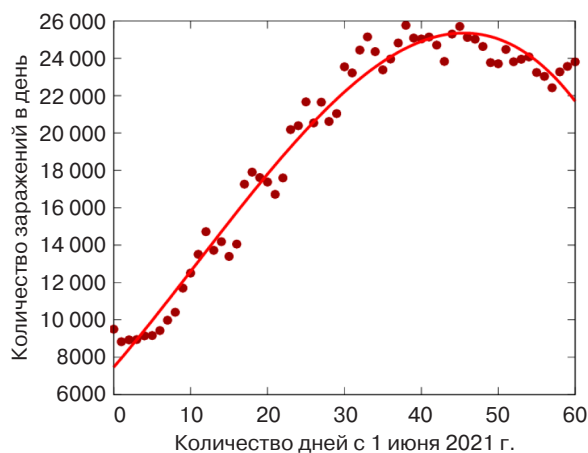
Можно заключить, что предсказания оказались оправданными. Максимум по г. Москве был достигнут к 25 июня. Поэтому мы предположили, что максимум по России должен быть достигнут (с учетом параметров, полученных на основе изучения предыдущих волн пандемии) примерно через 2–3 недели,

т.е. к середине июля. Кривые, построенные по методу наименьших квадратов (МНК), соответствовали прогнозу. Но интересно, что и «пики», т.е. абсолютные максимумы заражений по г. Москве и России смещены именно на такую величину: они приходятся примерно на 25 июня и 15 июля соответственно (вообще говоря, такие величины недостаточно показательны из-за не очень представительной статистики и случайных выбросов). С течением времени и появляющихся с каждым днем новых данных параболы по МНК смещаются несколько «вправо», такая тенденция отмечалась и ранее. Но величина запаздывания остается той же: максимумы соответственно теперь 5 июля и 25 июля.

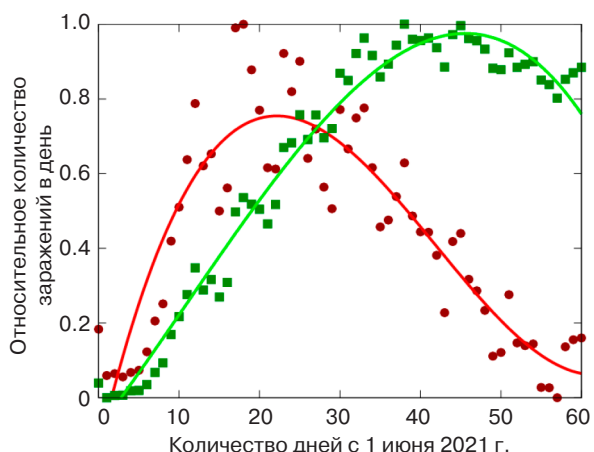
Для 3-й волны сдвиг по времени начала выздоровления – максимума заражения для г. Москвы и России в целом составляет три недели. Максимум для г. Москвы на рис. 4 соответствует примерно началу – середине третьей декады июня. Максимум по России отвечает примерно середине июля. На графиках рис. 4в представлены относительные величины заражений, эти величины находились по формуле $A_{\text{отн}} = (A - A_{\text{min}})/(A_{\text{max}} - A_{\text{min}})$, где A_{max} и A_{min} означают соответственно максимум и минимум этой величины (рис. 4а и 4б).



(а)

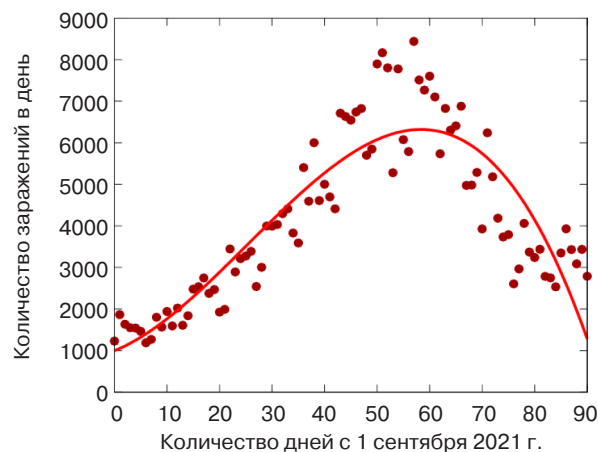


(б)

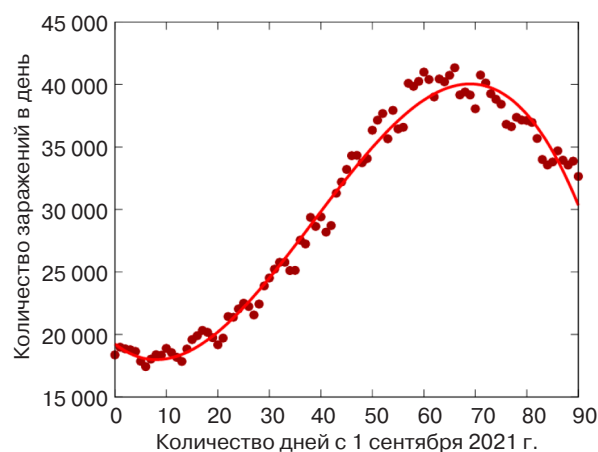


(в)

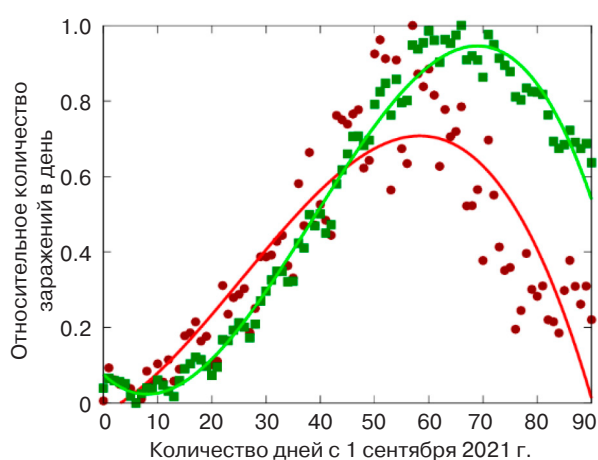
Рис. 4. Реальные данные и линии осреднения по МНК, общее количество инфицированных для г. Москвы (а) и для России в целом (б), относительное количество заражений для г. Москвы и для России в целом (в)



(а)



(б)



(в)

Рис. 5. Реальные данные и линии осреднения по МНК, общее количество инфицированных для г. Москвы (а) и для России в целом (б), относительное количество заражений для г. Москвы и для России в целом (в)

РАЗВИТИЕ 4-Й ВОЛНЫ ПАНДЕМИИ

Закономерности, выявленные для предыдущих волн, были использованы для предсказания поведения 4-й волны. В целом прогнозы подтвердились, это касается и характера кривых заражения, и сдвигов России относительно г. Москвы.

Для 4-й волны датами максимумов заражений для г. Москвы и России в целом оказались конец октября и середина ноября, разделенные 2–3 неделями (рис. 5). Причем для 4-й волны этот сдвиг несколько меньше, чем для 3-й, что допустимо объяснить возвращением заболевших из отпусков не только через г. Москву, но и через другие города.

ПРОВЕРКА МОДЕЛИ ДЛЯ 5-Й ВОЛНЫ ПАНДЕМИИ

Рассматриваются аналогичные процессы для 5-й волны. Заметим, что 10 февраля 2022 г. изменился метод подсчета заболевших, увеличивший показания, поэтому в расчеты вносился поправочный коэффициент. В связи с новым штаммом «омикрон» заражение происходило интенсивнее, но ожидалось, что на пространственное распространение это не должно было

сильно повлиять. Предполагалось, что поскольку максимум по г. Москве соответствует примерно 1 февраля, максимум по России будет около 14 февраля.

Развитие этой волны характеризуется своими особенностями. Здесь возможны резкие («взрывные») всплески заражения на местах, в частности в г. Москве, что соответствует большей способности к заражению данным штаммом. Так на рис. 6, где показано количество заражений в первые 3 недели развития новой волны, видно, что, начиная с 18 января, число заболевших за день резко возросло. Можно было ожидать, что такое течение болезни проявится для России в целом примерно с 2–3-недельным запаздыванием. Это действительно отмечается примерно с 3 февраля.

На графике рис. 6 приведено количество заболевших в день для России без учета московских показателей, что кажется более адекватным, чем с учетом г. Москвы.

Прогноз по результатам января заключался в том, что можно было ожидать спад заражения (начало выздоровления) для России в целом к середине февраля. Такой прогноз, оправдался, поскольку к 7 февраля (вплоть до этого числа показано распространение заболевания на рис. 7) обозначилось уменьшение заражения. Возможно, характер вируса влияет, хотя и в слабой степени, и на скорость распространения

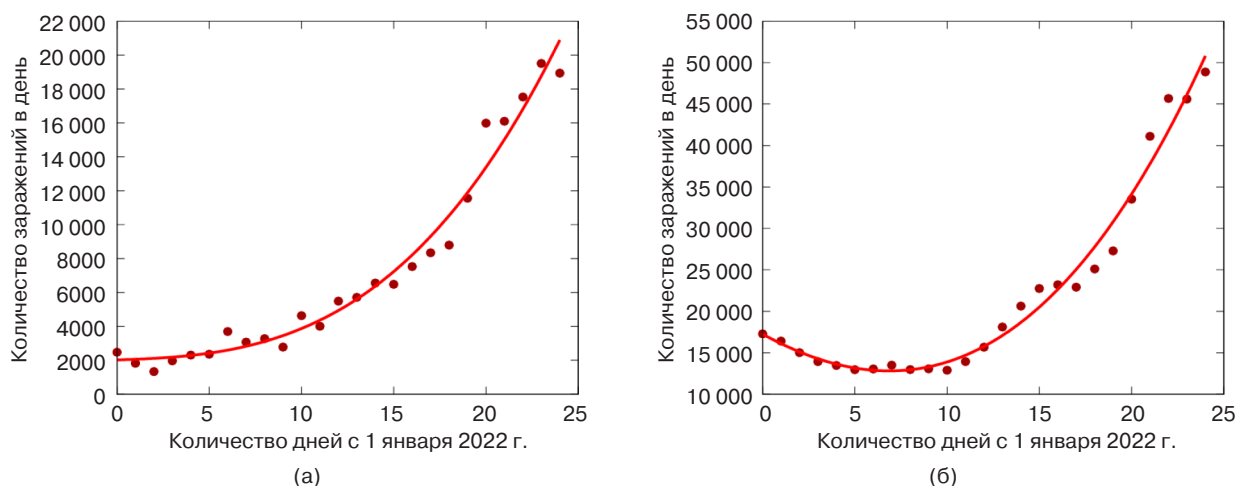


Рис. 6. Число заболеваний (абсолютная величина) в день по г. Москве (а) и по России в целом (б)

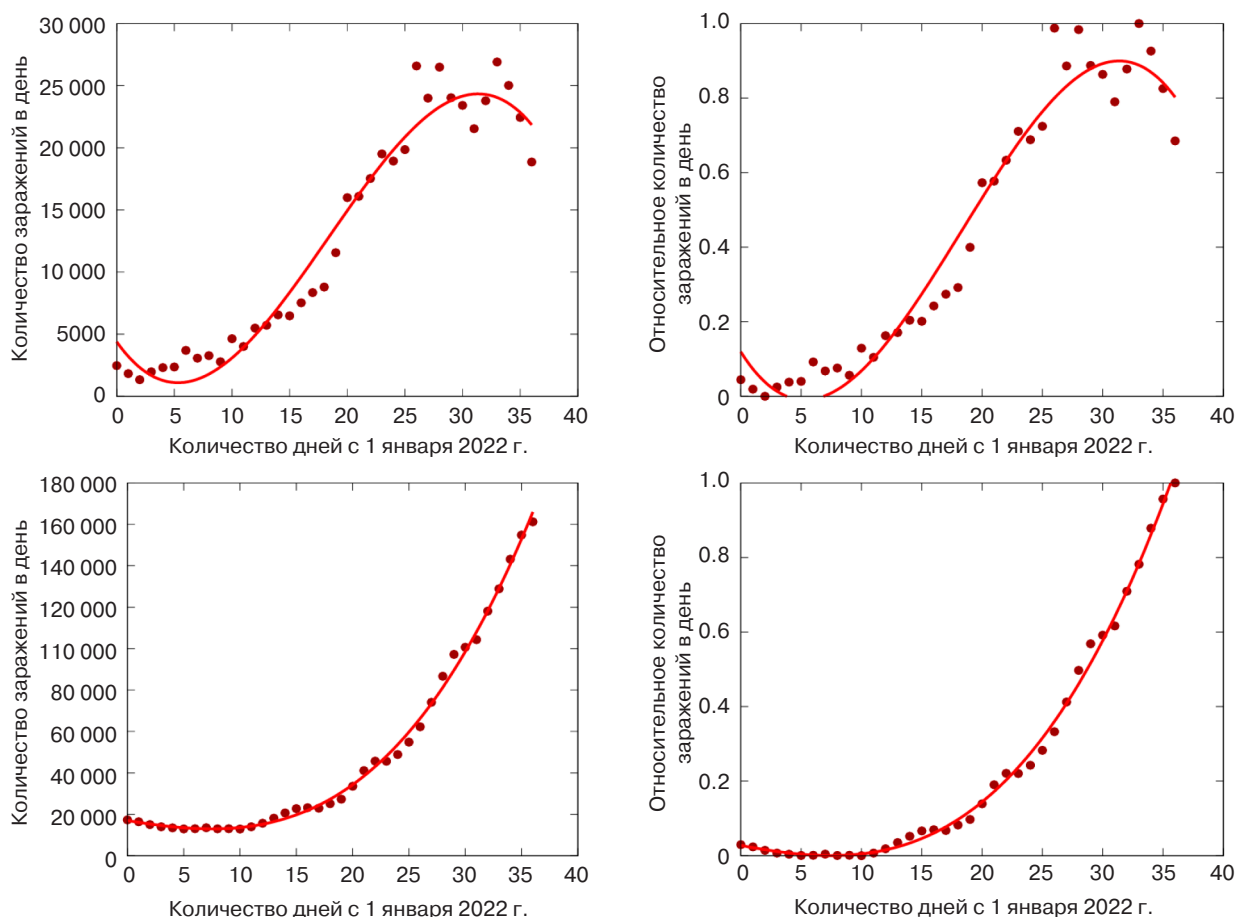


Рис. 7. Число заражений в день с 1 января по 7 февраля 2022 г., верхние рисунки относятся к г. Москве, нижние – к России без учета московских показателей

пространственной волны, несколько ускоряя ее (этот вопрос остается открытым). На рис. 7 приведены графики абсолютных приращений заражений в день, а также более показательные относительные величины заражений.

Построение относительных величин может способствовать выявлению универсальных свойств модели.

На рис. 8 представлены результаты числа заражений в день по фактическим данным на середину февраля. Видно, что, как и ожидалось, к середине февраля формируется максимум заражений по всей России. Отметим некий выброс данных, связанный с официальным пересчетом статистики. Прогноз по сдвигу в 12–14 дней подтвердился.

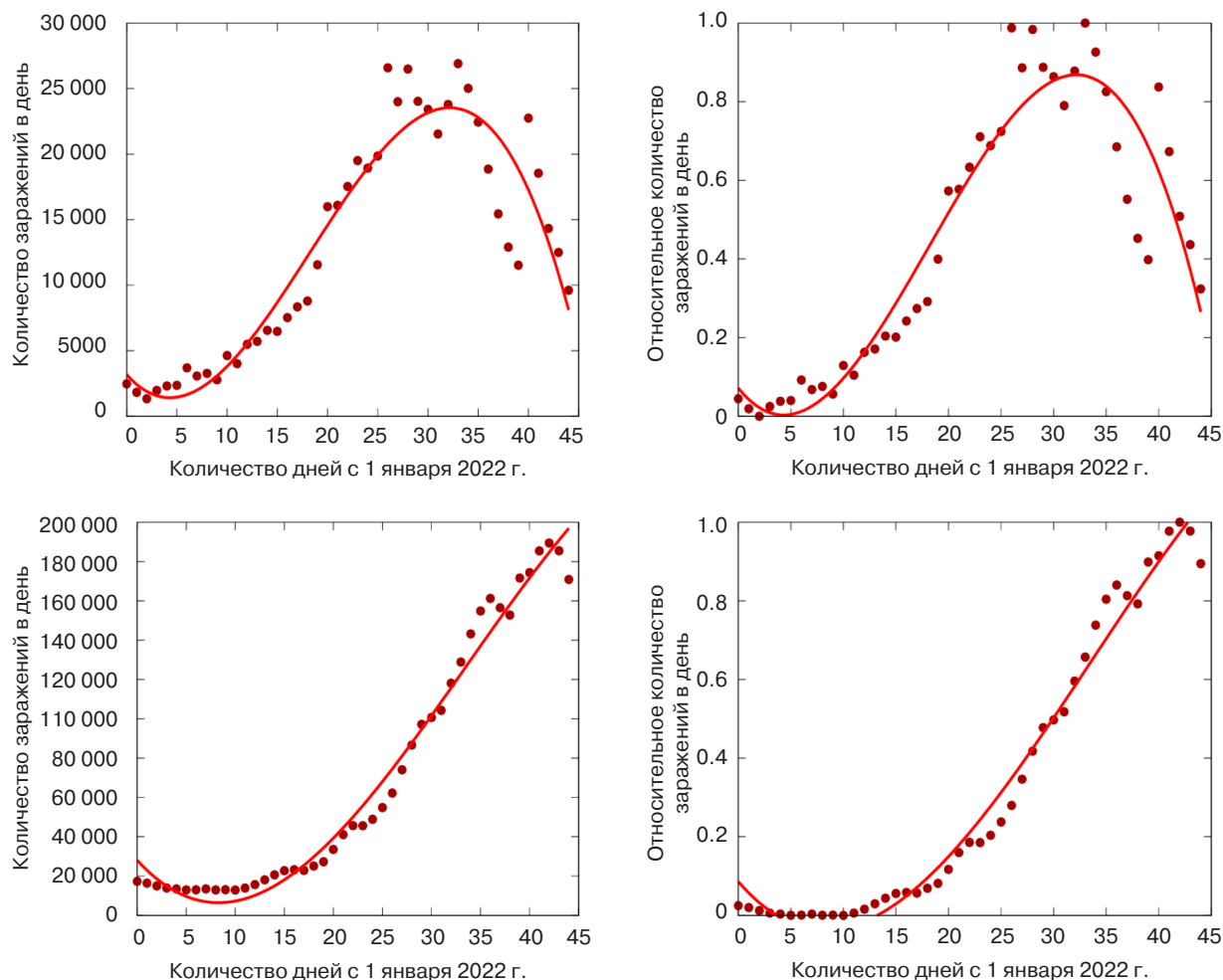


Рис. 8. Развитие пандемии в г. Москве (вверху) и России в целом без учета московских показателей (внизу) от 1 января до 15 февраля 2022 г.

Построены также распределения для г. Москвы и России в целом на конец февраля (рис. 9). Данные свидетельствуют о том, что кривые 5-й волны сформировались. Линии почти не отличаются от предыдущих, представленных на рис. 7, 8.

Таким образом, предсказания, сделанные для 5-й волны еще в конце января, подтверждаются реальными данными.

Полезно провести сравнение кривых для относительных величин в 4-й и 5-й волнах пандемии (рис. 10).

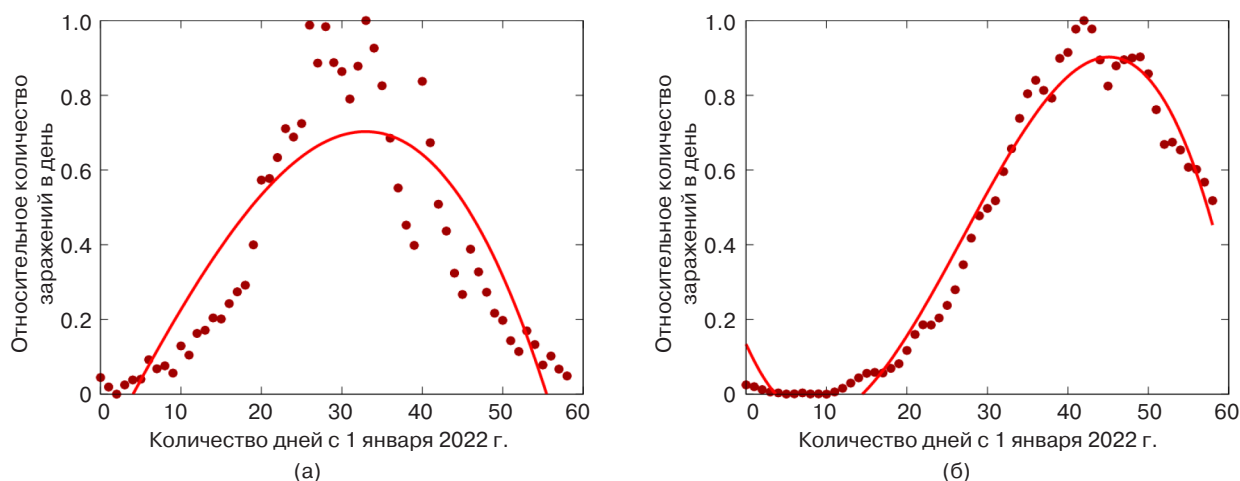


Рис. 9. Развитие пандемии в г. Москве (а) и России в целом без учета московских показателей (б) вплоть до начала марта

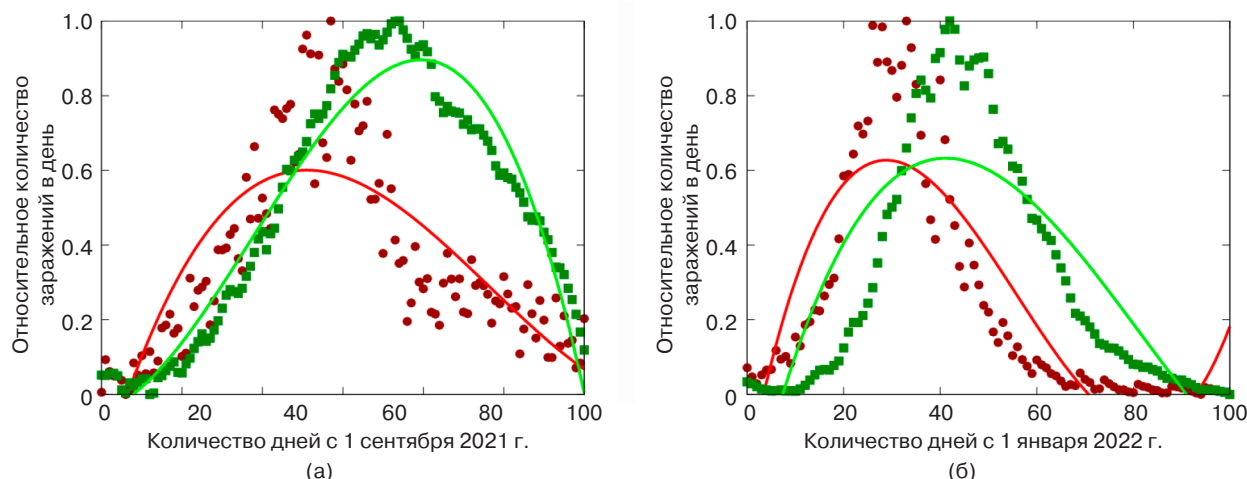


Рис. 10. Распределения количества заражений для 4-й (а) и 5-й (б) пандемических волн в г. Москве (красные линии) и России в целом без учета московских показателей (зеленые линии)

Отметим некоторые отличия в характере линий, что связано и со свойствами применяемого при осреднении МНК. Для 4-й волны гладкий график для России имеет заметно более высокий максимум, чем график для г. Москвы, что связано с относительно более медленным спадом количества заражений в России по сравнению с г. Москвой после достижения пика заражений 4-й волны. Для 5-й волны графики, построенные с помощью МНК, имеют сходную между собой амплитуду. Однако смещения по времени локальных максимумов для этих разных волн примерно одинаковы.

ПОИСК ОБЩНОСТИ ФОРМ ПАНДЕМИЧЕСКИХ ВОЛН И ИХ ИСПОЛЬЗОВАНИЕ

Наряду с величинами, определяющими смещение максимумов волн в разных пространственных точках страны, можно стараться выявить закономерности и в форме различных пандемических волн в отдельных точках, например, в г. Москве, учитывая, конечно, изменение характера заражения в новых эпидемических волнах.

Построение неких форм, реализующихся в последовательных волнах пандемии – важная задача. Отчасти она пересекается с тематикой работ [1–3].

Для этого используем полученные ранее закономерности. На рис. 11 показаны ожидаемые кривые заражения в 3-й волне с учетом данных по 2-й и 1-й волне. По ним можно судить, насколько верным будет прогноз. Согласно реальным данным для 2-й пандемической волны построена кривая на основе метода наименьших квадратов. Затем эта линия перенесена в правую часть рисунка и наложена на начало точек заражения в 3-й волне. Таким образом задается предсказание максимума заражения, которое ожидалось примерно к 15 июля 2021 г. Здесь же построена кривая МНК по полученным новым данным для 3-й волны пандемии. Согласно данному графику,

ожидаемый максимум тоже должен быть достигнут около 15 июля. На рис. 11 приведены соответствующие линии, но нанесены точки вплоть до 15 июля. Можно заключить, что такие полученные кривые примерно соответствуют прогнозам. Причем форма кривых отличается, но все же есть определенное сходство.

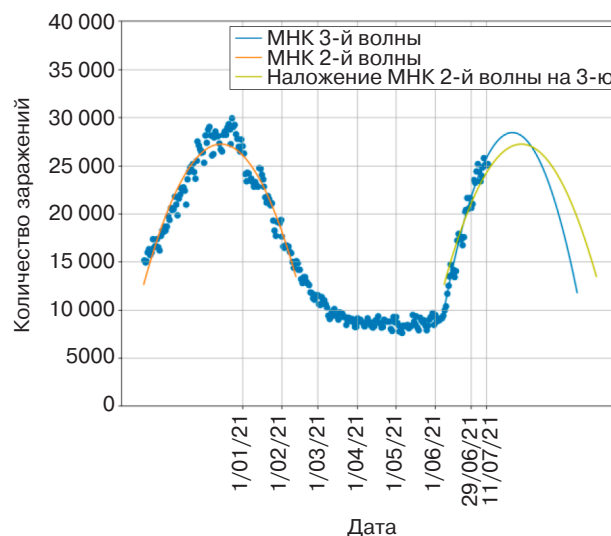


Рис. 11. Наложение кривой, полученной по 2-й волне, на ожидаемую кривую 3-ю волны вплоть до 15 июня

На рис. 12 показаны кривые для относительно-го количества заражений в зависимости от времени. Приведены реальные данные с характерными статистическими выбросами (а) и линии, осредненные по методу МНК (б). Они сдвинуты так, чтобы максимумы соответствовали нулю по времени. Для каждой волны взят день с максимальным количеством заражений и по 60 дней до и после (кроме 5-й волны). В масштабировании для каждого отрезка времени минимум оказался в 0, а максимум в 1. Можно

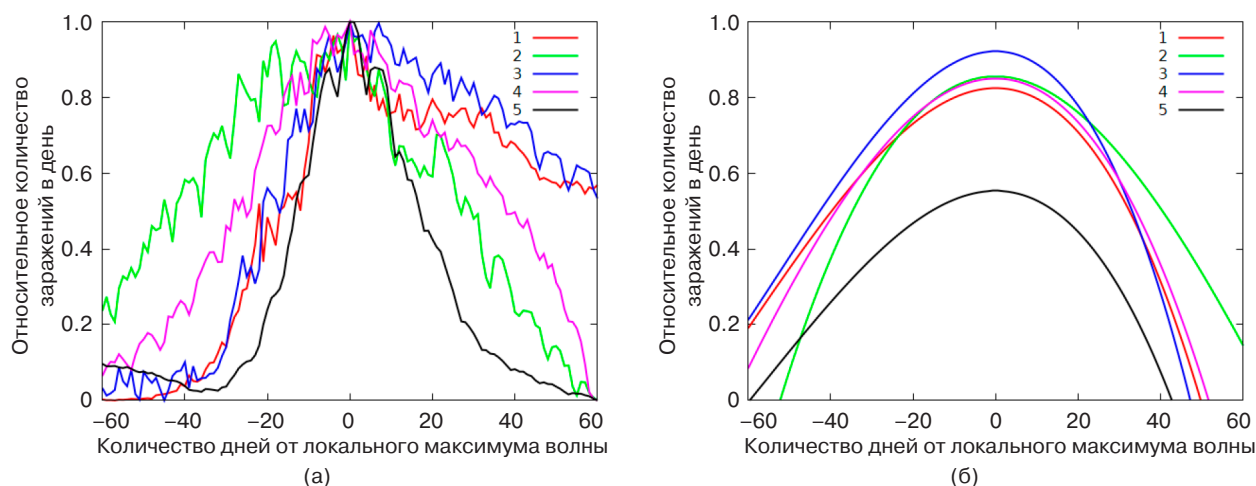


Рис. 12. Реальные данные для всех пяти волн (а) и сглаживающие кривые для них 3-й степени МНК (б)

отметить, что профили волн достаточно близко соответствуют друг другу за исключением 5-й волны, что, вероятно, связано с сильной интенсивностью распространения инфекции.

Можно отметить, что выявленное подобие волновых форм свидетельствует об определенном универсализме развития данной инфекции в России, отражающем и ситуацию в г. Москве, что позволяет с осторожностью делать прогнозы для последующих волн с учетом различной интенсивности штаммов и проведения вакцинации.

ЗАКЛЮЧЕНИЕ

В работе продемонстрирована пригодность предложенной ранее модели для описания распространения волн современной пандемии по территории России. Несмотря на различия в особенностях

новых волн, закономерности выполняются. Модель дает возможность предсказывать параметры будущих возможных волн. Сделанные в августе 2022 г. предположения для новой 6-й волны полностью подтвердились: 6-я волна развивалась аналогично предыдущим, запаздывание, равное примерно 2.5 неделям, реализовалось и здесь. Максимум заражения в день был достигнут к середине третьей декады августа, максимум заражения по России отчетливо фиксировался к середине сентября.

В настоящее время авторами изучается двумерная задача и соответствующая численная схема, что позволит учесть и другие основные центры заражения, прежде всего г. Санкт-Петербург.

Вклад авторов. Все авторы в равной степени внесли вклад в настоящую работу.

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

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Acioli P.H. Diffusion as a first model of spread of viral infection. *Am. J. Phys.* 2020;88(8):600–604. <https://doi.org/10.1119/10.0001464>
2. Bärwolff G. A Local and Time Resolution of the COVID-19 Propagation—A Two-Dimensional Approach for Germany Including Diffusion Phenomena to Describe the Spatial Spread of the COVID-19 Pandemic. *Physics*. 2021;3(3):536–548. <https://doi.org/10.3390/physics3030033>
3. Duffey R.B. Infection Waves in Pandemics and Risk Prediction: Physical Diffusion Theory and Data Comparisons. *Journal of Risk Analysis and Crisis Response*. 2021;11(2):67–74. <https://doi.org/10.2991/jracr.k.210609.001>
4. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*. 1927;115(772):700–721. <https://doi.org/10.1098/rspa.1927.0118>
5. Anderson R.M., May R.M. *Infectious Diseases of Humans*. Oxford, UK: Oxford University Press; 1992. 768 p.
6. Van Kampen N.G. *Stochastic Processes in Physics and Chemistry*. Amsterdam, The Netherlands: Elsevier North-Holland; 1981. 478 p.
7. Ivorra B., Ferrández M.R., Vela-Pérez M., Ramos A.M. Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China. *Commun. Nonlinear Sci. Numer. Simul.* 2020;88:105303. <https://doi.org/10.1016/j.cnsns.2020.105303>
8. Gross B., Zheng Z., Liu S., Chen X., Sela A., Li J., Li D., Havlin S. Spatio-temporal propagation of COVID-19 pandemics. *EPL (Europhysics Letters)*. 2020;131(5):58003. <https://doi.org/10.1209/0295-5075/131/58003>

9. Ramaswamy H., Oberai A.A., Yortsos Y.C. A comprehensive spatial-temporal infection model. *Chem. Eng. Sci.* 2021;233(100):116347. <https://doi.org/10.1016/j.ces.2020.116347>
10. Maier B., Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*. 2020;368(6492):742–746. <https://doi.org/10.1126/science.abb4557>
11. Bontempi E., Vergalli S., Squazzoni F. Understanding COVID-19 diffusion requires an interdisciplinary, multi-dimensional approach. *Environ. Res.* 2020;188:109814. <https://doi.org/10.1016/j.envres.2020.109814>
12. Neipel J., Bauermann J., Bo S., Harmon T., Jülicher F. Power-law population heterogeneity governs epidemic waves. *PLoS ONE* 2020;15(10):e0239678. <https://doi.org/10.1371/journal.pone.0239678>
13. Saffary T., Adegbeye O.A., Gayawan E., Elfaki F., Kuddus M.A., Saffary R. Analysis of COVID-19 Cases' Spatial Dependence in US Counties Reveals Health Inequalities. *Front Public Health*. 2020;12(8):579190. <https://doi.org/10.3389/fpubh.2020.579190>
14. Ascani A., Faggian A., Montresor S. The geography of COVID-19 and the structure of local economies: The case of Italy. *J. Regional Sci.* 2021;61(2):407–441. <https://doi.org/10.1111/jors.12510>
15. Prigogine I., Herman R. *Kinetic Theory of Vehicular Traffic*. New York, USA: Elsevier; 1971. 100 p.
16. Aristov V.V., Ilyin O.V. Kinetic Models for Historical Processes of Fast Invasion and Aggression. *Phys. Rev. E*. 2015;91(4):04286. <https://doi.org/10.1103/PhysRevE.91.042806>
17. Аристов В.В., Строганов А.В., Ястребов А.Д. Применение модели кинетического типа для изучения пространственного распространения COVID-19. *Доклады Российской академии наук. Физика, технические науки*. 2021;498(1):27–32. <https://doi.org/10.31857/S2686740021030020>
[Aristov V.V., Stroganov A.V., Yastrebov A.D. Application of a Kinetic Model for Studying the Spatial Spread of COVID-19. *Dokl. Phys.* 2021;66(5):129–133. <https://doi.org/10.1134/S1028335821050013>]
[Original Russian Text: Aristov V.V., Stroganov A.V., Yastrebov A. D. Application of a Kinetic Model for Studying the Spatial Spread of COVID-19. *Doklady Rossiiskoi Akademii Nauk. Fizika, Tekhnicheskie nauki*. 2021;498(1):27–32 (in Russ.). <https://doi.org/10.31857/S2686740021030020>]

Об авторах

Аристов Владимир Владимирович, д.ф.-м.н., профессор, кафедра высшей математики Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78); главный научный сотрудник, ФГУ Федеральный исследовательский центр «Информатика и управление» Российской академии наук (119333, Россия, Москва, ул. Вавилова, д. 44/2). E-mail: aristovvl@yandex.ru. Scopus Author ID 35517535600, SPIN-код РИНЦ 1481-8088, <https://orcid.org/0000-0003-2568-3453>

Строганов Андрей Валентинович, к.ф.-м.н., доцент, кафедра высшей математики Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78), E-mail: savthe@gmail.com. Scopus Author ID 36667697700, <https://orcid.org/0000-0001-6487-6093>

Ястребов Андрей Дмитриевич, аспирант, кафедра высшей математики Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78), E-mail: andr.yast711@gmail.com. Scopus Author ID 57314418000, <https://orcid.org/0000-0003-0072-6226>

About the authors

Vladimir V. Aristov, Dr. Sci. (Phys.-Math.), Professor, Department of Higher Mathematics, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia); Chief Researcher, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences (44/2, Vavilova ul., Moscow, 119333 Russia). E-mail: aristovvl@yandex.ru. Scopus Author ID 35517535600, RSCI SPIN-code 1481-8088, <https://orcid.org/0000-0003-2568-3453>

Andrey V. Stroganov, Cand. Sci. (Phys.-Math.), Assistant Professor, Department of Higher Mathematics, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: savthe@gmail.com. Scopus Author ID 36667697700. <https://orcid.org/0000-0001-6487-6093>

Andrey D. Yastrebov, Postgraduate Student, Department of Higher Mathematics, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: andr.yast711@gmail.com. Scopus Author ID 57314418000, <https://orcid.org/0000-0003-0072-6226>

Математическое моделирование
Mathematical modeling

УДК 519.857
<https://doi.org/10.32362/2500-316X-2023-11-4-72-83>



НАУЧНАЯ СТАТЬЯ

Комбинированные алгоритмы аппроксимации для интерактивного проектирования дорожных трасс в системах автоматизированного проектирования

Д.А. Карпов,
В.И. Струченков [®]

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[®] Автор для переписки, e-mail: str1942@mail.ru

Резюме

Цели. Цель работы состоит в создании алгоритмов аппроксимации последовательности точек на плоскости дугами клотоид и окружностей. Такая задача возникает в проектировании трасс железных и автомобильных дорог. План (проекция на горизонтальную плоскость) трассы дороги – это кривая (сплайн), состоящая из повторяющейся связки элементов «прямая + дуга клотоиды + дуга окружности + дуга клотоиды + ...». Такая комбинация элементов обеспечивает непрерывность не только кривой и касательной к ней, но и кривизны. Поскольку число элементов сплайна заранее неизвестно, а на их параметры накладываются ограничения, для этой задачи пока не опубликовано математически корректного алгоритма. Разработанная в РТУ МИРЭА двухэтапная схема решения задачи с определением числа элементов сплайна с помощью динамического программирования на первом этапе и оптимизацией его параметров с применением нелинейного программирования на втором, реализована только для сплайна с прямыми и окружностями (без клотоид). Ее реализация для сплайна с клотоидами много сложнее и пока не выполнена в силу ряда причин. В действующих системах автоматизированного проектирования (САПР) проектирование плана трассы выполняется в интерактивном режиме с последовательным подбором элементов. В этой связи имеет смысл разработка математически корректных алгоритмов поэлементной аппроксимации.

Метод. Задача поэлементной аппроксимации окружностью или клотоидой формализована как задача нелинейного программирования малой размерности. Целевая функция – сумма квадратов отклонений от исходных точек. Поскольку клотоида в декартовых координатах представляется степенными рядами, возникают трудности вычисления производных целевой функции по искомым параметрам элементов сплайна. Предложен математически корректный алгоритм вычисления этих производных на основе интегрального представления декартовых координат точек клотоиды как функций ее длины.

Результаты. Разработаны математическая модель и алгоритмы аппроксимации последовательности точек на плоскости клотоидой и окружностью с применением метода нелинейного программирования. Реализован алгоритм второго порядка с вычислением и обращением матрицы вторых производных (матрица Гессе).

Выводы. Для аппроксимации окружностью и клотоидой с применением нелинейного программирования не обязательно иметь аналитическое выражение целевой функции через искомые переменные. Предложенные алгоритмы позволяют вычислять не только первые, но и вторые производные в отсутствие таких выражений.

Ключевые слова: план трассы, сплайн, нелинейное программирование, клотоида, целевая функция, градиент, матрица Гессе

• Поступила: 20.12.2022 • Доработана: 12.04.2023 • Принята к опубликованию: 14.05.2023

Для цитирования: Карпов Д.А., Струченков В.И. Комбинированные алгоритмы аппроксимации для интерактивного проектирования дорожных трасс в системах автоматизированного проектирования. *Russ. Technol. J.* 2023;11(4):72–83. <https://doi.org/10.32362/2500-316X-2023-11-4-72-83>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Combined approximation algorithms for interactive design of road routes in CAD

Dmitry A. Karpov,
Valery I. Struchenkov @

MIREA – Russian Technological University, Moscow, 119454 Russia

@ Corresponding author, e-mail: str1942@mail.ru

Abstract

Objectives. The aim of the work is to create algorithms for approximating a sequence of points on a plane by arcs of clothoids and circles. Such a problem typically arises in the design of railroad and highway routes. The plan (projection onto a horizontal plane) of the road route is a curve (spline) consisting of a repeating bundle of elements “straight line + clothoid arc + circle arc + clothoid arc + ...”. Such a combination of elements provides continuity not only for the curve and its tangent, but also for the curvature. Since the number of spline elements is not known in advance, and their parameters are subject to restrictions, there is no mathematically consistent algorithm for this problem. The two-stage scheme for solving the problem is developed at RTU MIREA only for a spline with lines and circles (i.e., without clothoid elements). At the first stage, the scheme uses dynamic programming to determine the number of spline elements. At the second stage, the scheme optimizes parameters of the spline using nonlinear programming. This scheme has yet to be implemented for a spline with clothoids due to a significantly more complicated nature of this problem. Therefore, the design of route plans in existing computer aided design (CAD) systems is carried out in interactive mode using iterative selection of elements. In this regard, it makes sense to develop mathematically consistent algorithms for element-by-element approximation.

Methods. The problem of element-by-element approximation by a circle and a clothoid is formalized as a low-dimensional non-linear programming problem. The objective function is the sum of squared deviations from the original points. Since a clothoid can only be represented in Cartesian coordinates by power series, there are difficulties in calculating the derivatives of the objective function with respect to the desired parameters of the spline elements. The proposed mathematically consistent algorithm for calculating these derivatives is based on the integral representation of the Cartesian coordinates of the points of the clothoid as functions of its length.

Results. A mathematical model and algorithms have been developed for approximating a sequence of points on a plane by clothoids and circles using the method of nonlinear programming. A second-order algorithm is implemented with the calculation and inversion of the matrix of second derivatives (Hesse matrix).

Conclusions. For approximation by circles and clothoids using nonlinear programming, it is not necessary to have an analytical expression of the objective function in terms of the required variables. The proposed algorithms make it possible to calculate not only the first, but also the second derivatives in the absence of such expressions.

Keywords: route plan, spline, non-linear programming, clothoid, objective function, gradient, Hessian matrix

• Submitted: 20.12.2022 • Revised: 12.04.2023 • Accepted: 14.05.2023

For citation: Karpov D.A., Struchenkov V.I. Combined approximation algorithms for interactive design of road routes in CAD. *Russ. Technol. J.* 2023;11(4):72–83. <https://doi.org/10.32362/2500-316X-2023-11-4-72-83>

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

Задача аппроксимации кривой, заданной последовательностью точек на плоскости, окружностью, а тем более клотоидой, много сложнее, чем широко применяемая линейная или параболическая регрессия. В случае поиска аппроксимирующей окружности в декартовых координатах задача сводится к нелинейной системе уравнений, а в случае клотоиды не удастся получить даже этого, т.к. клотоида не представляется в виде $y = f(x)$. Математически корректного алгоритма аппроксимации клотоидой в литературе найти не удалось.

Вместо этого в действующих системах автоматизированного проектирования (САПР) используется метод перебора назначаемых проектировщиком вариантов или аппроксимация по методу эвольвент. Последний был предложен еще в докомпьютерную эпоху [1] при вычислении сдвижек плана трассы для приведения его в заданное проектное положение. Если проектное положение неизвестно, как это бывает при реконструкции плана трассы железной дороги, с изменением его параметров, таких как длины переходных кривых (дуг клотоид) или радиусов круговых кривых, то это проектное положение тем или иным способом задается проектировщиком, после чего компьютер вычисляет все проектные параметры плана трассы и сдвижки существующего пути. Метод эвольвент использовался и для приближенного решения задачи аппроксимации вместо ее точного решения в декартовых координатах¹. Этот метод успешно применялся и для проектирования плана трассы как единого целого в пределах перегона еще в 80-х гг. прошлого века на таких малоприспособленных для создания САПР ЭВМ, как Минск 32 и ЕС 1033¹. Уход от декартовых координат представлялся вынужденным. При этом было известно, что при наличии больших углов поворота и малых радиусов погрешность метода весьма существенна, особенно в случае сильно сбитых кривых и соответственно больших сдвижек или при проектировании

реконструкции с существенным изменением параметров плана трассы. Метод эвольвент использовался в 80-х гг. XX века и для оптимизации плана трассы новой железной дороги на участках напряженного хода [2, 3]. План трассы представлялся в виде ломаной линии для поиска числа элементов и их примерного расположения, которая затем преобразовывалась к требуемому виду с применением этого метода с последующей оптимизацией параметров полученного сплайна [3]. Широкое применение метода эвольвент для решения различных задач, несмотря на его недостатки, объясняется тем, что в отличие от использования декартовых координат, он позволяет работать с однозначными функциями, имеющими простые аналитические выражения: параболой второй степени вместо окружности и кубической параболой вместо клотоиды.

В работах зарубежных авторов предлагались различного рода эвристические алгоритмы [4–12]. Первоначально они основывались на анализе таких характеристик исходной ломаной, получаемой после соединения смежных точек съемки отрезками прямых, как углы поворота в вершинах ломаной и графики кривизны, затем перешли к использованию генетических алгоритмов. В нашей стране в действующих САПР используются различные программы, облегчающие поэлементный подбор и оценку вариантов трассы. Наиболее успешно это реализовано в системе *Robur* фирмы Топоматик², в которой используется полуавтоматический метод подбора параметров плана существующей трассы, основанный на графике кривизны. В последние годы периодически появляются сообщения о том, что задача проектирования реконструкции плана трассы железной дороги решена, но алгоритмы полностью не публикуются. Из недавно опубликованных эвристических алгоритмов иностранных авторов следует отметить представленные в работах китайского профессора Хао Пу с коллегами [13–15].

Трудно поверить в оптимальное решение задачи для нескольких круговых и переходных

¹ Методические рекомендации по расчету составных кривых плана трассы железных дорог. М.: ВНИИ транспортного строительства; 1985. 26 с. [*Methodical recommendations for the calculation of composite curves of the railroad route plan*. Moscow: All-Russian Research Institute of Transport Construction; 1985. 26 p. (in Russ.).]

² Документация по продуктам Топоматик Robur «Подбор пути с помощью графика кривизны». [Topomatic Robur product documentation “Path selection via curvature graph” (in Russ.).] http://help.topomatic.ru/v6/doku.php?id=rail:rail:tasks:selection_path:start. Дата обращения 18.04.2023. / Accessed April 18, 2023 (in Russ.).

кривых (клотоид), если нет алгоритма оптимального решения для одной клотоиды.

Рассматриваемая задача, как и проектирование плана и продольного профиля дорог в целом, характерна тем, что в сложных случаях оптимальное решение неизвестно, и в условиях отсутствия заинтересованности проектировщиков в его достижении, все «правдоподобное», что выдает компьютер, можно называть оптимальным. В этом ее принципиальное отличие от задач моделирования геометрических форм автомобильных дорог [16].

Тем не менее, разработка математически корректных алгоритмов аппроксимации окружностью и клотоидой остается актуальной как в теоретическом, так и в практическом плане, т.к. результат может быть полезен не только при проектировании дорог. Актуальный уровень развития вычислительной техники позволяет решить эту задачу в приемлемое время на общедоступных компьютерах, не используя паллиативные алгоритмы.

Цель настоящей статьи состоит в изложении математически корректных алгоритмов аппроксимации окружностью и клотоидой в декартовых координатах с применением алгоритмов нелинейного программирования, использующих метод эвольвент только для получения начального приближения с последующей оптимизацией параметров, определяющих положение окружности или клотоиды.

ПОСТАНОВКА ЗАДАЧИ

Для заданной последовательности точек на плоскости найти клотоиду (окружность), такую, что сумма квадратов расстояний h_i ($i = 1, 2, \dots, n$) от заданных точек до клотоиды (окружности) минимальна. Расстояния вычисляются по нормали от заданной точки к клотоиде (окружности).

Начальная точка искомой кривой задана. Может быть задано и направление касательной к искомой кривой в начальной точке и ограничение по минимальному и максимальному радиусам кривизны искомой кривой.

Итак, целевая функция:

$$F(\mathbf{h}) = \frac{1}{2} \sum_{i=1}^n h_i^2 \rightarrow \min, \quad (1)$$

где $\mathbf{h}(h_1, h_2, \dots, h_n)$ – вектор неизвестных, n – их количество.

МЕТОД ЭВОЛЬВЕНТ

В отличие от других кривых (парабол, окружностей, синусоид и т.д.) говорить об эвольвенте нельзя без указания другой кривой (эволюты), которая

порождает эвольвенту. В литературе можно встретить различные определения понятия «эвольвента». В математике эвольвента – это *кривая*, для которой заданная эволюта является геометрическим местом центров кривизны [17]. Следовательно, нормаль в каждой точке к эвольвенте является касательной к эволюте. Именно касательной, а не секущей, т.к. в каждой точке кривой может быть только один центр кривизны. В этой связи эвольвентой называют кривую, которую описывает конец гибкой нерастяжимой нити, сматываемой с эволюты (например, с окружности).

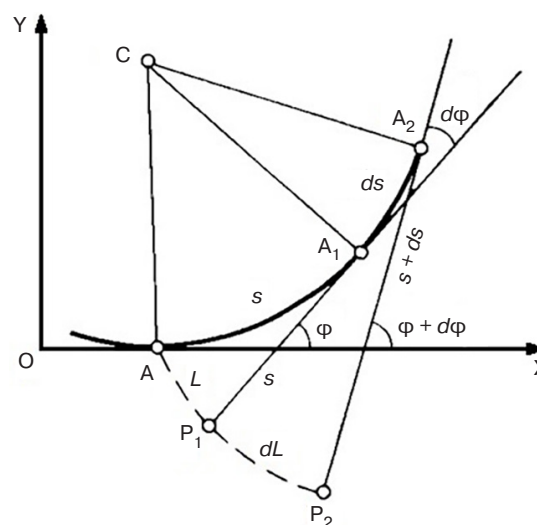


Рис. 1. Построение эвольвенты окружности.
 L – длина эвольвенты

Построение эвольвенты окружности показано на рис. 1. Здесь A – начальная точка на эволюте, A_1 и A_2 – ее новые положения на эволюте. Им соответствуют длины дуг, отсчитываемых от начальной точки, s и $s + ds$. Эти длины отложены по касательным в сторону, противоположную движению от начальной точки. Получены точки на эвольвенте P_1 и P_2 , соответственно. Длины дуг эвольвенты – AP_1 и AP_2 . При полном обороте ($\varphi = 2\pi$) длина эвольвенты $\mathcal{E} = 2\pi R$, где R – радиус окружности и т.д., т.е. фиксируется не начальная точка A , а ее новые положения на эволюте, в которых строятся касательные и на них разворачиваются дуги эволюты. Эвольвентой окружности является разворачивающаяся спираль.

Однако в [1, 18, 19] излагается другая трактовка понятия эвольвенты окружности и способ ее построения (рис. 2). Фиксируется точка A начала круговой кривой и для каждой точки окружности (a) длина дуги от начальной точки до текущей (Aa) разворачивается на касательную от точки A . Таким образом для каждой точки окружности получается своя эвольвента (aa' , bb' ...). Кривая aa' не является эвольвентой окружности, но ее длина равна длине

соответствующего куска эвольвенты. Отсюда и возникла путаница понятий.

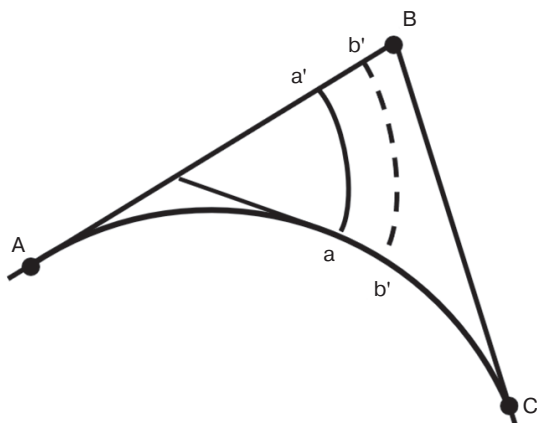


Рис. 2. Построение эвольвенты окружности по [1, 18, 19]

Если в [1, 18] эвольвентой окружности называются построенные описанным способом *кривые*, то в [19, с. 243, рис. 7.20] читаем: «*Эвольвентой* (курсив авторов) называется *длина дуги aa'*, которую опишет конец нерастяжимой нити, натянутой на окружность Aa и закрепленной в точке A, при выпрямлении нити».

Если принять, что эвольвента не *кривая*, о которой шла речь выше, а *длина дуги aa'*, то эта длина равна $K^2/(2R)$, где K – длина дуги Aa, а R – радиус окружности, и эвольвента окружности – это не спираль в декартовых координатах, а парабола второй степени как функция длины эволюты.

Если следовать способу построения эвольвенты окружности принятому в [1, 18, 19], то для каждой точки окружности (эволюты) получится своя эвольвента, и все они будут заканчиваться на касательной к окружности в начальной точке, которая фиксирована. Настоящую эвольвенту окружности следует строить, как показано на рис. 1.

Из рис. 1 следует $dL = \varphi(s)ds$, т.к. $dL = sd\varphi = sds/R = \varphi ds$, и далее:

$$L(S) = \int_0^S \varphi(s)ds. \quad (2)$$

Здесь S – длина эволюты от начальной точки, $\varphi(s)$ – зависимость угла касательной к эволюте с осью OX (или любым заданным направлением, что не принципиально), называемая угловой диаграммой (углограммой). Понятие «эвольвента» определялось для гладких кривых и формула (2) верна для всех таких кривых. Но в нашей задаче (1) исходная кривая – это ломаная линия и применять к ней приведенные выше определения эвольвенты, т.е. рассматривать ломаную как эволюту, нельзя.

Однако формулу (2) можно обобщить, если считать эвольвентой кривую, длина которой вычисляется через углограмму эволюты по формуле (2), т.е. длина эвольвенты – это площадь углограммы как функция длины исходной кривой (в частности, ломаной). Сама по себе эвольвента для рассматриваемой задачи не представляет интереса. Важно, что разность длин эвольвент двух эволют, имеющих общую точку, при некоторых дополнительных условиях приближенно равна расстоянию между эволютами по нормали. В нашей задаче первая эвольвента – это ломаная, соединяющая аппроксимируемые точки, а вторая – проектная кривая (окружность или клотоида). Отсюда и происходит метод эвольвент, который состоит в следующем:

1. Вычисляем углы исходной ломаной с осью OX и значения длины эвольвенты L_i исходной ломаной для точек съемки $i = 1, 2, \dots, n$. $L_1 = 0$, далее

$$L_i = \sum_{j=2}^i \varphi_j s_j. \quad (3)$$

Здесь φ_j – угол с осью OX j -го звена ломаной (от точки $j-1$ до точки j), s_j – длина этого звена.

Для окружности угол поворота линейно зависит от длины дуги, соответственно длина проектной эвольвенты $L_{пр}(S)$ – парабола второй степени.

2. Если проектная кривая задана, то разность $L_{пр}(S) - L_{исх}(S)$ приближенно дает искомую сдвигку h_j , где j и S должны соответствовать друг другу.
3. Если же требуется решить задачу аппроксимации, то неизвестные коэффициенты искомой параболы (a и b) определяются по методу наименьших квадратов.

Для окружности *длина* эвольвенты $L_{пр}(S) = aS^2 + bS$, где S , как и ранее, длина дуги от заданной начальной точки. Смысл искомых параметров следующий: $a = 1/(2R)$ и b – угол с осью OX касательной к окружности в начальной точке.

Получаем задачу:

$$F(a, b) = \sum_{i=2}^n (aS_i^2 + bS_i - L_i)^2 \rightarrow \min. \quad (4)$$

Условия $\frac{\partial F}{\partial a} = 0$ и $\frac{\partial F}{\partial b} = 0$ дают систему двух линейных уравнений, из которой находим a и b и затем $R = 1/(2a)$. Угол b может быть задан, и тогда останется неизвестным только a .

При аппроксимации клотоидой возможны варианты: клотоида сопрягает прямую с окружностью; окружность с прямой; две окружности разного радиуса.

Если длины сопрягающих клотоид существенно меньше длин круговых кривых, то важно подобрать положение окружности, а затем «пристроить клотоиды». Однако возможна и обратная ситуация, при которой клотоида является самостоятельным элементом. Например, в проектировании автодорог есть термин «клотоидное трассирование» [20]. Не следует думать, что в этом случае трасса состоит из одних клотоид и прямых, но клотоиды в процентном отношении по длине преобладают.

По определению кривизна клотоиды $\sigma = \sigma_0 + kS$, где σ_0 – кривизна в начальной точке (при длине $S=0$), k – скорость изменения кривизны. Текущий угол касательной с осью ОХ $\varphi = \varphi_0 + \sigma_0 S + kS^2/2$, где φ_0 – начальный угол касательной с осью ОХ. Наконец, длина эвольвенты клотоиды равна

$$L = \varphi_0 S + \sigma_0 S^2/2 + kS^3/6.$$

Аппроксимация клотоидой по методу эвольвент в общем случае сводится к задаче:

$$F(\varphi_0, \sigma_0, k) = \sum_{i=2}^n \left(\varphi_0 S_i + \frac{\sigma_0 S_i^2}{2} + \frac{k S_i^3}{6} - L_i \right)^2 \rightarrow \min. \quad (5)$$

Здесь S_i и L_i имеют тот же смысл, что и в задаче аппроксимации окружностью (4). Одна или две переменные могут быть фиксированы. В любом случае задача (5) решается просто. Дифференцируя (5) по искомым переменным и приравнявая производные к нулю, получаем систему линейных уравнений (при двух фиксированных неизвестных – одно уравнение). После решения системы нужно вычислить кривизну в конечной точке клотоиды $\sigma_n = \sigma_0 + kS_n$.

Если $R_n = 1/\sigma_n$ окажется вне допустимых пределов, надо принять $R_n = R_{\text{пред}}$ и при заданном σ_0 вычислить $k = (1/R_{\text{пред}} - \sigma_0)S_n$. Далее неизвестное φ_0 получаем, подставляя в (5) найденное k и решая задачу с одним неизвестным. Если же σ_0 не задана, то подставляем в (5) $\sigma_0 = \sigma_n - kS_n$ и решаем задачу (5) с оставшимися неизвестными и найденным σ_n .

Таким образом, использовать метод эвольвент весьма просто для аппроксимации как окружностью, так и клотоидой. К сожалению, точность метода может оказаться недостаточной.

Для проверки этого утверждения в качестве координат исходных точек съемки вычислялись декартовы координаты концов хорд заданной длины на заданной окружности. По этим координатам вычислялись углы хорд с осью ОХ и длины эвольвенты по формуле (3). Далее решалась задача (4). Очевидно, что ее решением являются переменные,

соответствующие исходной окружности. При этом отклонения всех исходных точек от аппроксимирующей окружности должны быть равны нулю. Однако этого не получалось. Если при радиусе окружности $R = 500$ м и более, хордах по $l_x = 20$ м и длине окружности $S \leq 500$ м, т.е. углах поворота менее 1 рад отклонения получаемой окружности от исходной были менее 0.01 м, то при $R = 200$ м, $l_x = 20$ м, $S = 200$ м радиус по методу эвольвент получался 199.9167 м вместо 200 м и максимальное отклонение $D_{\text{max}} = 0.0383$ м вместо 0. При $S = 400$ м (угол поворота 2 рад) для той же окружности $D_{\text{max}} = 0.1180$ м вместо 0. При этом во всех случаях разности длин проектной и исходной эвольвент аппроксимируемой окружности во всех точках съемки были равны нулю. При $l_x = 10$ м точность метода существенно выше. Так при $R = 300$ м, $S = 200$ м при разбивке по 10 м получаем $R = 299.9861$ и $D_{\text{max}} = 0.00297$, а при разбивке той же окружности по 20 м получим $R = 299.9444$ м и $D_{\text{max}} = 0.01190$ м.

Аналогичные расчеты были выполнены применительно к использованию метода эвольвент для аппроксимации клотоидой. Использовались дуги равной длины, а не стягивающие их хорды, что упрощает вычисление декартовых координат точек по концам дуг. Эти точки рассматривались как точки съемки, вычислялись длины хорд, и, как и для окружностей, вычислялись углы и длины эвольвенты L_i по формуле (3). Далее решалась задача (5) при $\sigma_0 = 0$ и неизвестных φ_0 и k .

Как и следовало ожидать, метод эвольвент хорошо работает при малых значениях параметра k и коротких клотоидах. В противном случае результаты неудовлетворительны.

Пример 1. Клотоида длиной 400 м разбита на 20 равных частей через 20.0 м; $k = 3.333333 \cdot 10^{-5}$. Получено $\varphi_0 = -0.00027412$ вместо 0 и $k = 3.337930 \cdot 10^{-5}$ вместо исходного $k = 3.333333 \cdot 10^{-5}$.

Аппроксимация длины эвольвенты как функции длины ломаной кубической параболы (решение задачи 5) выполнена с отклонениями не более 0.041 м. Однако максимальное отклонение исходных точек от полученной клотоиды по нормали к ней $D_{\text{max}} = 0.255$ м.

Пример 2. Та же задача, но длина клотоиды 200 м.

Аппроксимация кубической параболой выполнена с отклонениями не более 0.001 м. Получено $\varphi_0 = -1.839272551 \cdot 10^{-5}$ вместо 0 и $k = 3.334527 \cdot 10^{-5}$ вместо исходного $k = 3.333333 \cdot 10^{-5}$, $D_{\text{max}} = 0.048$ м.

Следует отметить, что в этих расчетах показано не столько как метод эвольвент находит оптимум, а скорее – как он уходит от него. Для «сильно сбитой» исходной трассы отклонения от оптимума решений, получаемых по методу эвольвент, могут

быть существенно больше, особенно при малых радиусах и больших углах поворота. Однако в любом случае эти решения можно использовать как начальные приближения для оптимальной аппроксимации окружностью.

ОПТИМИЗАЦИЯ НАЧАЛЬНОГО ПРИБЛИЖЕНИЯ ПРИ АППРОКСИМАЦИИ ОКРУЖНОСТЬЮ

Будем считать, что заданы координаты точек, подлежащих аппроксимации окружностью, и ее начальная точка А. Обычно фиксируют и направление касательной в этой точке (угол α с осью ОХ), но мы без ограничения общности будем считать неизвестными этот угол и радиус окружности R . Координаты центра окружности (рис. 3):

$$x_c = x_a - R \cdot \sin \alpha \text{ и } y_c = y_a + R \cdot \cos \alpha. \quad (6)$$

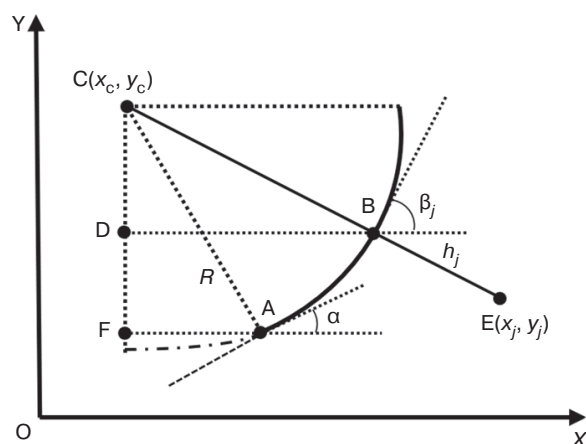


Рис. 3. К вычислению производных отклонений h_j от окружности

Отклонение h_j (BE на рис. 3) произвольной точки съемки с координатами x_j, y_j от окружности

$$h_j = \sqrt{(x_j - x_c)^2 + (y_j - y_c)^2} - R. \quad (7)$$

Используя (6) и (7), получаем производные:

$$\frac{\partial h_j}{\partial R} = \cos(\beta_j - \alpha) - 1 \text{ и } \frac{\partial h_j}{\partial \alpha} = R \sin(\beta_j - \alpha).$$

Здесь β_j – угол касательной к окружности в точке В. Далее вычисляем градиент целевой функции (1):

$$\begin{aligned} \frac{\partial F}{\partial R} &= \sum_{j=1}^n h_j \cdot \frac{\partial h_j}{\partial R} = \sum_{j=1}^n h_j (\cos(\beta_j - \alpha) - 1), \\ \frac{\partial F}{\partial \alpha} &= \sum_{j=1}^n h_j \cdot \frac{\partial h_j}{\partial \alpha} = R \sum_{j=1}^n h_j \sin(\beta_j - \alpha). \end{aligned} \quad (8)$$

Задача аппроксимации окружностью с целевой функцией (1) сведена к двумерной задаче минимизации с ограничением на R и начальным приближением, полученным по методу эволюент в результате решения задачи (4). Наличие хорошего начального приближения особенно важно, т.к. нет оснований полагать, что задача одноэкстремальна.

Можно использовать различные методы для решения задачи [21–23], в частности градиентные или покоординатный спуск, меняя поочередно R и α .

Отметим, что если фиксированы начальная точка А и угол α , то задание еще одной точки окружности (например, конечной) определяет окружность однозначно, и перебор нескольких таких точек с вычислением значения целевой функции для каждой из них может оказаться достаточным, особенно при больших радиусах, когда метод эволюент дает приемлемый результат и нужно убедиться в этом, пытаясь его улучшить.

ОПТИМИЗАЦИЯ НАЧАЛЬНОГО ПРИБЛИЖЕНИЯ ПРИ АППРОКСИМАЦИИ КЛОТОИДОЙ

Как и при аппроксимации окружностью, при аппроксимации клотоидой основной вопрос состоит в вычислении производных расстояний заданных точек съемки от клотоиды по параметрам, ее определяющим. Покажем, как решается эта задача на примере перехода от прямой к окружности с использованием клотоиды.

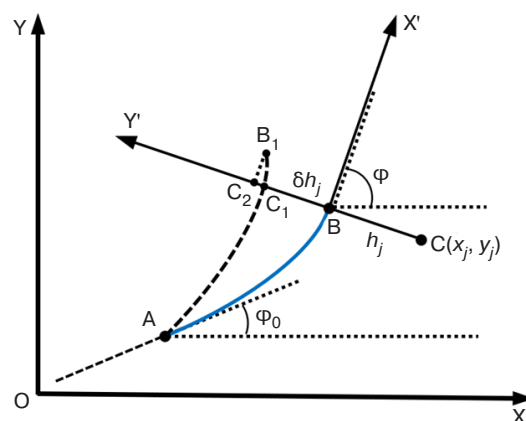


Рис. 4. К вычислению производных отклонений h_j от клотоиды

На рис. 4 АВ – исходное положение куска клотоиды; АВ₁ – его новое положение при изменении одного из параметров клотоиды; ϕ_0 и ϕ углы касательной с осью ОХ; СВ = h_j – отклонение точки съемки С от клотоиды; ВС₁ – его приращение.

Обозначим приращения координат точки В при ее переходе в новое положение В₁ из-за изменения какого-либо из параметров клотоиды через δx_B и δy_B . В системе координат с центром в точке В и

осями, направленными по касательной и нормали к клотоиде соответственно, координата y' точки B_1 , т.е. $BC_2 = \delta y_B \cos \varphi - \delta x_B \sin \varphi$. Величиной $C_1 C_2$ можно пренебречь по сравнению с $C_1 B$ и считать, что линейная часть приращения

$$\delta h_j = \delta y_B \cos \varphi - \delta x_B \sin \varphi. \quad (9)$$

Поскольку начальная точка A фиксирована, то параметрами, определяющими положение клотоиды, а значит и качество аппроксимации, являются в общем случае φ_0 – угол касательной с осью OX в начальной точке (при $L = 0$), σ_0 – начальная кривизна и k – параметр клотоиды (скорость изменения кривизны, иначе говоря – производная кривизны по длине).

В случае перехода с прямой на окружность $\sigma_0 = 0$, остаются только две переменных φ_0 и k , и фактически задача сводится к вычислению производных координат точек клотоиды по этим переменным, т.к. еще до вычисления производных с использованием итерационного алгоритма [24] вычисляются координаты точек пересечения с клотоидой нормалей из каждой точки съемки, длины клотоиды L_j от начальной точки до точек пересечения и углы касательных с осью OX .

Декартовы координаты точек клотоиды как функции ее длины выражаются степенными рядами, но мы воспользуемся их интегральным представлением

$$\begin{aligned} x_B &= x_A + \int_0^L \cos \left(\varphi_0 + \frac{kt^2}{2} \right) dt, \\ y_B &= y_A + \int_0^L \sin \left(\varphi_0 + \frac{kt^2}{2} \right) dt. \end{aligned} \quad (10)$$

Дифференцируя (10), получаем

$$\begin{aligned} \frac{\partial x_B}{\partial \varphi_0} &= - \int_0^L \sin \left(\varphi_0 + \frac{kt^2}{2} \right) dt = -(y_B - y_A), \\ \frac{\partial y_B}{\partial \varphi_0} &= \int_0^L \cos \left(\varphi_0 + \frac{kt^2}{2} \right) dt = (x_B - x_A), \\ \frac{\partial x_B}{\partial k} &= - \int_0^L \sin \left(\varphi_0 + \frac{kt^2}{2} \right) \frac{t^2}{2} dt = \\ &= \frac{1}{2k} \int_0^L t dt \cos \left(\varphi_0 + \frac{kt^2}{2} \right) = \frac{L \cos \varphi - (x_B - x_A)}{2k}, \\ \frac{\partial y_B}{\partial k} &= \int_0^L \cos \left(\varphi_0 + \frac{kt^2}{2} \right) \frac{t^2}{2} dt = \\ &= \frac{1}{2k} \int_0^L t dt \sin \left(\varphi_0 + \frac{kt^2}{2} \right) = \frac{L \sin \varphi - (y_B - y_A)}{2k}. \end{aligned}$$

Используя (9), получаем искомые производные отклонений точек съемки от клотоиды:

$$\begin{aligned} \frac{\partial h_j}{\partial \varphi_0} &= (x_B - x_A) \cos \varphi + (y_B - y_A) \sin \varphi, \\ \frac{\partial h_j}{\partial k} &= \frac{(x_B - x_A) \sin \varphi - (y_B - y_A) \cos \varphi}{2k}. \end{aligned} \quad (11)$$

Здесь A – начальная точка клотоиды, h_j – отклонение произвольной j -й точки съемки от клотоиды по нормали к ней, B – точка пересечения этой нормали с клотоидой, φ – угол касательной в этой точке с осью OX , L – длина клотоиды от точки A до точки B .

Формулы (11) позволяют найти градиент целевой функции (1)

$$\frac{\partial F}{\partial k} = \sum_{j=1}^n h_j \cdot \frac{\partial h_j}{\partial k}; \quad \frac{\partial F}{\partial \varphi_0} = \sum_{j=1}^n h_j \cdot \frac{\partial h_j}{\partial \varphi_0}. \quad (12)$$

Далее можно использовать градиентные методы оптимизации, например, метод сопряженных градиентов [22–24]. Однако формулы (10)–(12) позволяют вычислить матрицу вторых производных целевой функции $\mathbf{G}$ (матрицу Гессе) и применить более эффективный метод второго порядка [23, 24]:

$$\begin{aligned} \frac{\partial^2 F}{\partial k^2} &= \sum_{j=1}^n \left(\left(\frac{\partial h_j}{\partial k} \right)^2 + h_j \cdot \frac{\partial^2 h_j}{\partial k^2} \right) = G_{11}, \\ \frac{\partial^2 F}{\partial k \partial \varphi_0} &= \sum_{j=1}^n \left(\frac{\partial h_j}{\partial \varphi_0} \cdot \frac{\partial h_j}{\partial k} + \frac{\partial^2 h_j}{\partial k \partial \varphi_0} \right) = G_{12} = G_{21}, \\ \frac{\partial^2 F}{\partial \varphi_0^2} &= \sum_{j=1}^n \left(\left(\frac{\partial h_j}{\partial \varphi_0} \right)^2 + h_j \cdot \frac{\partial^2 h_j}{\partial \varphi_0^2} \right) = G_{22}. \end{aligned} \quad (13)$$

Вычисление матрицы Гессе сведено к вычислению вторых производных отклонений h_j аппроксимируемых точек от клотоиды по искомым переменным φ_0 и k , т.к. первые производные уже вычислены.

В соответствии с (9)

$$\begin{aligned} \frac{\partial^2 h_j}{\partial k^2} &= -\sin \varphi \cdot \frac{\partial^2 x_B}{\partial k^2} + \cos \varphi \cdot \frac{\partial^2 y_B}{\partial k^2} = \\ &= \sin \varphi \int_0^L \cos \left(\varphi_0 + \frac{kt^2}{2} \right) \frac{t^4}{4} dt - \cos \varphi \int_0^L \sin \left(\varphi_0 + \frac{kt^2}{2} \right) \frac{t^4}{4} dt. \end{aligned}$$

Применяя интегрирование по частям и используя вычисленные ранее интегралы, получаем

$$\begin{aligned}\frac{\partial^2 h_j}{\partial k^2} &= \frac{1}{4k} \left\{ L^3 + \frac{3}{k} (\cos \varphi (y_B - y_A) - \sin \varphi (x_B - x_A)) \right\}, \\ \frac{\partial^2 h_j}{\partial \varphi_0^2} &= -\sin \varphi \cdot \frac{\partial^2 x_B}{\partial \varphi_0^2} + \cos \varphi \cdot \frac{\partial^2 y_B}{\partial \varphi_0^2} = \\ &= \sin \varphi (x_B - x_A) - \cos \varphi (y_B - y_A), \\ \frac{\partial^2 h_j}{\partial k \partial \varphi_0} &= \sin \varphi \int_0^L \cos \left(\varphi_0 + \frac{kt^2}{2} \right) \frac{t^2}{2} dt - \\ &- \cos \varphi \int_0^L \sin \left(\varphi_0 + \frac{kt^2}{2} \right) \frac{t^2}{2} dt = \\ &= \frac{1}{2k} \{ L - \sin \varphi (y_B - y_A) - \cos \varphi (x_B - x_A) \}.\end{aligned}\quad (14)$$

Используя полученные производные, вычисляем матрицу Гессе и обратную к ней G^{-1} .

Обозначим: $x(\varphi_0, k)$ – вектор неизвестных, $g\left(\frac{\partial F}{\partial \varphi_0}, \frac{\partial F}{\partial k}\right)$ – градиент. Для начального приближения, полученного по методу эвольвент (нулевая итерация), это x^0 и g^0 . Вычислим следующую итерационную точку:

$$x^1 = x^0 - G^{-1} \cdot g^0. \quad (15)$$

Для положительно определенной матрицы Гессе это точка минимума целевой функции, если ее разложить в ряд Тейлора и ограничиться вторыми производными. В общем случае одного шага до точки минимума на луче $x^0 - \lambda G^{-1} \cdot g^0$ недостаточно. В квадратичной задаче минимум достигается при $\lambda = 1$. В общем случае при $\lambda = 1$ получаем точку минимума не исходной функции, а аппроксимирующей ее квадратичной формы с матрицей Гессе, что не одно и то же. Поэтому $\lambda = 1$ следует рассматривать лишь как приближенное значение и искать точное значение путем решения задачи одномерной оптимизации функции $F(x)$ на луче $x^0 - \lambda G^{-1} \cdot g^0$, рассматривая ее как функцию единственного параметра λ .

В полученной точке минимума по направлению приходится снова вычислять градиент и матрицу, обратную к матрице Гессе и т.д.

В целом алгоритм клотоидной аппроксимации состоит из следующих пунктов.

1. Построение углограммы исходной ломаной по точкам съемки.
2. Вычисление длин эвольвенты ломаной последовательно по точкам съемки по формуле (3).
3. Решение задачи аппроксимации полученной ломаной. В общем случае аппроксимация кубической параболой при наличии трех неизвестных параметров клотоиды или квадратной параболой при двух параметрах, как рассмотрено выше.

При этом решается система не более чем трех линейных уравнений.

4. Для клотоиды, соответствующей полученному решению, по специальному итерационному алгоритму определяются точки пересечения с нормальными из каждой точки съемки, углы касательных в этих точках с осью OX, соответствующие длины от начальной точки клотоиды до каждой из них и отклонения h_j точек съемки от клотоиды.
5. Вычисляются первые и вторые производные отклонений h_j по искомым параметрам клотоиды (12), (13).
6. Вычисляется градиент целевой функции (14) и проверяются условия окончания счета (например, малость нормы градиента). Если условия окончания счета не выполнены, то далее:
7. Вычисляется матрица Гессе и обратная к ней.
8. Определяется точка минимума соответствующей квадратичной формы (15), решается задача одномерной оптимизации, т.е. корректировки шага по направлению поиска, и осуществляется переход в новую итерационную точку и далее с новыми значениями неизвестных к п. 4.

Для отладки и проверки алгоритма клотоидной аппроксимации сначала использовались результаты решения тестовых задач по методу эвольвент, для которых оптимум был известен, но решения по методу эвольвент были неудовлетворительны.

В рассмотренном выше примере 1 по методу эвольвент максимальное отклонение $D_{\max} = 0.255$ м вместо 0 , $\varphi_0 = -0.00027412$ вместо 0 и $k = 3.337930 \cdot 10^{-5}$ вместо $k = 3.333333 \cdot 10^{-5}$.

Целевая функция $F^0 = 0.10298$. В точке минимума квадратичной формы (15) $F = 1.08524 \cdot 10^{-7}$, $\varphi_0 = 3.099241 \cdot 10^{-7}$, $k = 3.333335 \cdot 10^{-5}$. После выполнения первой итерации $F = 3.70261 \cdot 10^{-15}$, $\varphi_0 = -2.422261 \cdot 10^{-14}$, $k = 3.33333333 \cdot 10^{-5}$. Все отклонения не превышали 0.00006 м.

В примере 2 решалась та же задача, но при длине клотоиды 200 м. По методу эвольвент получено $\varphi_0 = -1.839272551 \cdot 10^{-5}$, вместо 0 и $k = 3.334527 \cdot 10^{-5}$ вместо $k = 3.333333 \cdot 10^{-5}$. $D_{\max} = 0.048$ м. В точке минимума квадратичной формы (15) $F = 1.80438 \cdot 10^{-14}$, $\varphi_0 = -9.592883457 \cdot 10^{-10}$, $k = 3.333334 \cdot 10^{-5}$, $D_{\max} = 0.0002$ м. Итерация была прервана из-за достижения требуемой точности по градиенту.

ЗАКЛЮЧЕНИЕ

Возможности современных общедоступных компьютеров, теории и методов компьютерной выработки проектных решений далеко не полностью используются в действующих САПР линейных сооружений, которые основаны на идеях более чем

50-летней давности. Переход к разработке и внедрению интеллектуальных САПР, в которых проектные решения выдает компьютер в результате решения задач оптимизации, вполне возможен уже сейчас применительно ко многим проектным задачам, но в силу отсутствия заинтересованных инстанций в создании таких систем и больших затрат труда и средств на их разработку такой переход в нашей стране в ближайшее время маловероятен. Наиболее продвинутыми в этом отношении являются упоминавшиеся работы китайских ученых. Однако действующие САПР могут быть усовершенствованы и путем применения программ оптимизации в интерактивном процессе выработки проектных решений. Так, изложенные в статье алгоритмы аппроксимации можно применить вместо назначения решений вручную. Кроме того, изложенный алгоритм клотоидной

аппроксимации может быть востребован при решении задач, не имеющих отношения к проектированию дорожных трасс.

Ближайшей задачей является обобщение изложенного метода вычисления производных целевой функции по параметрам, определяющим кривую в целом (а не одну клотоиду), при отсутствии аналитического выражения этой функции через искомые параметры. Речь идет о переходе от решения задачи аппроксимации одной клотоидой к сплайн-аппроксимации последовательностью нескольких связок «прямая + клотоида + окружность + клотоида + ...».

Вклад авторов. Все авторы в равной степени внесли вклад в настоящую работу.

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

СПИСОК ЛИТЕРАТУРЫ

1. Горинов А.В. *Изыскания и проектирование железных дорог*. Т. 2. М.: Трансжелдориздат; 1961. 338 с.
2. Шейдвассер Д.М. Оптимизация трасс железных дорог на напряженных ходах. В кн.: *Автоматизация проектирования объектов транспортного строительства*. М.: Транспорт; 1986. С. 16–29.
3. Струченков В.И., Шейдвассер Д.М. Оптимизация на ЭВМ трассы новой железной дороги на напряженных ходах. *Транспортное строительство*. 1987;3:7–9.
4. Jha M.K., McCall C., Schonfeld P. Using GIS, genetic algorithms, and visualization in highway development. *Computer-Aided Civil and Infrastructure Engineering*. 2001;16(6):399–414. <https://doi.org/10.1111/0885-9507.00242>
5. Jha M.K., Schonfeld P. A highway alignment optimization model using geographic information systems. *Transp. Res. Part A. Policy Pract.* 2004;8(6):455–481. <https://doi.org/10.1016/j.tra.2004.04.001>
6. Jong J.C., Jha M.K., Schonfeld P. Preliminary highway design with genetic algorithms and geographic information systems. *Computer-Aided Civil and Infrastructure Engineering*. 2000;15(4):261–271. <https://doi.org/10.1111/0885-9507.00190>
7. Kang M.W., Schonfeld P., Yang N. Prescreening and repairing in a genetic algorithm for highway alignment optimization. *Computer-Aided Civil and Infrastructure Engineering*. 2009;24(2):109–119. <https://doi.org/10.1111/j.1467-8667.2008.00574.x>
8. Pushak Y., Hare W., Lucet Y. Multiple-path selection for new highway alignments using discrete algorithms. *Eur. J. Oper. Res.* 2016;248(2):415–427. <https://doi.org/10.1016/j.ejor.2015.07.039>
9. Sarma K.C., Adeli H. Bilevel parallel genetic algorithms for optimization of large steel structures. *Computer Aided Civil and Infrastructure Engineering*. 2001;16(5):295–304. <https://doi.org/10.1111/0885-9507.00234>

REFERENCES

1. Gorinov A.V. *Izyskaniya i proektirovanie zheleznikh dorog (Research and Design of Railways)*. V. 2. Moscow: Transzheldorizdat; 1961. 338 p. (in Russ.).
2. Sheidwasser D.M. Optimization of railroad tracks on busy tracks. In: *Avtomatizatsiya proektirovaniya ob'ektov transportnogo stroitel'stva (Automation of the Design of Transport Construction Objects)*. Moscow: Transport; 1986. P. 16–29 (in Russ.).
3. Struchenkov V.I., Sheidwasser D.M. Optimization on a computer of the route of a new railway on stressful passages. *Transportnoe stroitel'stvo = Transport Construction*. 1987;3:7–9 (in Russ.).
4. Jha M.K., McCall C., Schonfeld P. Using GIS, genetic algorithms, and visualization in highway development. *Computer-Aided Civil and Infrastructure Engineering*. 2001;16(6):399–414. <https://doi.org/10.1111/0885-9507.00242>
5. Jha M.K., Schonfeld P. A highway alignment optimization model using geographic information systems. *Transp. Res. Part A. Policy Pract.* 2004;8(6):455–481. <https://doi.org/10.1016/j.tra.2004.04.001>
6. Jong J.C., Jha M.K., Schonfeld P. Preliminary highway design with genetic algorithms and geographic information systems. *Computer-Aided Civil and Infrastructure Engineering*. 2000;15(4):261–271. <https://doi.org/10.1111/0885-9507.00190>
7. Kang M.W., Schonfeld P., Yang N. Prescreening and repairing in a genetic algorithm for highway alignment optimization. *Computer-Aided Civil and Infrastructure Engineering*. 2009;24(2):109–119. <https://doi.org/10.1111/j.1467-8667.2008.00574.x>
8. Pushak Y., Hare W., Lucet Y. Multiple-path selection for new highway alignments using discrete algorithms. *Eur. J. Oper. Res.* 2016;248(2):415–427. <https://doi.org/10.1016/j.ejor.2015.07.039>

10. Shafahi Y., Bagherian M. A customized particle swarm method to solve highway alignment optimization problem. *Computer-Aided Civil and Infrastructure Engineering*. 2013;28(1):52–67. <https://doi.org/10.1111/j.1467-8667.2012.00769.x>
11. Bosurgi G., D'Andrea A. A polynomial parametric curve (PPC-curve) for the design of horizontal geometry of highways. *Computer-Aided Civil and Infrastructure Engineering*. 2012;27(4):303–312. <https://doi.org/10.1111/j.1467-8667.2011.00750.x>
12. Cerf R. *The quasispecies regime for the simple genetic algorithm with roulette wheel selection*. arXiv:1506.0981v2. <https://doi.org/10.48550/arXiv.1506.09081>
13. Pu H., Li W., Schonfeld P., et al. A Method for Automatically Recreating the Horizontal Alignment Geometry of Existing Railways. *Computer-Aided Civil and Infrastructure Engineering*. 2019;34(1):71–94. <https://doi.org/10.1111/mice.12392>
14. Li W., Zhen S., Schonfeld P., et al. Recreating Existing Railway Horizontal Alignments Automatically Using Overall Swing Iteration. *J. Transport. Eng. Part A: Systems*. 2022;148(8). <https://doi.org/10.1061/JTEPBS.0000691>
15. Pu H., Fu H., Schonfeld P., et al. Modelling and optimization of constrained alignments for existing railway reconstruction. *Int. J. Rail Transportat.* 2023;11(3):428–447. <https://doi.org/10.1080/23248378.2022.2081878>
16. Сальков Н.А. *Моделирование геометрических форм автомобильных дорог*: монография. М.: ИНФРА-М; 2019. 162 с.
17. Смирнов В.И. *Курс высшей математики*. Т. 2. М.: Наука; 1967. 479 с.
18. Горинов А.В., Кантор И.И., Кондратенко А.П., Турбин И.В. *Изыскания и проектирование железных дорог*. М.: Транспорт; 1979. 319 с.
19. Кантор И.И. *Изыскания и проектирование железных дорог*. М.: ИКЦ «Академкнига»; 2003. 288 с.
20. Федотов Г.А., Поспелов И.И. *Изыскания и проектирование автомобильных дорог*. Кн. 1. М.: Высшая школа; 2009. 650 с.
21. Гилл Ф., Мюррей У., Райт М. *Практическая оптимизация*: пер. с англ. М.: Мир; 1985. 509 с.
22. Кохендерфер М.Д., Уилер Т.А. *Алгоритмы оптимизации*. М.: Вильямс; 2020. 528 с.
23. Черноуцкий И.Г. *Методы оптимизации. Компьютерные технологии*. СПб.: БХВ-Петербург; 2011. 329 с.
24. Струченков В.И. Новый алгоритм поэлементного расчета трасс в САПР линейных сооружений. *Информационные технологии*. 2015;21(4):271–276.
9. Sarma K.C., Adeli H. Bilevel parallel genetic algorithms for optimization of large steel structures. *Computer Aided Civil and Infrastructure Engineering*. 2001;16(5): 295–304. <https://doi.org/10.1111/0885-9507.00234>
10. Shafahi Y., Bagherian M. A customized particle swarm method to solve highway alignment optimization problem. *Computer-Aided Civil and Infrastructure Engineering*. 2013;28(1):52–67. <https://doi.org/10.1111/j.1467-8667.2012.00769.x>
11. Bosurgi G., D'Andrea A. A polynomial parametric curve (PPC-curve) for the design of horizontal geometry of highways. *Computer-Aided Civil and Infrastructure Engineering*. 2012;27(4):303–312. <https://doi.org/10.1111/j.1467-8667.2011.00750.x>
12. Cerf R. *The quasispecies regime for the simple genetic algorithm with roulette wheel selection*. arXiv:1506.0981v2. <https://doi.org/10.48550/arXiv.1506.09081>
13. Pu H., Li W., Schonfeld P., et al. A Method for Automatically Recreating the Horizontal Alignment Geometry of Existing Railways. *Computer-Aided Civil and Infrastructure Engineering*. 2019;34(1):71–94. <https://doi.org/10.1111/mice.12392>
14. Li W., Zhen S., Schonfeld P., et al. Recreating Existing Railway Horizontal Alignments Automatically Using Overall Swing Iteration. *J. Transport. Eng. Part A: Systems*. 2022;148(8). <https://doi.org/10.1061/JTEPBS.0000691>
15. Pu H., Fu H., Schonfeld P., et al. Modelling and optimization of constrained alignments for existing railway reconstruction. *Int. J. Rail Transportat.* 2023;11(3):428–447. <https://doi.org/10.1080/23248378.2022.2081878>
16. Sal'kov N.A. *Modelirovanie geometricheskikh form avtomobil'nykh dorog*: Monografiya (Modeling Geometric Shapes of Highways: Monograph). Moscow: INFRA-M; 2019. 162 p. (in Russ.).
17. Smirnov V.I. *Kurs vysshei matematiki (Higher Mathematics Course)*. V. 2. Moscow: Nauka; 1979. 479 p. (in Russ.).
18. Gorinov A.V., Kantor I.I., Kondratenko A.P., Turbin I.V. *Izyskaniya i proektirovanie zheleznykh dorog (Research and Design of Railways)*. Moscow: Transport; 1979. 319 p. (in Russ.).
19. Kantor I.I. *Izyskaniya i proektirovanie zheleznykh dorog (Research and Design of Railways)*. Moscow: Akademkniga; 2003. 288 p. (in Russ.).
20. Fedotov G.A., Pospelov I.I. *Izyskaniya i proektirovanie avtomobil'nykh dorog (Research and Design of Highways)*. V. 1. Moscow: Vysshaya shkola; 2009. 650 p. (in Russ.).
21. Gill Ph.E., Murray W., Wright M.H. *Prakticheskaya optimizatsiya (Practical Optimization)*: transl. from Engl. Moscow: Mir; 1985. 509 p. (in Russ.). [Gill Ph.E., Murray W., Wright M.H. *Practical Optimization*. London: Academic Press; 1981. 402 p.]
22. Kochenderfer M.D., Wheeler T.A. *Algoritmy optimizatsii (Algorithms for Optimization)*. Moscow: Vil'yams; 2020. 528 p. (in Russ.). [Kochenderfer M.D., Wheeler T.A. *Algorithms for Optimization*. London: MIT Press; 2019. 520 p.]

23. Chernorutskii I.G. *Metody optimizatsii. Komp'yuternye tekhnologii (Methods of optimization. Computer technologies)*. St. Petersburg: BHV-Petersburg; 2011. 329 p. (in Russ.).
24. Struchenkov V.I. New algorithm for perelement calculation of line structures routes. *Informacionnye tekhnologii = Information Technologies*. 2015;21(4): 271–276 (in Russ.).

Об авторах

Карпов Дмитрий Анатольевич, к.т.н., заведующий кафедрой общей информатики Института искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: karpov@mirea.ru. SPIN-код РИНЦ 2619-7100, <https://orcid.org/0000-0003-3734-7182>

Струченков Валерий Иванович, д.т.н., профессор, кафедра общей информатики Института искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: str1942@mail.ru. SPIN-код РИНЦ 4581-4698, <https://orcid.org/0000-0002-9801-7454>

About the authors

Dmitry A. Karpov, Cand. Sci. (Eng.), Head of the General Informatics Department, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: karpov@mirea.ru. RSCI SPIN-code 2619-7100, <https://orcid.org/0000-0003-3734-7182>

Valery I. Struchenkov, Dr. Sci. (Eng.), Professor, General Informatics Department, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: str1942@mail.ru. RSCI SPIN-code 4581-4698, <https://orcid.org/0000-0002-9801-7454>

Математическое моделирование
Mathematical modeling

УДК 621.372.8
<https://doi.org/10.32362/2500-316X-2023-11-4-84-93>



НАУЧНАЯ СТАТЬЯ

Модели волноводов, сочетающих градиентные и нелинейно-оптические слои

С.Е. Савотченко[@]

Белгородский государственный технологический университет им. В.Г. Шухова, Белгород, 308012
Россия

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

Резюме

Цели. Теоретические исследования волноводных свойств границ раздела нелинейно-оптических и градиентных сред являются важными для использования в оптоэлектронике. Комбинированные волноводы, сочетающие слои с различными оптическими свойствами, представляются наиболее перспективными, поскольку для них можно подобрать оптимальные характеристики с помощью широкого ряда управляющих параметров. Цель работы – разработка теории композитных оптически-нелинейных градиентных волноводов с произвольным профилем, в рамках которой возможно получение точных аналитических выражений для поверхностных волн и волноводных мод в явном виде. Основной особенностью предлагаемой в данной работе теории является то, что она применима для описания поверхностных волн и волноводных мод, поле в которых сосредоточено внутри градиентного слоя и не выходит за его границу, не контактирующую с нелинейным слоем.

Методы. Использованы аналитические методы теории оптических волноводов, нелинейной оптики.

Результаты. Проведено теоретическое описание волноводных свойств границы раздела двух сред с принципиально различными оптическими характеристиками. Сформулированная модель плоского волновода применима для сред с произвольным распределением пространственного профиля диэлектрической проницаемости. Получено аналитическое выражение, описывающее поверхностную волну, распространяющуюся вдоль границы раздела среды со ступенчатой нелинейностью и градиентного слоя с произвольным профилем диэлектрической проницаемости. Также получены аналитические выражения для поверхностных волн, распространяющихся вдоль границы раздела среды с керровской нелинейностью (как самофокусирующей, так и дефокусирующей) с градиентными средами, характеризующимися экспоненциальным и линейным профилями диэлектрической проницаемости.

Выводы. Предложенная теория позволяет наглядно описать в явном аналитическом виде узко локализованные световые потоки в таких волноводах. Показано, что сочетание различных полупроводниковых кристаллов в композитном волноводе позволяет получить с одной стороны от волноведущего интерфейса нелинейно-оптический слой, а с другой – слой с градиентным профилем диэлектрической проницаемости.

Ключевые слова: нелинейная оптика, нелинейные волны, оптическая нелинейность, керровская нелинейность, оптический волновод, градиентный волновод

• Поступила: 13.12.2022 • Доработана: 17.02.2023 • Принята к опубликованию: 18.05.2023

Для цитирования: Савотченко С.Е. Модели волноводов, сочетающих градиентные и нелинейно-оптические слои. *Russ. Technol. J.* 2023;11(4):84–93. <https://doi.org/10.32362/2500-316X-2023-11-4-84-93>

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

Автор заявляет об отсутствии конфликта интересов.

RESEARCH ARTICLE

Models of waveguides combining gradient and nonlinear optical layers

Sergey E. Savotchenko [®]

V.G. Shukhov Belgorod State Technological University, Belgorod, 308012 Russia

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

Abstract

Objectives. Theoretical studies of the waveguide properties of interfaces between nonlinear optical and graded-index media are important for application in optoelectronics. Waveguides combining layers with different optical properties seem to be the most promising, since they can be matched to optimal characteristics using a wide range of control parameters. The paper aims to develop a theory of composite optically nonlinear graded-index waveguides with an arbitrary profile, within which it is possible to obtain exact analytical expressions for surface waves and waveguide modes in an explicit form. The main feature of the theory proposed in this paper is its applicability for describing surface waves and waveguide modes, in which the field is concentrated inside the gradient layer and does not exceed its boundary, avoiding contact with the nonlinear layer.

Methods. Analytical methods of the theory of optical waveguides and nonlinear optics are used.

Results. A theoretical description of the waveguide properties of the interface between two media having significantly different optical characteristics is carried out. The formulated model of a plane waveguide is applicable to media having an arbitrary spatial permittivity profile. An analytical expression describing a surface wave propagating along the interface between a medium having stepwise nonlinearity and a gradient layer with an arbitrary permittivity profile is obtained. Additionally, analytical expressions for surface waves propagating along the interface between a medium with Kerr nonlinearity (both self-focusing and defocusing), as well as graded-index media characterized by exponential and linear permittivity profiles, are obtained.

Conclusions. The proposed theory supports a visual description in an explicit analytical form of a narrowly localized light beam within such waveguides. It is shown that by combining different semiconductor crystals in a composite waveguide, it is possible to obtain a nonlinear optical layer on one side of the waveguide interface and a layer with a graded-index dielectric permittivity profile on the other.

Keywords: nonlinear optics, nonlinear waves, optical nonlinearity, Kerr nonlinearity, optical waveguide, graded-index waveguide

• Submitted: 13.12.2022 • Revised: 17.02.2023 • Accepted: 18.05.2023

For citation: Savotchenko S.E. Models of waveguides combining gradient and nonlinear optical layers. *Russ. Technol. J.* 2023;11(4):84–93. <https://doi.org/10.32362/2500-316X-2023-11-4-84-93>

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

Волноводные структуры обычно основаны на различиях оптических свойств их слоев [1, 2]. В оптоэлектронике широко используются волноводы с постоянными значениями показателя преломления слоев [3], с переменным профилем пространственного распределения (градиентные) [4], а также с нелинейно-оптическим откликом, когда показатель преломления зависит от интенсивности светового потока [5]. Комбинированные волноводы, сочетающие такие слои, представляются наиболее перспективными для оптоэлектроники, поскольку для них можно подобрать оптимальные характеристики с помощью широкого ряда управляющих параметров [6, 7].

В последние годы активизировались теоретические исследования волноводных свойств границ раздела нелинейно-оптических и градиентных сред [8, 9]. В частности, получены волноводные моды слоистой структуры, состоящей из градиентного слоя с линейным [10, 11] и экспоненциальным [12, 13] профилями, контактирующими с керровской нелинейной средой. В наших недавних работах представлены результаты теоретического исследования волноводных свойств структур, сочетающих пары различных градиентных и нелинейных сред. В частности, рассмотрены линейный [14–19], параболический [20, 21] и экспоненциальный профили [22, 23] показателя преломления / диэлектрической проницаемости сред, контактирующих со средами различной нелинейности: ступенчатой [14, 22], керровской [20] и ее обобщений [16–18, 20, 21], фоторефрактивной диффузионного типа [21, 23]. Описаны также симметричные трехслойные структуры, состоящие из линейно-градиентного слоя в среде с керровской нелинейностью [24] и в среде с фоторефрактивной нелинейностью [25].

В работе предлагается теоретическое описание поверхностных волн, распространяющихся вдоль плоской границы раздела сред, одна из которых обладает нелинейным откликом, а другая характеризуется пространственным градиентом диэлектрической проницаемости. Новизна работы заключается в том, что предложена теория таких композитных оптически-нелинейных градиентных волноводов с произвольным профилем, в рамках которой получены

аналитические выражения для поверхностных волн. Основная особенность данной теории – это ее применимость для описания поверхностных волн и волноводных мод, поле в которых сосредоточено внутри градиентного слоя и не выходит за его границу, не контактирующую с нелинейным слоем. Предложенная теория позволяет наглядно описать в явном аналитическом виде узколокализованные световые потоки в таких волноводах. Рассмотрены также примеры полупроводниковых материалов, которые могут быть использованы для проектирования волноводов рассматриваемого типа.

1. ТЕОРИЯ И МЕТОДЫ ИССЛЕДОВАНИЯ

В качестве простейшей волноводной структуры рассмотрим плоскую границу раздела двух немагнитных сред принципиально различных оптических типов. В частности, с одной стороны границы будет среда с плавным изменением пространственного профиля показателя преломления / диэлектрической проницаемости (т.е. $n = n(x)$ или $\varepsilon = \varepsilon(x)$ соответственно; x – пространственная координата в направлении, перпендикулярном плоскости контакта $x = 0$), а с другой – нелинейно-оптическая среда, в которой показатель преломления / диэлектрическая проницаемость зависят от интенсивности света I (т.е. $n = n(I)$ или $\varepsilon = \varepsilon(I)$, $I = |E|^2$, E – амплитуда напряженности электрического поля). Будем считать, что рассматриваемые среды не обладают диэлектрическими потерями, и поэтому все их оптические характеристики являются действительными величинами.

Начало координат располагается в плоскости yz так, что ось x перпендикулярна границе раздела, расположенной в плоскости $x = 0$. Будем рассматривать распространяющуюся вдоль границы раздела поперечную волну с отличной от нуля компонентой напряженности электрического поля: $E_y = \psi(x) \times \exp(i\beta z - i\omega t)$, где $\beta = kn$ – константа распространения; $n = ck/\omega$ – эффективный показатель преломления; c – скорость света в вакууме; ω – частота; $k = 2\pi/\lambda$ – волновое число; λ – длина волны; $\psi(x)$ – пространственное распределение напряженности электрического поля в направлении, поперечном границе раздела. Как хорошо известно [1, 2], функция $\psi(x)$

подчиняется стационарному уравнению (магнитная проницаемость считается равной единице):

$$\psi''(x) + \{\varepsilon(x, I) - n^2\}k^2\psi(x) = 0, \quad (1)$$

где диэлектрическую проницаемость волноводной системы представим в виде:

$$\varepsilon(x, I) = \begin{cases} \varepsilon_G(x), & x < 0, \\ \varepsilon_N(I), & x > 0. \end{cases} \quad (2)$$

И пусть для определенности градиентная среда располагается в левом полупространстве и характеризуется диэлектрической проницаемостью $\varepsilon_G(x)$, а оптически нелинейная среда располагается в правом полупространстве и характеризуется диэлектрической проницаемостью $\varepsilon_N(I)$.

Если поперечное распределение поля представить в виде:

$$\psi(x) = \begin{cases} \psi_G(x), & x < 0, \\ \psi_N(x), & x > 0, \end{cases} \quad (3)$$

то с учетом (2) уравнение (1) распадается на два:

$$\psi_G''(x) + \{\varepsilon_G(x) - n^2\}k^2\psi_G(x) = 0, \quad x < 0, \quad (4)$$

$$\psi_N''(x) + \{\varepsilon_N(I) - n^2\}k^2\psi_N(x) = 0, \quad x > 0. \quad (5)$$

Уравнения (4) и (5) следует дополнить граничными условиями, описывающими требование непрерывности компонент поля на границе раздела сред:

$$\psi_N(+0) = \psi_G(-0), \quad \psi_N'(+0) = \psi_G'(-0), \quad (6)$$

а также ограниченности на бесконечности: $|\psi(x)| \rightarrow 0$, $|x| \rightarrow \infty$.

Поскольку локализация энергии светового пучка вдоль волноведущего слоя является важной задачей, то основное внимание будет уделено таким волнам, для которых глубина проникновения поля меньше толщины градиентного слоя. В ряде наших работ [14–22] было показано, что в определенных интервалах значений параметров волноводной системы амплитуда поля на расстояниях порядка толщины градиентного слоя может быть намного меньше амплитуды на границе раздела, и поэтому ею можно пренебречь. Следовательно, в таком случае нет необходимости использовать граничные условия на расстоянии толщины градиентного слоя, можно использовать только граничные условия на границе раздела сред для описания узко локализованных поперечных волн.

Решение уравнения (4) представим в виде:

$$\psi_G(x) = \psi_0 \frac{F(g(x))}{F(g(0))}, \quad (7)$$

где ψ_0 – амплитуда поля на границе раздела; $F(g)$ представляет собой специальную функцию с вспомогательным аргументом $g(x)$, разрешающую аналитически уравнение (4) при заданном профиле диэлектрической проницаемости $\varepsilon_G(x)$ и удовлетворяющую уравнению

$$F''g' + F'g'' + \{\varepsilon_G(x) - n^2\}k^2F = 0. \quad (8)$$

Явный вид зависимости $g(x)$ связан с заменой переменных, которую необходимо сделать, чтобы привести уравнение (8) к виду, для которого известно, что точное решение выражается через какую-либо специальную функцию. К примеру, в случае линейной замены $g(x) = ax + b$ в (8) будет $g' = a$ и $g'' = 0$ и тогда оно упростится: $aF'' + \{\varepsilon_G(x) - n^2\}k^2F = 0$.

При выборе представления решения в форме (7) обязательно следует учесть требование на бесконечности: $|F(g(x))| \rightarrow 0$, $|x| \rightarrow \infty$.

Резюмируем кратко основные положения теории рассматриваемых композитных волноводов. Волновод представляет собой ультратонкую границу, разделяющую среду с градиентом диэлектрической проницаемости и нелинейно-оптическую среду. Поперечная волна локализована так, что поле не выходит за границу градиентного слоя. Распределение электрического поля в направлении, поперечном границе раздела, определяется из стационарного одномерного нелинейного уравнения, удовлетворяет условиям сопряжения на границе раздела и условию исчезновения на бесконечности.

Явный вид решения уравнения (5) зависит от выбора модели нелинейности среды. Рассмотрим далее возникающие в таких волноводах волны с профилями, узко локализованными вдоль границы раздела сред, на примерах керровской и ступенчатой нелинейности и градиентного экспоненциального профиля.

2. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

2.1. Керровская нелинейность

Наиболее распространенной формой нелинейного отклика оптической системы является керровская, для которой диэлектрическая проницаемость линейно зависит от интенсивности света:

$$\varepsilon_N(I) = \varepsilon_{0N} + \alpha I, \quad (9)$$

где ε_{0N} – невозмущенная диэлектрическая константа (положительная); α – коэффициент керровской

нелинейности, положительное значение которого соответствует самофокусирующей среде, а отрицательное – дефокусирующей. Тогда с учетом диэлектрической проницаемости (9) уравнение (5) примет вид:

$$\psi_N''(x) - q_N^2 \psi_N(x) + \alpha k^2 \psi_N^3(x) = 0, \quad (10)$$

где $q_N^2 = k^2(n^2 - \varepsilon_{0N})$.

Решение нелинейного уравнения (10) при $n^2 > \varepsilon_{0N}$, удовлетворяющее условию $|\psi(x)| \rightarrow 0$, $|x| \rightarrow \infty$, можно записать в виде:

$$\psi_N(x) = \begin{cases} \sqrt{\frac{2}{\alpha}} \cdot \frac{q_N}{kchq_N(x - x_N)}, & \alpha > 0, \\ \sqrt{\frac{2}{|\alpha|}} \cdot \frac{q_N}{kshq_N(x - x_N)}, & \alpha < 0, \end{cases} \quad (12)$$

где величина x_N характеризует положение центра «солитона» и определяется из граничных условий (6). Подставив (7) и (12) в (6), получим:

$$x_N = \begin{cases} \frac{1}{k\sqrt{n^2 - \varepsilon_{0N}}} \cdot \operatorname{arctg}\left(\frac{\varepsilon_{G\text{eff}}}{n^2 - \varepsilon_{0N}}\right)^{1/2}, & \alpha > 0, \\ \frac{1}{k\sqrt{n^2 - \varepsilon_{0N}}} \cdot \operatorname{arctg}\left(\frac{\varepsilon_{G\text{eff}}}{n^2 - \varepsilon_{0N}}\right)^{1/2}, & \alpha < 0, \end{cases} \quad (13)$$

где введена эффективная диэлектрическая проницаемость градиентного слоя

$$\varepsilon_{G\text{eff}} = \frac{F'(g(0))}{F(g(0))} \cdot \frac{g'(0)}{k}. \quad (14)$$

Аналогично, после подстановки (7) и (12) в (6) получается интенсивность поля на границе раздела:

$$I_0 = \psi_0^2 = \begin{cases} 2(n^2 - \varepsilon_{0N} - \varepsilon_{G\text{eff}}) / \alpha, & \alpha > 0, \\ 2(\varepsilon_{G\text{eff}} + \varepsilon_{0N} - n^2) / |\alpha|, & \alpha < 0. \end{cases} \quad (15)$$

В качестве примера рассмотрим экспоненциальный профиль диэлектрической проницаемости:

$$\varepsilon_G(x) = \varepsilon_e + (\varepsilon_0 - \varepsilon_e)e^{2x/a}, \quad (16)$$

где ε_0 и ε_e диэлектрические константы (положительные) на границе раздела и на конце градиентного слоя характерной толщины a .

Подставив профиль (16) в уравнение (4), можно получить:

$$\psi_G''(x) + (V/a)^2 \exp(2x/a) - q^2 \psi_G(x) = 0, \quad (17)$$

где $V^2 = a^2 k^2 (\varepsilon_0 - \varepsilon_e)$, $q^2 = k^2(n^2 - \varepsilon_e)$.

В этом случае $F(g(x)) = J_v(g(x))$, где $J_v(g)$ – функция Бесселя первого рода, $g(x) = Ve^{x/a}$, и тогда выражение (7), определяющее решение уравнения (17) при $n^2 > \varepsilon_e$ и удовлетворяющее условию $|\psi(x)| \rightarrow 0$, $x \rightarrow -\infty$, можно записать в виде:

$$\psi_G(x) = \psi_0 J_{aq}(Ve^{x/a}) / J_{aq}(V). \quad (18)$$

Эффективная диэлектрическая проницаемость экспоненциального градиентного слоя:

$$\varepsilon_{G\text{eff}} = \frac{J'_{aq}(V)}{J_{aq}(V)} \cdot \frac{V}{ak}. \quad (19)$$

Таким образом, используя выражения (3), (12) и (18), можно записать волну, распространяющуюся вдоль границы раздела самофокусирующей керровской нелинейной среды и экспоненциального градиентного слоя, в виде:

$$\psi(x) = \begin{cases} \sqrt{\frac{2(n^2 - \varepsilon_{0N} - \varepsilon_{G\text{eff}})}{\alpha}} \cdot \frac{J_{aq}(Ve^{x/a})}{J_{aq}(V)}, & x < 0, \\ \sqrt{\frac{2}{\alpha}} \cdot \frac{q_N}{kchq_N(x - x_N)}, & x > 0, \end{cases} \quad (20)$$

где диэлектрическая проницаемость экспоненциального градиентного слоя определяется выражением (19).

В качестве второго примера рассмотрим линейный профиль диэлектрической проницаемости:

$$\varepsilon_G(x) = \varepsilon_0 + (\varepsilon_0 - \varepsilon_e)(x/h). \quad (21)$$

где h – толщина слоя в случае линейной функции.

Подставив профиль (21) в уравнение (4), можно получить:

$$\psi_G''(x) + \{\varepsilon_0 - n^2 + (\varepsilon_0 - \varepsilon_e)(x/h)\} k^2 \psi_G(x) = 0. \quad (22)$$

В этом случае $F(g(x)) = \operatorname{Ai}(g(x))$, где $\operatorname{Ai}(g)$ – функция Эйри первого рода, $g(x) = -x/x_G + \delta$ и $x_G = \{a/k_0^2(\varepsilon_0 - \varepsilon_e)\}^{1/3}$, $\delta = -(\varepsilon_0 - n^2)h/x_G(\varepsilon_0 - \varepsilon_e)$. Тогда выражение (7), определяющее решение уравнения (22) при $\varepsilon_e < n^2 < \varepsilon_0$ и удовлетворяющее условию $|\psi(x)| \rightarrow 0$, $x \rightarrow -\infty$, можно записать в виде:

$$\psi_G(x) = \psi_0 \operatorname{Ai}(-x/x_L + \delta) / \operatorname{Ai}(\delta). \quad (23)$$

Эффективная диэлектрическая проницаемость линейного градиентного слоя:

$$\varepsilon_{G\text{eff}} = -\frac{1}{kx_L} \cdot \frac{\operatorname{Ai}'(\delta)}{\operatorname{Ai}(\delta)}. \quad (24)$$

Таким образом, используя выражения (3), (12) и (24), можно записать волну, локализованную вдоль

границы раздела самофокусирующей керровской нелинейной среды и линейного градиентного слоя, в виде:

$$\psi(x) = \begin{cases} \sqrt{\frac{2(n^2 - \varepsilon_{0N} - \varepsilon_{Geff})}{\alpha}} \cdot \frac{\text{Ai}(-x/x_L + \delta)}{\text{Ai}(\delta)}, & x < 0, \\ \sqrt{\frac{2}{\alpha}} \cdot \frac{q_N}{kchq_N(x - x_N)}, & x > 0, \end{cases} \quad (25)$$

где диэлектрическая проницаемость экспоненциально-го градиентного слоя определяется выражением (24).

2.2. Ступенчатая нелинейность

Пусть теперь с градиентным слоем контактирует нелинейная среда, диэлектрическая проницаемость которой описывается ступенчатой функцией (модель простейшей нелинейной среды [26] или модель «резкой ступеньки» [27]):

$$\varepsilon_N(|E|) = \begin{cases} \varepsilon_1, & |E| < E_s, \\ \varepsilon_2, & |E| > E_s, \end{cases} \quad (26)$$

где E_s – значение поля, при достижении которого происходит резкое (мгновенное) переключение от одного значения диэлектрической константы ε_1 к другому ε_2 ($\varepsilon_2 > \varepsilon_1$). Таким образом, вблизи контакта в нелинейной среде, где $|E| > E_s$, возникает область (приповерхностный домен) ширины x_s , в которой диэлектрическая константа имеет значение ε_2 , а за ее пределами, где $|E| < E_s$, диэлектрическая константа имеет значение ε_1 . Положение границы приповерхностного домена x_s , определяется условиями:

$$\begin{aligned} \psi_N(x_s + 0) &= \psi_N(x_s - 0) = E_s, \\ \psi'_N(x_s + 0) &= \psi'_N(x_s - 0). \end{aligned} \quad (27)$$

Как показано в [26], в модели ступенчатой нелинейности уравнение (5) с диэлектрической проницаемостью (26) распадается на два:

$$\psi''_N(x) - (n^2 - \varepsilon_1)k^2\psi_N(x) = 0, \quad |E| < E_s, \quad (28)$$

$$\psi''_N(x) + (\varepsilon_2 - n^2)k^2\psi_N(x) = 0, \quad |E| > E_s. \quad (29)$$

Решение уравнения (28) при $n^2 > \varepsilon_1$ имеет вид:

$$\psi_N(x) = E_s e^{-q_1(x-x_s)}, \quad (30)$$

где $q_1^2 = (n^2 - \varepsilon_1)k^2$, а решение уравнения (28) при $n^2 < \varepsilon_2$ имеет вид:

$$\psi_N(x) = \psi_0 \cos(p_2(x - x_m))/\cos(p_2x_m), \quad (31)$$

где $p_2^2 = (\varepsilon_2 - n^2)k^2$, а величины x_s, x_m определяются из граничных условий. Подставив решения (7), (30) и (31) в граничные условия (6) и (27), получим:

$$x_m = \frac{1}{k\sqrt{\varepsilon_2 - n^2}} \cdot \arctg\left(\frac{\varepsilon_{Geff}}{\varepsilon_2 - n^2}\right)^{1/2}, \quad (32)$$

$$x_s = x_m + q_1/p_2^2, \quad (33)$$

$$I_0 = \psi_0^2 = E_s^2 \cdot \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_{Geff} + \varepsilon_2 - n^2}, \quad (34)$$

где диэлектрическая проницаемость экспоненциального градиентного слоя определяется выражением (19). Таким образом, поверхностная волна, распространяющаяся вдоль границы раздела среды со ступенчатой нелинейностью и градиентного слоя с произвольным профилем, может быть записана в виде:

$$\psi(x) = E_s \cdot \begin{cases} \left(\frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_{Geff} + \varepsilon_2 - n^2}\right)^{1/2} \frac{F(g(x))}{F(g(0))}, & x < 0, \\ \left(\frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_2 - n^2}\right)^{1/2} \cos(p_2(x - x_m)), & 0 < x < x_s, \\ e^{-q_1(x-x_s)}, & x > x_s, \end{cases} \quad (35)$$

где положение границы приповерхностного домена определяется выражением (33).

В работе [22] был рассмотрен случай контакта среды со скачкообразной нелинейностью и среды с экспоненциальным профилем диэлектрической проницаемости. Было подробно проанализировано (и проиллюстрировано) влияние оптических параметров соответствующих волноводных структур на профили волн и их управляемую локализацию.

2.3. Обсуждение

Полученное выражение (20) для поверхностных волн отличается от приведенного в работе [14]. Волна (20) может распространяться при произвольно изменяемом в допустимой области значений эффективном показателе преломления (константе распространения). В [14] авторы приводят и анализируют дисперсионное уравнение, связывающее эффективный показатель преломления и оптические

характеристики волновода, т.е. его значение фиксировано при заданных значениях характеристик волновода. В частности, ими проанализирована зависимость эффективного показателя преломления n от параметра нелинейности αI_0 , представляющего собой произведение коэффициента керровской нелинейности на интенсивность поля на границе раздела сред.

Как известно, константа распространения связана с углом падения возбуждающего поверхностную волну луча, который может меняться в экспериментах. Поэтому такой параметр должен считаться управляющим, т.е. варьируемым в экспериментах на выбранной волноводной структуре. В данной работе получено выражение (15), определяющее зависимость параметра нелинейности αI_0 от эффективного показателя преломления n . Эта зависимость является, на наш взгляд, более практичной с экспериментальной точки зрения.

Также следует подчеркнуть отличие полученного выражения (25), описывающего волну, узколокализованную вдоль границы раздела самофокусирующей керровской нелинейной среды и линейного градиентного слоя. Полученный в работах [10, 12] профиль поверхностных волн, поперечный границе раздела таких сред, может выходить за пределы градиентного слоя. Нами же разработана теория для описания волн, пространственное распределение интенсивности которых полностью сосредоточено в градиентном слое и не выходит за его пределы.

Далее приведем примеры материалов, на которых могут быть основаны композитные волноводы, комбинирующие нелинейно-оптические слои и градиентные слои с пространственным распределением оптических характеристик.

В нелинейных оптических кристаллах локализация светового пучка происходит вследствие нелинейного отклика среды. При определенных условиях локализация поля происходит не только в среде с самофокусирующей нелинейностью, но и с дефокусирующей [28].

Диэлектрическая проницаемость, или квадрат показателя преломления, зависят от квадрата напряженности электрического поля, т.е. линейно от интенсивности света (керровская нелинейность), что характерно для кристаллов KDP и LiNbO_3 при определенных длинах волн. Такая форма нелинейного отклика наблюдается в многослойных нанокompозитных пленках Co/TiO_2 в диапазоне длин волн 400–1000 нм при температурах 10–50 °C [29]. В [30] отмечалось усиление эффекта Керра в тонких пленках на основе кобальта в том же диапазоне длин волн.

В [31, 32] указано, что высокоинтенсивный световой пучок может изменять оптические

характеристики кристаллов в узких областях вдоль направления его распространения. В результате наблюдалось образование приповерхностного слоя с отличающимися от остального кристалла оптическими характеристиками. Подобные изменения обусловлены нелинейным откликом кристалла, зависящим от интенсивности распространяющегося вдоль поверхности светового пучка. Теоретическое описание таких явлений основано на моделях, таких как «резкая» ступенчатая нелинейность [26, 33], «гладкая» ступенчатая нелинейность [27] и насыщаемая нелинейность в различных формулировках [34, 35].

Как отмечалось в п. 2.2, модель «резкой» ступенчатой нелинейности описывает изменение диэлектрической проницаемости скачком от одного постоянного значения к другому при достижении интенсивности светового пучка определенного порогового значения. В рамках такой модели многие авторы смогли получить в явном аналитическом виде результаты, примененные для теоретического описания особенностей распространения поверхностных волн [26], самоотражения [33], самолокализованных оптических импульсов [27], оптических бистабильностей [36]. Такая зависимость может рассматриваться как предельный случай модели «гладкой» ступенчатой нелинейности в случае резкого возрастания диэлектрической проницаемости при малом росте интенсивности светового пучка.

В [37] было показано, что изменение диэлектрической проницаемости полупроводника с экситон-экситонным взаимодействием в определенном спектральном диапазоне может происходить практически скачкообразно. Авторы [38, 39] отмечали, что такое поведение может наблюдаться в полупроводниковых кристаллах, таких как кристаллы CdS и CdSe с малой энергией связи биэкситона в диапазоне примерно 0.5–3.0 мэВ. Такое явление авторами объясняется тем, что когерентные фотоны, проходя через полупроводниковую пленку, возбуждают когерентные экситоны с такими же значениями волновых векторов и фаз, как и у фотонов. В результате такого оптического взаимодействия возникают биэкситоны, которые определяют как поляризацию кристалла, так и концентрацию квазичастиц, отвечающих за его оптические свойства. При этом оптические нелинейные эффекты могут наблюдаться при относительно малых интенсивностях падающего светового пучка.

Следует отметить, что в некоторых средах при интенсивном освещении происходит быстрое изменение их оптических свойств, в частности, в легированных полупроводниками стеклах CdSSe и Schott OG 550 [40, 41], легированных ионами

кристаллах $\text{GdAlO}_3:\text{Cr}^{3+}$ [42], тонких пленках, сформированных из фотохромного белка бактериородопсина [43]. В них наблюдалось в пределе при малых временах релаксации практически скачкообразное изменение показателя преломления от одного значения к другому в зависимости от интенсивности светового пучка.

С другой стороны, для многих полупроводниковых гетероструктур, используемых в современной оптоэлектронике, наблюдается зависимость показателя преломления или диэлектрической функции от пространственного расстояния [3]. В частности, показатели преломления полупроводниковых фотонных кристаллов GaAs/GaAlAs [44, 45], InGaAs/InAlAs [46], InGaAsP/InP [47] описываются пространственно-градиентными профилями. Поэтому их можно отнести к оптически градиентным средам. Следует отметить, что использование арсенида галлия и основанных на нем гетероструктур типа $\text{Ga}_{1-x}\text{Al}_x\text{As}$, $\text{Ga}_{1-x}\text{In}_x\text{As}$ и $\text{Ga}_{1-x}\text{Al}_x\text{N}$, $\text{Ga}_{1-x}\text{In}_x\text{N}$ представляется весьма перспективным в связи с их повышенной радиационной стойкостью [48] по сравнению с другими кристаллами, используемыми в полупроводниковой оптоэлектронике.

Пространственные распределения показателей преломления часто создаются путем имплантации в стекла ионов, которые индуцируют ионообменные процессы, создающие индукцию напряжений вблизи поверхности из-за большой разницы между ионными радиусами обмениваемых ионов. К примеру, градиентные профили показателя преломления формируются в стеклах ВК7 введением ионов K^+-Na^+ и в известково-натриевом стекле – ионов Ag^+-Na^+ [49]. Отметим, что к экспоненциальному профилю близкими являются профили, получаемые методом термодиффузии ионов металла в стеклянную подложку [50].

Таким образом, сочетание различных полупроводниковых кристаллов в композитном волноводе позволяет получить с одной стороны от волноводущего интерфейса нелинейно-оптический слой, а с другой – слой с градиентным профилем диэлектрической проницаемости.

ЗАКЛЮЧЕНИЕ

В работе проведено теоретическое описание волноводных свойств границы раздела двух сред с принципиально различными оптическими характеристиками. Сформулированная модель плоского волновода применима для сред с произвольным распределением пространственного профиля диэлектрической проницаемости.

Получено аналитическое выражение, описывающее поверхностную волну, распространяющуюся вдоль границы раздела среды со ступенчатой нелинейностью и градиентного слоя с произвольным профилем диэлектрической проницаемости. Также получены аналитические выражения для поверхностных волн, распространяющихся вдоль границы раздела среды с керровской нелинейностью (как самофокусирующей, так и дефокусирующей) с градиентными средами, характеризующимися экспоненциальным и линейным профилями диэлектрической проницаемости.

Приведенный анализ материалов показал, что возможно подобрать такие полупроводниковые кристаллы, на которых могут быть основаны композитные волноводы, комбинирующие нелинейно-оптические слои и градиентные слои с пространственным распределением оптических характеристик.

Полученные в работе результаты могут иметь значение при проектировании элементов оптических устройств, основанных на использовании возможностей управления локализацией световых пучков вдоль волноводущих поверхностей раздела контактирующих сред.

Благодарности

Работа выполнена с использованием оборудования центра высоких технологий БГТУ им. В.Г. Шухова.

Acknowledgments

The study was performed using the equipment of the High Technology Center at the V.G. Shukhov Belgorod State Technological University.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Adams M.J. *An Introduction to Optical Waveguides*. Chichester: Wiley; 1981. 401 p.
2. Chen C.-L. *Foundations for Guided-Wave Optics*. New York: John Wiley & Sons Inc.; 2005. 462 p. <https://doi.org/10.1002/0470042222>
3. Dragoman D., Dragoman M. *Advanced Optoelectronic devices*. Berlin: Springer; 1999. 424 p.
4. Bednarik M., Cervenka M. Electromagnetic waves in graded-index planar waveguides. *J. Opt. Soc. Am. B*. 2020;37(12):3631–3643. <https://doi.org/10.1364/JOSAB.408679>
5. Čada M., Qasymeh M., Pištora J. Optical Wave Propagation in Kerr Media. In: *Wave Propagation. Theories and Applications*. London: IntechOpen; 2013. P. 175–192. <http://doi.org/10.5772/51293>

6. Shvartsburg A.B., Maradudin A. *Waves in Gradient Metamaterials*. Singapore: World Scientific; 2013. 339 p. <https://doi.org/10.1142/8649>
7. Al-Bader S.J., Jamid H.A. Graded-index optical waveguides with nonlinear cladding. *J. Opt. Soc. Am. A*. 1988;5(3):374–379. <https://doi.org/10.1364/JOSAA.5.000374>
8. Taya S.A., Kullab H.M., Qadoura I.M. Dispersion properties of slab waveguides with double negative material guiding layer and nonlinear substrate. *J. Opt. Soc. Am. B*. 2013;30(7):2008–2013. <https://doi.org/10.1364/JOSAB.30.002008>
9. Almawgani A.H.M., Taya S.A., Hussein A.J., Colak I. Dispersion properties of a slab waveguide with a graded-index core layer and a nonlinear cladding using the WKB approximation method. *J. Opt. Soc. Am. B*. 2022;39(6):1606–1613. <https://doi.org/10.1364/JOSAB.458569>
10. Hussein A.J., Nassar Z.M., Taya S.A. Dispersion properties of slab waveguides with a linear graded-index film and a nonlinear substrate. *Microsyst. Technol.* 2021;27(7):2589–2594. <https://doi.org/10.1007/s00542-020-05016-z>
11. Taya S.A., Hussein A.J., Colak I. An exact solution of a slab waveguide dispersion relation with a linear graded-index guiding layer (TM case). *Microsyst. Technol.* 2022;28(22):1213–1219. <https://doi.org/10.1007/s00542-022-05281-0>
12. Taya S.A., Hussein A.J., Ramahi O.M., Colak I., Chaouche Y.B. Dispersion curves of a slab waveguide with a nonlinear covering medium and an exponential graded-index thin film (transverse magnetic case). *J. Opt. Soc. Am. B*. 2021;38(11):3237–3243. <https://doi.org/10.1364/JOSAB.439034>
13. Hussein A.J., Taya S.A., Vigneswaran D., Udiyakumar R., Upadhyay A., Anwar T., Amiri I.S. Universal dispersion curves of a planar waveguide with an exponential graded-index guiding layer and a nonlinear cladding. *Results in Physics*. 2021;20:103734. <https://doi.org/10.1016/j.rinp.2020.103734>
14. Savotchenko S.E. The surface waves propagating along the contact between the layer with the constant gradient of refractive index and photorefractive crystal. *J. Opt.* 2022;24(4):045501. <https://doi.org/10.1088/2040-8986/ac51e9>
15. Savotchenko S.E. The composite planar waveguide structure consisting of the linearly graded-index layer and the nonlinear layer formed with an increasing the electric field. *Optik*. 2022;252:168542. <https://doi.org/10.1016/j.ijleo.2021.168542>
16. Savotchenko S.E. Light localization in a linearly graded-index substrate covered by intensity dependent nonlinear self-focusing cladding. *J. Opt.* 2022;24(6):065503. <https://doi.org/10.1088/2040-8986/ac6bab>
17. Savotchenko S.E. Discrete spectrum of waveguide modes of a linearly graded-index film introduced into a medium with a stepwise nonlinearity. *Optik*. 2023;281(6):170835. <https://doi.org/10.1016/j.ijleo.2023.170835>
18. Savotchenko S.E. Guided waves in a graded-index substrate covered by an intensity-dependent defocusing nonlinear medium. *Appl. Phys. B: Lasers and Optics*. 2022;128(8):153. <https://doi.org/10.1007/s00340-022-07872-1>
19. Savotchenko S.E. Nonlinear surface waves propagating along the contact between the graded-index layer and the medium with near surface layer where Kerr nonlinearity disappears with increasing light intensity. *Optik*. 2023;272:170373. <https://doi.org/10.1016/j.ijleo.2022.170373>
20. Savotchenko S.E. Surface waves propagating along the interface between a parabolic graded-index medium and a self-focusing nonlinear medium: exact analytical solution. *J. Opt.* 2022;24(10):105501. <https://doi.org/10.1088/2040-8986/ac8e80>
21. Savotchenko S.E. Surface waves propagating along the interface between parabolic graded-index medium and photorefractive crystal with diffusion nonlinearity. *Phys. B: Condensed Matter*. 2023;648(2):414434. <https://doi.org/10.1016/j.physb.2022.414434>
22. Savotchenko S.E. Surface waves propagating along the interface separating an exponential graded-index medium and the medium with a step change in the dielectric constant. *Optik*. 2022;271(12):170092. <https://doi.org/10.1016/j.ijleo.2022.170092>
23. Savotchenko S.E. Waveguide properties of interface separating a photorefractive crystal with diffusion nonlinearity and an exponential graded-index medium. *Phys. Lett. A*. 2022;455(12):128516. <https://doi.org/10.1016/j.physleta.2022.128516>
24. Savotchenko S.E. New types of transverse electric nonlinear waves propagating along a linearly graded-index layer in a medium with Kerr nonlinearity. *Opt. Quant. Electron.* 2023;55(1):74. <https://doi.org/10.1007/s11082-022-04323-1>
25. Savotchenko S.E. Temperature controlled waveguide properties of the linearly graded-index film in semiconductor crystal with the photorefractive nonlinearity. *Appl. Phys. B: Lasers and Optics*. 2023;129(1):7. <https://doi.org/10.1007/s00340-022-07950-4>
26. Khadzhi P.I., Fedorov L.V., Torstveit S. Nonlinear surface waves for the simplest model of nonlinear medium. *Phys. Tech. Lett.* 1991;61:110–113.
27. Kaplan A.E. Multistable self-trapping of light and multistable soliton pulse propagation. *IEEE J. Quant. Electron.* 1985;21(9):1538–1543. <https://doi.org/10.1109/JQE.1985.1072828>
28. Kartashov Y.V., Malomed B.A., Torner L. Solitons in nonlinear lattices. *Rev. Mod. Phys.* 2011;83(1):247–305. <http://doi.org/10.1103/RevModPhys.83.247>
29. Laudyn U.A., Rutkowska K.A., Rutkowski R.T., Karpierz M.A., Woliński T.R., Wójcik J. Nonlinear effects in photonic crystal fibers filled with nematic liquid crystals. *Cent. Eur. J. Phys.* 2008;6(3):612–618. <https://doi.org/10.2478/s11534-008-0096-z>
30. Polyakov V.V., Polyakova K.P., Seredkin V.A., Patrino G.S. The enhanced magneto-optical Kerr effect in Co/TiO₂ multilayer films. *Tech. Phys. Lett.* 2012;38(10):921–923. <https://doi.org/10.1134/S1063785012100227>
31. Jarque E.C., Malyshev V.A. Nonlinear reflection from a dense saturable absorber: from stability to chaos. *Opt. Commun.* 1997;14291(3):66–70. [https://doi.org/10.1016/S0030-4018\(97\)00275-7](https://doi.org/10.1016/S0030-4018(97)00275-7)
32. Schuzgen A., Peyghambarian N., Hughes S. Doppler Shifted Self Reflection from a Semiconductor. *Phys. Stat. Sol. (B)*. 1999;206(1):125–130. [https://doi.org/10.1002/\(SICI\)1521-3951\(199803\)206:1<125::AID-PSSB125>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1521-3951(199803)206:1<125::AID-PSSB125>3.0.CO;2-8)

33. Ляхомская К.Д., Хаджи П.И. Эффект самоотражения и простейшие модели нелинейной среды. *Журн. техн. физики*. 2000;70(11):86–90.
Lyakhomskaya K.D., Khadzhi P.I. Self-reflection effect in naïve model of nonlinear media. *Tech. Phys.* 2000;45(11):1457–1461. <https://doi.org/10.1134/1.1325030> [Original Russian Text: Lyakhomskaya K.D., Khadzhi P.I. Self-reflection effect in naïve model of nonlinear media. *Zhurnal Tekhnicheskoi Fiziki*. 2000;70(11):86–90 (in Russ.).]
34. Christian J.M., McDonald G.S., Chamorro-Posada P. Bistable Helmholtz bright solitons in saturable materials. *J. Opt. Soc. Am. B*. 2009;26(12):2323–2330. <https://doi.org/10.1364/JOSAB.26.002323>
35. Korovai O.V. Nonlinear s-polarized quasi-surface waves in the symmetric structure with a metamaterial core. *Phys. Solid State*. 2015;57(7):1456–1462. <https://doi.org/10.1134/S1063783415070197>
36. Enns R.H., Rangnekar S.S., Kaplan A.E. Bistable-soliton pulse propagation: Stability aspect. *Phys. Rev. A*. 1987;36(3):1270–1279. <https://doi.org/10.1103/PhysRevA.36.1270>
37. Khadzhi P.I., Rusanov A.M., Gaivan S.L. Cavity-free optical bistability of a thin semiconductor film in the exciton region of the spectrum. 1999;29(6):539–541. <https://doi.org/10.1070/QE1999v029n06ABEH001526>
38. Khadzhi P.I., Gaivan S.L. Nonlinear interaction of an ultrashort light pulse with a thin semiconductor film under conditions of two-photon excitation of biexcitons. *Quantum Electron.* 1995;25(9):897–900. <https://doi.org/10.1070/QE1995v025n09ABEH000497>
39. Corovai A.V., Khadzhi P.I. Optical properties of a semiconductor upon two-photon excitation of biexcitons by a powerful pump pulse and one-photon probing in the M band. *Quantum Electron.* 2001;31(10):937–939. <https://doi.org/10.1070/QE2001v031n10ABEH002080>
40. Roussignol P., Ricard D., Flytzanis C. Nonlinear optical properties of commercial semiconductor-doped glasses. *Appl. Phys. A*. 1987;44:285–292. <https://doi.org/10.1007/BF00624594>
41. Vanhaudenarde A., Trespidi M., Frey R. Refractive-index changes during photodarkening in semiconductor-doped glasses. *J. Opt. Soc. Am. B*. 1994;11(8):1474–1479. <https://doi.org/10.1364/JOSAB.11.001474>
42. Catunda T., Cury L.A. Transverse self-phase modulation in ruby and $\text{GdAlO}_3:\text{Cr}^{3+}$ crystals. *J. Opt. Soc. Am. B*. 1990;7(8):1445–1455. <https://doi.org/10.1364/JOSAB.7.001445>
43. Wang S.Q., Wang X., Birge R., Downie J.D., Timucin D., Gary C. Propagation of a Gaussian beam in a bacteriorhodopsin film. *J. Opt. Soc. Am. B*. 1998;15(5):1602–1609. <https://doi.org/10.1364/JOSAB.15.001602>
44. Mendoza-Alvarez J.G., Nunes F.D., Patel N.B. Refractive index dependence on free carriers for GaAs. *J. Appl. Phys.* 1980;51(8):4365–4367. <https://doi.org/10.1063/1.328298>
45. Ravindran S., Datta A., Alameh K., Lee Y.T. GaAs based long-wavelength microring resonator optical switches utilising bias assisted carrier-injection induced refractive index change. *Opt. Express*. 2012;20(14):15610–15627. <https://doi.org/10.1364/OE.20.015610>
46. Zucker J.E., Chang T.Y., Wegener M., Sauer N.J., Jones K.L., Chemla D.S. Large refractive index changes in tunable-electron-density InGaAs/InAlAs quantum wells. *IEEE Photon. Technol. Lett.* 1990;2(1):29–31. <https://doi.org/10.1109/68.47032>
47. Ishida K., Nakamura H., Matsumura H. InGaAsP/InP optical switches using carrier induced refractive index change. *Appl. Phys. Lett.* 1987;50(3):141–142. <https://doi.org/10.1063/1.97695>
48. Вигдорович Е.Н. Радиационная стойкость эпитаксиальных структур на основе GaAs. *Russ. Technol. J.* 2019;7(3):41–49. <https://doi.org/10.32362/2500-316X-2019-7-3-41-49>
[Vigdorovich E.N. Radiation resistance of epitaxial structures based on GaAs. *Russ. Technol. J.* 2019;7(3):41–49 (in Russ.). <https://doi.org/10.32362/2500-316X-2019-7-3-41-49>]
49. Karasiński P., Rogoziński R. Influence of refractive profile shape on the distribution of modal attenuation in planar structures with absorption cover. *Opt. Commun.* 2007;269(1):76–88. <https://doi.org/10.1016/j.optcom.2006.07.067>
50. Shutyi A., Sementsov D., Kazakevich A.V., Sannikov D. Waveguide regimes of a graded-index planar waveguide with cladding. *Tech. Phys.* 1999;44(1):1329–1333. <https://doi.org/10.1134/1.1259518>

Об авторе

Савотченко Сергей Евгеньевич, д.ф.-м.н, доцент, профессор кафедры высшей математики, ФГБОУ ВО «Белгородский государственный технологический университет им. В.Г. Шухова» (308012, Россия, Белгород, ул. Костюкова, д. 46). E-mail: savotchenkose@mail.ru. Scopus Author ID 6603577988, SPIN-код РИНЦ 2552-4344, <https://orcid.org/0000-0002-7158-9145>

About the author

Sergey E. Savotchenko, Dr. Sci. (Phys.-Math.), Associate Professor, Professor, High Mathematics Department, V.G. Shukhov Belgorod State Technological University (46, Kostyukova ul., Belgorod, 308012 Russia). E-mail: savotchenkose@mail.ru. Scopus Author ID 6603577988, RSCI SPIN-code 2552-4344, <https://orcid.org/0000-0002-7158-9145>

Математическое моделирование
Mathematical modeling

УДК 621.391:53.08
<https://doi.org/10.32362/2500-316X-2023-11-4-94-104>



НАУЧНАЯ СТАТЬЯ

Восстановление смазанного фотографического изображения движущегося объекта, получаемого на пределе разрешающей способности

В.Б. Федоров[®],
С.Г. Харламов,
А.И. Стариковский

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

Резюме

Цели. Задача восстановления смазанного изображения движущегося объекта имеет большое практическое значение, в частности, при обработке изображений поверхности Земли, получаемых со спутников. Целью работы является исследование возможности повышения качества восстановления смазанных изображений, получаемых на пределе разрешающей способности фотоаппарата.

Методы. Использованы методы цифровой обработки сигналов, методы теории некорректных и плохо обусловленных задач.

Результаты. Предложен метод восстановления «смазанного» фотографического изображения движущегося объекта, отличающийся от традиционных подходов тем, что уравнение дискретной свертки, к решению которого сводится задача восстановления смазанного изображения, получается путем аппроксимации соответствующего интегрального уравнения на основе интерполяционного ряда Котельникова, а не на основе квадратурной формулы, как это делается традиционно. В работе получены формулы для вычисления ядра свертки, получаемой с применением интерполяционного ряда Котельникова. Как известно, задача обращения дискретной свертки относится к классу некорректных задач и требует регуляризации. Дано сравнение результатов восстановления смазанных изображений (с использованием регуляризации по Тихонову), осуществляемого как традиционным путем, т.е. с применением квадратурной формулы, так и предлагаемым способом, основывающимся на интерполяционном ряде Котельникова. Показано, что качество восстановления смазанного изображения в обоих случаях получается практически одинаковым. Однако использование квадратурной формулы предполагает, что величина «смаза» выражена целым числом пикселей, в то время как в случае использования ряда Котельникова эта величина может задаваться и долями пикселя.

Выводы. Показано, что дискретизацию свертки, описывающей искажение изображения типа «смаз», целесообразно осуществлять на основе интерполяционного ряда Котельникова в случае, когда осуществляется обработка смазанного изображения, получаемого на пределе разрешающей способности фотоаппарата. Это обусловлено тем, что в этом случае величина «смаза» может составлять доли пикселя. Такая ситуация характерна, например, для спутниковой фотосъемки поверхности Земли.

Ключевые слова: «смаз» изображения движущегося объекта, разрешающая способность, восстановление изображения, тихоновская регуляризация, обратная задача, интерполяционный ряд Котельникова, краевой эффект

• Поступила: 05.12.2022 • Доработана: 09.03.2023 • Принята к опубликованию: 06.06.2023

Для цитирования: Федоров В.Б., Харламов С.Г., Стариковский А.И. Восстановление смазанного фотографического изображения движущегося объекта, получаемого на пределе разрешающей способности. *Russ. Technol. J.* 2023; 11(4):94–104. <https://doi.org/10.32362/2500-316X-2023-11-4-94-104>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Restoration of a blurred photographic image of a moving object obtained at the resolution limit

Victor B. Fedorov[@],
Sergey G. Kharlamov,
Anatoly I. Starikovskiy

MIREA – Russian Technological University, Moscow, 119454 Russia

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

Abstract

Objectives. When processing images of the Earth's surface obtained from satellites, the problem of restoring a blurry image of a moving object is of great practical importance. The aim of this work is to study the possibility of improving the quality of restoration of blurry images obtained at the limit of the resolution of the camera.

Methods. Digital signal processing methods informed by the theory of incorrect and ill-conditioned problems were used.

Results. The proposed method for restoring a blurred photographic image of a moving object differs from traditional approaches in that the discrete convolution equation, to which the problem of restoring a blurred image is reduced, is obtained by approximating the corresponding integral equation based on the Kotelnikov interpolation series rather than on the traditional basis of the quadrature formula. In the work, formulas are obtained for calculating the kernel of the convolution obtained using the Kotelnikov interpolation series. The discrete convolution inversion problem, which belongs to the class of ill-posed problems, requires regularization. Results of traditional approaches to restoring blurred images using the quadrature formula with Tikhonov regularization and the proposed method based on the Kotelnikov interpolation series are compared. Although the quality of the blurred image restoration is almost the same in both cases, in the quadrature formula the blur value is expressed as an integer number of pixels, while, when using the Kotelnikov series, this value can also be specified in fractions of a pixel.

Conclusions. The expediency of discretizing the convolution describing the image distortion of the blur type on the basis of the Kotelnikov interpolation series when processing a blurred image obtained at the limit of the resolution of the camera is demonstrated. In this case, the amount of blur can be expressed in fractions of a pixel. This situation typically arises when processing satellite photography of the Earth's surface.

Keywords: blurred image of a moving object, resolution, image restoration, Tikhonov regularization, inverse problem, Kotelnikov interpolation series, edge effect

• Submitted: 05.12.2022 • Revised: 09.03.2023 • Accepted: 06.06.2023

For citation: Fedorov V.B., Kharlamov S.G., Starikovskiy A.I. Restoration of a blurred photographic image of a moving object obtained at the resolution limit. *Russ. Technol. J.* 2023;11(4):94–104. <https://doi.org/10.32362/2500-316X-2023-11-4-94-104>

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

Задача восстановления смазанного изображения движущегося объекта достаточно хорошо изучена¹ [1–14]. Так, в случае, когда скорость движения объекта априорно известна, ее решение сводится к обращению дискретной свертки.

При этом указанная дискретная свертка получается в результате аппроксимации соответствующей интегральной свертки. Эта аппроксимация интегральной свертки может быть осуществлена путем замены интеграла квадратурной формулой (обычно квадратурной формулой трапеций). Однако существует и другая возможность, предлагаемая в настоящей работе. Она основывается на замене подынтегральной функции ее интерполяционным рядом Котельникова. При этом подынтегральная функция выражает зависимость яркости точки изображения от ее координат.

Если рассматривается изображение, самые мелкие детали которого существенно больше размеров пикселя, то при его восстановлении после «смаза» величина «смаза» может задаваться с точностью до размера пикселя, т.е. выражаться целым числом пикселей. В этом случае, как показано в работе, результат восстановления практически не зависит от того, каким способом была дискретизирована свертка. Однако дело обстоит иначе, когда требуется восстановить смазанное изображение, полученное на пределе разрешающей способности фотоаппарата, т.е. когда самые мелкие детали изображения по размерам близки к размерам пикселя. В этом случае для обеспечения максимального разрешения может потребоваться выразить величину «смаза» в долях пикселя. Фактически это может означать, что смазанное изображение перед восстановлением должно быть проинтерполировано, т.е. виртуально представлено как бы более мелкими «пикселями». После чего точную величину «смаза» уже можно будет выразить целым числом этих более мелких

«пикселей» и затем применить традиционный метод восстановления.

Но такой подход сопряжен с дополнительным расходом памяти компьютера. Так, если размеры пикселя потребуются уменьшить в 10 раз, то размер соответствующего массива с изображением увеличится в 100 раз.

При этом предлагаемый метод аппроксимации интегральной свертки на основе интерполяционного ряда Котельникова допускает задавать величину «смаза» в долях пикселя, и, следовательно, интерполяция изображения не потребуется, т.е. не потребуется дополнительного расходования памяти компьютера.

МЕТОД

Основное интегральное уравнение

Основное интегральное уравнение, связывающее яркость точек смазанного изображения $Q(x, y)$ с яркостью точек восстанавливаемого изображения $P(x, y)$, имеет вид:

$$Q(x, y) = w^2 \int_0^T P(x - v_x t; y - v_y t) dt,$$

где (x, y) – декартовы координаты текущей точки, (v_x, v_y) – декартовы координаты скорости смещения изображения по поверхности ПЗС-матрицы (по предположению все точки изображения движутся с одинаковыми скоростями), w – размер пикселя, T – время экспозиции [1–3]. Если вертикальная составляющая скорости отсутствует, то имеем интегральное уравнение Фредгольма 1-го рода типа свертки:

$$Q(x) = \frac{w^2}{v_x} \int_0^{v_x T} P(x - \xi) d\xi. \quad (1)$$

Дискретизация свертки на основе квадратурной формулы

Заменяя интеграл в формуле (1) квадратурной формулой трапеций, получим

$$Q(x) \approx \frac{w^2}{v_x} \sum_{k=0}^n c_k P(x - kw),$$

¹ Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. *Цифровая обработка изображений в информационных системах*: учебное пособие. Новосибирск: Изд-во НГТУ; 2002. 352 с. [Gruzman I.S., Kirichuk V.S., Kosykh V.P., Peretyagin G.I., Spector A.A. *Digital Image Processing in Information Systems*: Tutorial. Novosibirsk: NSTU; 2002. 352 p. (in Russ.).]

где $c_0 = c_n = 0.5$, $c_1 = \dots = c_{n-1} = 1$ – коэффициенты квадратурной формулы; $n = v_x/w$ – величина смещения изображения за время экспозиции, выраженная целым числом пикселей; $k \in \{0, \dots, n\}$.

Или, поскольку в пределах одного пикселя яркость должна считаться постоянной, имеем

$$Q(m) \approx \sum_{k=0}^{\min\{n,m\}} \frac{w^2}{v_x} c_k P(m-k), \quad (2)$$

где m – номер пикселя, $m \in \{0, \dots, M-1\}$, M – число пикселей в одном ряду, $k \in \{0, \dots, m\}$, $Q[m] = Q(mw)$, $P[m] = P(mw)$.

Далее равенство в (2) будет считаться точным.

Свертку (2) можно записать также в матричном виде

$$\mathbf{Q} = \mathbf{A} \cdot \mathbf{P}, \quad (3)$$

где, например, при $M = 8$, $n = 3$ теплицева матрица $\mathbf{A}$ будет иметь вид:

$$\mathbf{A} = \frac{w^2}{v_x} \begin{bmatrix} 0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 0.5 & 0.0 & 0.0 & 0.0 \\ 0.5 & 1.0 & 1.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.5 & 1.0 & 1.0 & 0.5 & 0.0 \\ 0.0 & 0.0 & 0.5 & 1.0 & 1.0 & 0.5 \end{bmatrix}.$$

Дискретизация свертки на основе интерполяционного ряда Котельникова

Предполагая, что максимальная пространственная частота в преобразовании Фурье изображения $P(x)$ не превосходит $1/2w$, подынтегральную

функцию в (1) можно представить интерполяционным рядом Котельникова:

$$P(x) = \sum_{k=-\infty}^{+\infty} P[k] \operatorname{sinc}\left(\frac{x}{w} - k\right).$$

Подставляя это выражение в (1), получим

$$Q(x) = \frac{w^2}{v_x} \sum_{k=-\infty}^{+\infty} P[k] \int_0^{v_x T} \operatorname{sinc}\left(\frac{x}{w} - \frac{\xi}{w} - k\right) d\xi$$

или

$$Q[m] = \sum_{k=-\infty}^{+\infty} P[k] a[m-k] = \sum_{k=-\infty}^{+\infty} a[k] P[m-k].$$

Но поскольку фактически горизонтальные размеры изображения ограничены M пикселями, то приходим к приближенному равенству (которое далее условимся считать точным):

$$Q[m] \approx \sum_{k=0}^{M-1} P[k] a[m-k], \quad (4)$$

где ядро этой дискретной свертки определяется формулой

$$a[k] = \frac{w^2}{v_x \pi} \left(\operatorname{Si}\left(\pi \left(\frac{v_x T}{w} - k\right)\right) + \operatorname{Si}(\pi k) \right). \quad (5)$$

Соответствующие графики представлены на рис. 1.

Как видно из приведенных графиков, с ростом величины смещения изображения за время экспозиции ядро свертки (4) приближается к ядру свертки (2). Отметим

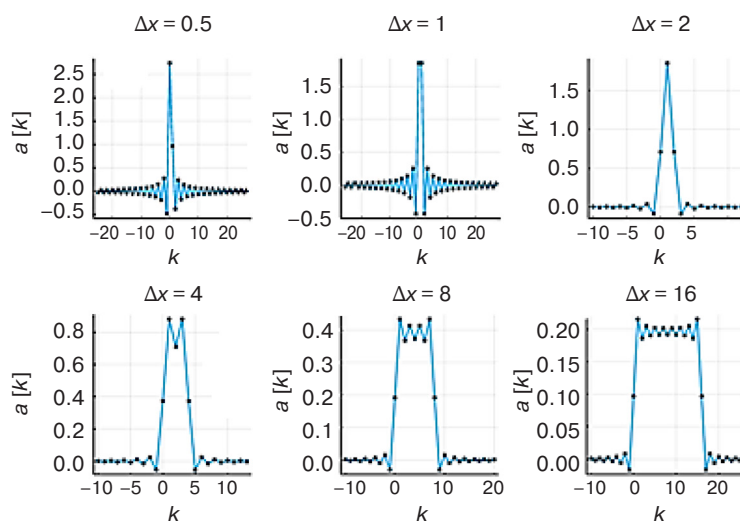


Рис. 1. Графики ядра дискретной свертки (4), полученной с использованием интерполяционной теоремы Котельникова, соответствующие различным величинам «смаза» изображения $\Delta x = v_x T/w$

при этом, что в случае свертки (4) величина смещения может выражаться не только целым числом пикселей, но и долями пикселя. В случае же свертки (2) эта величина всегда выражается целым числом пикселей.

Свертка (4) также может быть записана в матричном виде:

$$Q = A \cdot P. \quad (6)$$

Однако здесь, в отличие от (3), теплицева матрица A уже не будет нижней треугольной матрицей (у нее вообще не будет нулевых диагоналей).

Моделирование смазанного изображения (решение прямой задачи)

Для проведения численных экспериментов по восстановлению смазанного изображения необходимо располагать смазанными нужным образом изображениями. Получать такие изображения можно, используя соотношения (2) или (4), считая при этом функцию $P[m]$ заданной, а функцию $Q[m]$ – неизвестной. Указанные соотношения, как известно, могут быть сведены к циклической свертке, которая эффективно вычисляется (и обращается) на основе быстрого преобразования Фурье (БПФ). Например, на рис. 2 и 3 показано исходное и полученное указанным способом горизонтально смазанное на 20 пикселей изображение соответственно.

При моделировании смазанного изображения существенным моментом является наличие или отсутствие на смазанном изображении «переходных краев». У горизонтально смазанного изображения, представленного на рис. 3, такие «переходные края» (слева и справа) имеются. Каждый из этих «переходных краев» представляет собой полосу шириной, равной величине смещения изображения, яркость которой постепенно уменьшается до нуля.

Однако у реального смазанного фотографического изображения такие края будут отсутствовать, т.к. прообраз получаемого изображения представляет собой как бы бесконечную в оба конца ленту, у которой нет краев. Поэтому для приведения полученного смазанного изображения к реальности его края следует «обрезать».

РЕЗУЛЬТАТЫ

Результаты восстановления смазанного изображения без использования регуляризации

Обращение смоделированного смазанного изображения, также, как и решение прямой задачи, осуществимо на основе БПФ. Так, дискретную свертку вида (2) или (4) можно представить в виде циклической свертки, дополнив каждую из конечных последовательностей, входящих в свертку, требуемым числом нулей

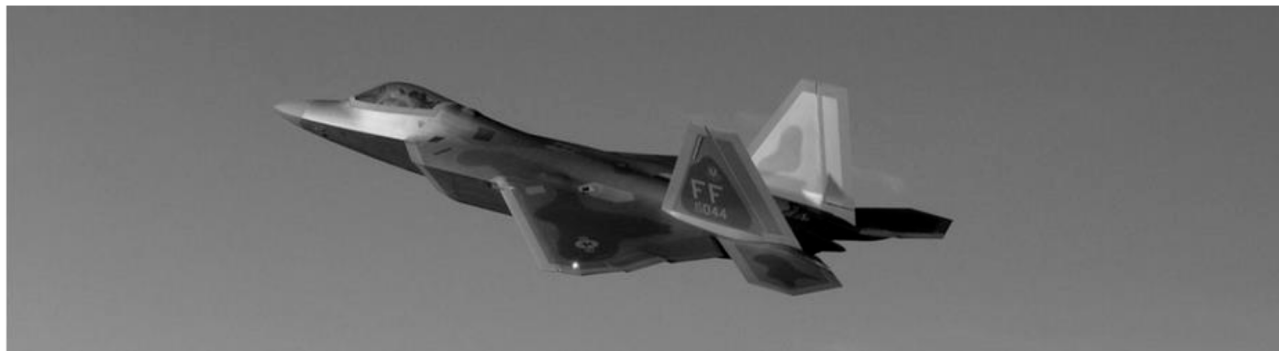


Рис. 2. Исходное изображение



Рис. 3. Смоделированное смазанное изображение с необрезанными краями (величина горизонтального «смаза» – 20 пикселей)

$$Q[m] = \sum_{k=0}^{M-1} P[k]a[m-k], \quad (7)$$

где Q , P , a – массивы длины, равной $\text{length}(P) + \text{length}(a) - 1$, получаемые из соответствующих массивов Q , P , a (входящих в (2) и (4)) добавлением требуемого числа нулей.

Тогда, для обращения этой свертки может быть использован так называемый инверсный фильтр, действие которого в частотной области имеет следующий вид:

$$\mathcal{F}[P] = \frac{1}{\mathcal{F}[a]} \mathcal{F}[Q],$$

где $\mathcal{F}[\cdot]$ обозначает дискретное преобразование Фурье, $1/\mathcal{F}[a]$ – так называемая передаточная функция инверсного фильтра.

Если при этом «переходные края» смоделированного смазанного изображения не обрезать, то результат восстановления будет получаться практически идеальным даже без применения регуляризации. Пример восстановленного таким образом изображения показан на рис. 4. Эксперименты показывают, что в случае, когда края смазанного изображения остаются не обрезанными, при величине смещения изображения $\Delta x = \frac{v_x T}{w} = 20$ для получения данного

результата длина «хвостов» ядра Котельникова должна быть не меньше 3. Под длиной «хвостов» понимается число учитываемых отсчетов ядра с отрицательными индексами, которое также равно числу учитываемых отсчетов с индексами, большими величины Δx .

Однако, как уже отмечалось, реальное смазанное изображение отличается тем, что не имеет «переходных краев». Если же в соответствии с реальностью получившиеся «переходные края» обрезать, то результат восстановления (без регуляризации) окажется чрезвычайно плохим. Соответствующий пример результата восстановления приведен на рис. 5.

Ситуацию можно попытаться исправить путем предварительной реконструкции обрезанных краев. Такая реконструкция может представлять собой просто добавление слева и справа к обрезанному изображению двух вертикальных полос шириной равной величине «смаза», яркость пикселей горизонтальных рядов которых будет увеличиваться/уменьшаться («по непрерывности») по линейному закону. На рис. 6 приведен пример смоделированного смазанного изображения с обрезанными «переходными краями», а на рис. 7 и 8 – результаты его восстановления после реставрации обрезанных краев.

Как можно видеть, результат восстановления путем обращения дискретной свертки с ядром Котельникова оказался значительно лучше. Однако, как показывают численные эксперименты, качество

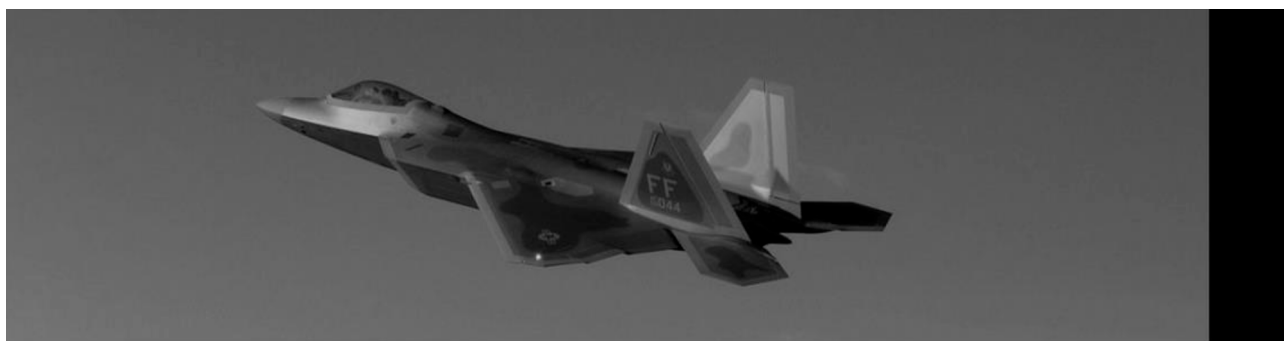


Рис. 4. Смазанное изображение с необрезанными краями, восстановленное с помощью инверсного фильтра при использовании ядра Котельникова и отсутствии регуляризации

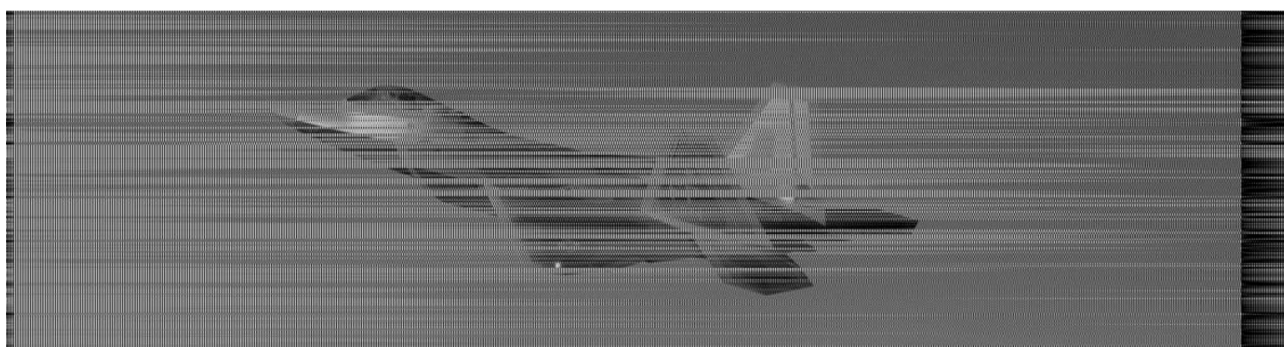


Рис. 5. Смазанное изображение с необрезанными краями, восстановленное с помощью инверсного фильтра при использовании квадратурной формулы трапеций и отсутствии регуляризации



Рис. 6. Смоделированное смазанное изображение с обрезанными «переходными краями» при величине горизонтального «смаза», равной 20 пикселям

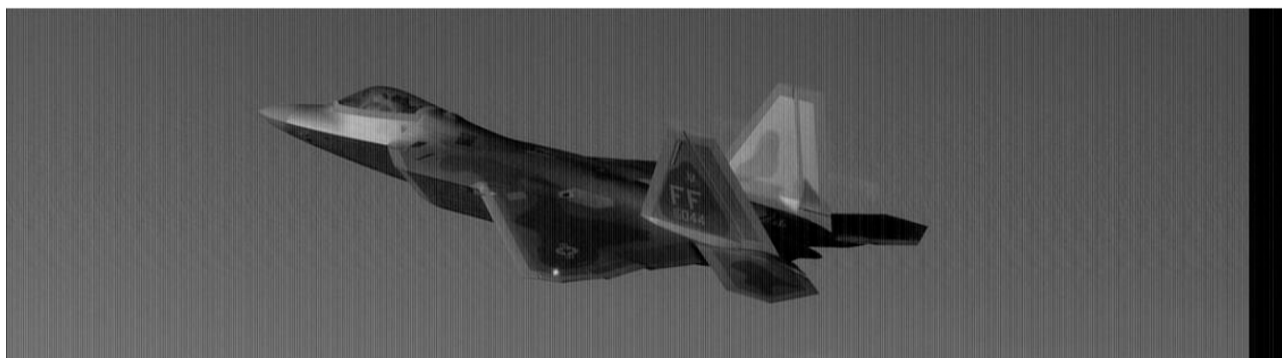


Рис. 7. Смазанное изображение с обрезанными краями, восстановленное с помощью инверсного фильтра при использовании ядра Котельникова и отсутствии регуляризации

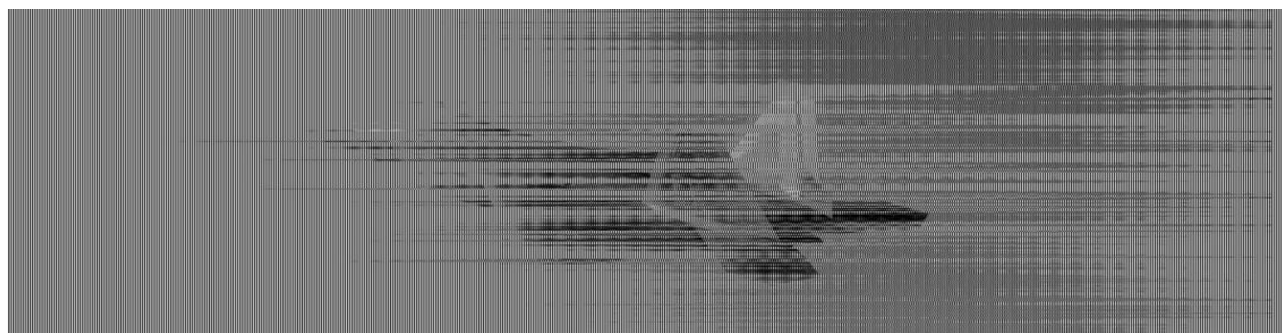


Рис. 8. Смазанное изображение с обрезанными краями, восстановленное с помощью инверсного фильтра при использовании квадратурной формулы трапеций и отсутствии регуляризации

восстановления существенно зависит от того, сколько отсчетов в «хвостах» ядра Котельникова, определяемого формулой (5), учитывается. В данном случае учитывались лишь отсчеты с индексами от 0 до 20 (при величине «смаза» 20 пикселей), т.е. «хвосты» ядра полностью отбрасывались. В целом нельзя утверждать, что обращение дискретной свертки с ядром Котельникова без применения регуляризации во всех случаях гарантирует удовлетворительный результат.

Как можно видеть из примеров, реконструкция краев несколько улучшает результат восстановления изображения, но в общем случае его качество

все равно остается неудовлетворительным. Причина этого в том, что обращение дискретных сверток (2) или (4) можно рассматривать как решение системы линейных алгебраических уравнений (СЛАУ) вида (3) или (6) соответственно.

Погрешности реконструкции (а тем более полное отсутствие такой реконструкции) надо рассматривать как погрешности в задании исходных данных (т.е. погрешностей в нескольких крайних значениях $Q[0], \dots, Q[M-1]$). Оценка относительной погрешности полученного решения, как известно, пропорциональна относительной погрешности правой части СЛАУ с коэффициентом

пропорциональности, равным числу обусловленности матрицы СЛАУ. А это число обычно весьма велико, его величина зависит как от размера матрицы, так и от величины «смаза». Таким образом, даже довольно малая относительная погрешность исходных данных может приводить к очень значительным относительным погрешностям получаемого решения. Для предотвращения этого явления при обращении свертки обычно используется метод регуляризации Тихонова [1–3].

Результаты восстановления смазанного изображения методом инверсной фильтрации с тихоновской регуляризацией

Задача восстановления изображения, как известно, является очень чувствительной к погрешностям исходных данных (решение интегрального уравнения Фредгольма 1-го рода (1) относится к классу некорректно поставленных задач). Поэтому при наличии сколько-нибудь существенных погрешностей в исходных данных обращение свертки требует применения регуляризации.

Передачная функция регуляризованного по Тихонову дискретного инверсного фильтра имеет вид:

$$\frac{A[m]}{|A[m]|^2 + \alpha W^{2p}[m]},$$

где $A = \mathcal{F}[a]$; W^{2p} – стабилизирующая функция порядка $p = 0, 1, 2, \dots$; α – параметр регуляризации ($\alpha \geq 0$, при $\alpha = 0$ регуляризации нет, этот параметр можно подбирать, например, экспериментально); массив a определен в формуле (7) [1–3].

Стабилизирующая функция нулевого порядка есть $W^0 = \mathcal{F}([1, 0, \dots, 0])$, т.е. тождественно равна 1. Регуляризирующая функция первого порядка определяется как $W^2 = \mathcal{F}([1, -2, 1, 0, \dots, 0])$ – преобразование Фурье конечно-разностного фильтра 2-го порядка. Регуляризирующие функции более высоких порядков – суть степени W^2 . Результаты восстановления с реконструкцией краев и применением регуляризации горизонтально смазанного на 20 пикселей изображения представлены на рис. 9 и 10.

Как можно видеть из полученных результатов, качество восстановления смазанного изображения, как в случае использования квадратурной формулы, так и в случае использования ряда Котельникова одинаково хорошее.

Результаты численных экспериментов по восстановлению изображения при величинах «смаза», выраженных долями пикселя

Рассмотрим теперь ситуацию, когда требуется восстановить смазанное изображение, полученное на пределе разрешающей способности фотоаппарата,

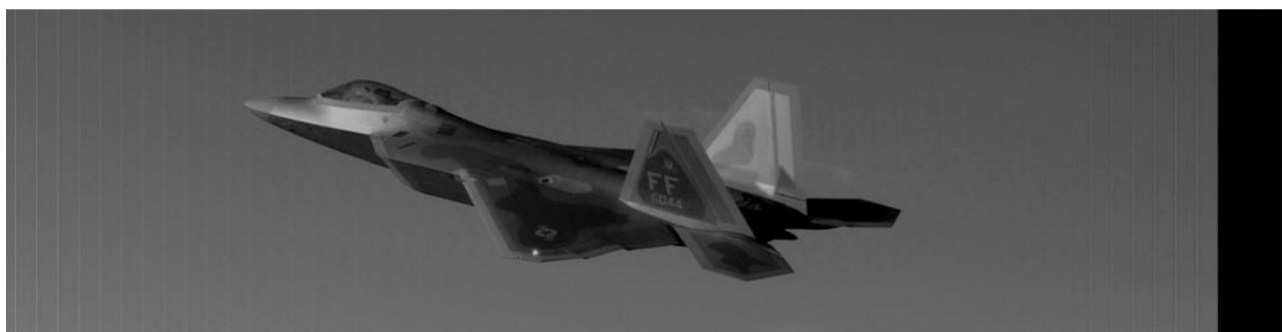


Рис. 9. Смазанное изображение после реставрации обрезанных краев, восстановленное с помощью инверсного фильтра при использовании ядра Котельникова ($\alpha = 10^{-3}$, $p = 1$)



Рис. 10. Смазанное изображение после реставрации обрезанных краев, восстановленное с помощью инверсного фильтра при использовании квадратурной формулы трапеций ($\alpha = 10^{-2}$, $p = 1$)

т.е. когда мгновенные образы самых мелких деталей движущегося изображения по размерам близки к размерам пикселя. В этом случае в качестве модели объекта съемки будем рассматривать миру, состоящую из параллельных полос уменьшающейся ширины, вплоть до размера одного пикселя (имеется в виду – в пространстве образов). На рис. 11 представлен мгновенный образ такой миру (с учетом конечности размеров пикселей ПЗС-матрицы фотоаппарата). На рис. 12 представлено горизонтально смазанное изображение на 5.5 пикселя той же самой миру.

На рис. 13 показано восстановленное смазанное изображение миру с необрезанными краями с помощью инверсного фильтра при использовании ядра Котельникова и отсутствии регуляризации. Для получения наилучшего результата восстановления при заданной величине «смаза» изображения длина «хвостов» ядра Котельникова должна подбираться экспериментально. Так, представленный на рис. 13 результат получен при длине «хвостов», равной 2.



Рис. 11. Мгновенное (несмазанное) изображение миру, получаемое на пределе разрешающей способности



Рис. 12. Смазанное по горизонтали на 5.5 пикселей изображение миру (края не обрезаны)



Рис. 13. Смазанное изображение миру с необрезанными краями, восстановленное с помощью инверсного фильтра при использовании ядра Котельникова и отсутствии регуляризации

Как видно из сравнения рис. 14 и 15, контрастность восстановленного смазанного изображения миру с использованием ядра Котельникова

получилась выше, чем при использовании квадратурной формулы, что неудивительно, поскольку величина «смаза» равна 5.5 пикселей, а при использовании квадратурной формулы возможно учитывать только смещение на целое число пикселей. При этом, в случае использования квадратурной формулы, инверсный фильтр был «настроен» не на величину «смаза», равной 5.5, а на величину 6 (т.к. при использовании квадратурной формулы нет возможности учитывать дробную часть величины «смаза»).



Рис. 14. Результат восстановления обрезанного смазанного изображения миру с помощью инверсного фильтра при использовании ядра Котельникова с предварительной реставрацией обрезанных краев ($\alpha = 10^{-3}$, $\rho = 1$)



Рис. 15. Результат восстановления обрезанного смазанного изображения миру с необрезанными краями с помощью инверсного фильтра при использовании квадратурной формулы трапеций с предварительной реставрацией обрезанных краев ($\alpha = 10^{-2}$, $\rho = 1$)

ЗАКЛЮЧЕНИЕ

Результаты численных экспериментов показывают, что в случае, если величина «смаза» изображения выражена целым числом пикселей, то предложенный подход, основанный на использовании ядра Котельникова, работает так же хорошо, как и традиционный, основанный на использовании квадратурных формул. Однако применение ряда Котельникова будет иметь преимущество в случае, если величина «смаза» изображения выражена долями пикселя, поскольку в этом случае для полного учета априорной информации о величине «смаза» исключается необходимость интерполяции исходного изображения, как это было бы при использовании квадратурных формул.

Таким образом, использование ядра Котельникова предпочтительно при обработке смазанного изображения, получаемого на пределе разрешающей

способности фотоаппарата, когда важно устранять «смаз», величина которого может составлять доли пикселя. Такая ситуация характерна, например, для спутниковой фотосъемки поверхности Земли.

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

1. Sizikov V., Dovgan A. Reconstruction of images smeared uniformly and non-uniformly. *CEUR Workshop Proceedings*. 2019;2344: paper 2. URL: <https://ceur-ws.org/Vol-2344/paper2.pdf>
2. Сизиков В.С. *Прямые и обратные задачи восстановления изображений, спектроскопии и томографии с MatLab*. СПб.: Лань; 2011. 410 с. ISBN 978-5-8114-2754-3
3. Василенко Г.И., Тараторин А.М. *Восстановление изображений*. М.: Радио и связь; 1986. 302 с.
4. Бакушинский А.Б., Гончарский А.В. *Некорректные задачи. Численные методы и приложения*. М.: Изд-во МГУ; 1989. 199 с.
5. Бейтс Р., Мак-Доннелл М. *Восстановление и реконструкция изображений*: пер с англ. М.: Мир; 1989. 336 с.
6. Гонсалес Р., Вудс Р. *Цифровая обработка изображений*. М.: Техносфера; 2012. 1104 с.
7. Гончарский А.В., Леонов А.С., Ягола А.Г. Методы решения интегральных уравнений Фредгольма 1-го рода типа свертки. В: *Некоторые вопросы автоматизированной обработки и интерпретации физических экспериментов*. Вып. 1. М.: Изд-во МГУ. 1973. С. 170–191.
8. Медофф Б.П. Реконструкция изображений по ограниченному данным: Теория и применение в компьютерной томографии. В: *Реконструкция изображений*; под ред. Г. Старка. М.: Мир; 1992. С. 384–436.
9. Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. *Регуляризирующие алгоритмы и априорная информация*. М.: Наука; 1983. 198 с.
10. Арефьева М.В., Сысоев А.Ф. Быстрые регуляризирующие алгоритмы цифрового восстановления изображений. В: *Вычислительные методы и программирование: сборник работ науч.-исслед. центра МГУ*. 1983. Вып. 39. С. 40–55.
11. Василенко Г.И. *Теория восстановления сигналов: О редукции к идеальному прибору в физике и технике*. М.: Сов. радио; 1979. 272 с.
12. Тихонов А.Н., Арсенин В.Я. *Методы решения некорректных задач*. М.: URSS; 2022. 288 с. ISBN 978-5-9710-9341-1
13. Тихонов А.Н., Гончарский А.В., Степанов В.В. Обратные задачи обработки фотоизображений. В: *Некорректные задачи естествознания*; под ред. А.Н. Тихонова, А.В. Гончарского. М.: Изд-во МГУ; 1987. С. 185–195.
14. Russ J.C. *The Image Processing Handbook*. Boca Raton: CRC Press; 2007. 852 p.

REFERENCES

1. Sizikov V., Dovgan A. Reconstruction of images smeared uniformly and non-uniformly. *CEUR Workshop Proceedings*. 2019;2344: paper 2. Available from URL: <https://ceur-ws.org/Vol-2344/paper2.pdf>
2. Sizikov V.S. *Pryamye i obratnye zadachi vosstanovleniya izobrazhenii, spektroskopii i tomografii s MatLab (Direct and Inverse Problems of Image Reconstruction, Spectroscopy and Tomography with MATLAB)*. St. Petersburg: Lan'; 2011. 410 p. (in Russ.). ISBN 978-5-8114-2754-3
3. Vasilenko G.I., Taratorin A.M. *Vosstanovlenie izobrazhenii (Image Recovery)*. Moscow: Radio i svyaz'; 1986. 302 p. (in Russ.).
4. Bakushinskii A.B., Goncharskii A.V. *Nekorrektnye zadachi. Chislennye metody i prilozheniya (Incorrect Tasks. Numerical Methods and Applications)*. Moscow: Moscow University Press; 1989. 199 p. (in Russ.).
5. Bates R., McDonnell M. *Vosstanovlenie i rekonstruktsiya izobrazhenii (Image Restoration and Reconstruction)*: transl. from Engl. Moscow: Mir; 1989. 336 p. (in Russ.). [Bates R., McDonnell M. *Image Restoration and Reconstruction*. NY: Oxford University Press; 1986. 312 p.]
6. Gonzalez R., Woods R. *Tsifrovaya obrabotka izobrazhenii (Digital Image Processing)*. Moscow: Tekhnosfera; 2012. 1104 p. (in Russ.). [Gonzalez R., Woods R. *Digital Image Processing*. Prentice Hall; 2008. 954 p.]
7. Goncharskii A.V., Leonov A.S., Yagola A.G. Methods for solving Fredholm integral equations of the 1st kind of convolution type. In: *Some Problems of Automated Processing and Interpretation of Physical Experiments*. V. 1. Moscow: Moscow University Press; 1973. P. 170–191 (in Russ.).
8. Medoff B.P. Image Reconstruction from Limited Data: Theory and Application in Computed Tomography. In: G. Stark (Ed.). *Image Reconstruction*. Moscow: Mir; 1992. P. 384–436 (in Russ.).
9. Tikhonov A.N., Goncharskii A.V., Stepanov V.V., Yagola A.G. *Regulyariziruyushchie algoritmy i apriornaya informatsiya (Regularizing Algorithms and A Priori Information)*. Moscow: Nauka; 1983. 198 p. (in Russ.).
10. Aref'eva M.V., Sysoev A.F. Fast regularizing algorithms for digital image restoration. In: *Vychislitel'nye metody i programmirovaniye (Numerical Methods and Programming)*. 1983. Is. 39. P. 40–55 (in Russ.).
11. Vasilenko G.I. *Teoriya vosstanovleniya signalov: O reduktsii k ideal'nomu priboru v fizike i tekhnike (Theory of Signal Recovery: On the Reduction to an Ideal Device in Physics and Technology)*. Moscow: Sovetskoe Radio; 1979. 272 p. (in Russ.).

12. Tikhonov A.N., Arsenin V.Ya. *Metody resheniya nekorrektnykh zadach (Methods for Solving Ill-Posed Problems)*. Moscow: URSS; 2022. 288 p. (in Russ.). ISBN 978-5-9710-9341-1
13. Tikhonov A.N., Goncharskii A.V., Stepanov V.V. Inverse problems of photo image processing. In: Tikhonov A.N., Goncharskii A.V. (Eds.). *Incorrect Problems of Natural Sciences*. Moscow: Moscow University Press; 1987. P. 185–195 (in Russ.).
14. Russ J.C. *The Image Processing Handbook*. Boca Raton: CRC Press; 2007. 852 p.

Об авторах

Федоров Виктор Борисович, к.т.н., доцент, кафедра высшей математики Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: feodorov@mirea.ru. Scopus Author ID 57208924592, SPIN-код РИНЦ 2622-7666, <https://orcid.org/0000-0003-1011-5453>

Харламов Сергей Григорьевич, магистрант, кафедра высшей математики Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: serhar2000@mail.ru. <https://orcid.org/0000-0003-4470-6323>

Стариковский Анатолий Иванович, к.т.н., доцент, профессор кафедры радиоэлектронных систем и комплексов Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: starikovski@mirea.ru. Scopus Author ID 57208926243, ResearcherID AAH-2239-2020, SPIN-код РИНЦ 1126-4471, <https://orcid.org/0000-0003-4040-3843>

About the authors

Victor B. Fedorov, Cand. Sci. (Eng.), Associate Professor, Department of Higher Mathematics, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: feodorov@mirea.ru. Scopus Author ID 57208924592, RSCI SPIN-code 2622-7666, <https://orcid.org/0000-0003-1011-5453>

Sergey G. Kharlamov, Master Student, Department of Higher Mathematics, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: serhar2000@mail.ru. <https://orcid.org/0000-0003-4470-6323>

Anatoly I. Starikovskiy, Cand. Sci. (Eng.), Associate Professor, 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: starikovski@mirea.ru. Scopus Author ID 57208926243, ResearcherID AAH-2239-2020, RSCI SPIN-code 1126-4471, <https://orcid.org/0000-0003-4040-3843>

Экономика наукоемких и высокотехнологичных предприятий и производств.
Управление в организационных системах
Economics of knowledge-intensive and high-tech enterprises and industries.
Management in organizational systems

УДК 658.51
<https://doi.org/10.32362/2500-316X-2023-11-4-105-115>



НАУЧНАЯ СТАТЬЯ

Организация инжинирингового центра для импортозамещения в промышленности

Д.Х. Михайлиди ¹,
А.В. Рагуткин ²,
Д.О. Скобелев ^{1, 2},
А.Б. Сухатерин ^{2, @}

¹ Научно-исследовательский институт «Центр экологической и промышленной политики», Москва, 115054 Россия

² МИРЭА – Российский технологический университет, Москва, 119454 Россия

@ Автор для переписки, e-mail: suhaterin@mirea.ru

Резюме

Цели. После введения санкций против Российской Федерации и запрета поставки иностранной электронной техники, в т.ч. систем автоматизации, российским предприятиям, важнейшей задачей развития науки и техники в России является обеспечение технологического суверенитета. Один из «кирпичиков» в фундаменте решения данной задачи – это импортозамещение. Согласно планам, разработанным Министерством промышленности и торговли Российской Федерации, курс на импортозамещение поможет произвести замену импортного оборудования отечественными аналогами.

Методы. Описываются подходы, положенные в основу совместного проекта РТУ МИРЭА и НИИ «ЦЭПП» по решению задач импортозамещения. Рассмотрены имеющиеся в мировом опыте стратегии замещения, а также объективные и субъективные препятствия для его проведения в России, среди которых недостаточная функциональность нормативно-правовых актов и сформированная привязанность к импортным технологиям и правилам. Показана особенность современных внешних взаимоотношений России как причина необходимости и срочности формирования технологического суверенитета. Описаны основные функциональные требования к программно-аппаратной платформе для построения современных автоматизированных систем управления (АСУ) для машиностроения, а также возможности инжинирингового центра для решения прикладных задач по преодолению импортозависимости. Показаны составные части производства средств производства (машиностроения) и его роль в жизненном цикле продукции.

Результаты. Обоснован выбор пилотного объекта инжиниринга – секционной стеклоформирующей машины, предмета разработки – программно-аппаратного комплекса, включающего элементы промышленной электроники и АСУ, показаны основные функциональные элементы АСУ и возникающие между ними связи.

Выводы. Подтверждается, что в создании цифровых продуктов необходимо добиться полного импортозамещения. Представлены перспективы сотрудничества с заинтересованными организациями.

Ключевые слова: жизненный цикл продукции, технологический суверенитет, импортозамещение, машиностроение, программно-аппаратная платформа, стеклоформирующая машина, автоматизированная система управления, реверс-инжиниринг

• Поступила: 24.04.2023 • Доработана: 08.06.2023 • Принята к опубликованию: 21.06.2023

Для цитирования: Михайлиди Д.Х., Рагуткин А.В., Скобелев Д.О., Сухатерин А.Б. Организация инженерингового центра для импортозамещения в промышленности. *Russ. Technol. J.* 2023;11(4):105–115. <https://doi.org/10.32362/2500-316X-2023-11-4-105-115>

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

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Organization of an engineering center for industrial import substitution

Dmitry Kh. Mikhailidi ¹,
Alexander V. Ragutkin ²,
Dmitry O. Skobelev ^{1, 2},
Alexey B. Sukhaterin ^{2, @}

¹ Environmental Industrial Policy Center, Moscow, 115054 Russia

² MIREA – Russian Technological University, Moscow, 119454 Russia

@ Corresponding author, e-mail: suhaterin@mirea.ru

Abstract

Objectives. Following the imposition of sanctions against the Russian Federation, which included a ban on the supply of foreign electronic equipment—including automation systems—to Russian enterprises, the continuing development of science and technology in Russia became a question of ensuring technological sovereignty according to the principle of import substitution. According to plans developed by the Ministry of Industry and Trade of the Russian Federation, the policy of import substitution, including automation systems, will ensure the replacement of imported equipment with domestic counterparts.

Methods. Approaches underlying the joint project of MIREA – Russian Technological University and Environmental Industrial Policy Center to solve the problems of import substitution are described. Various substitution strategies available in the world experience, as well as objective and subjective obstacles to their implementation in Russia, including the insufficiency of domestic regulatory legal acts and previously formed attachments to imported technologies and regulatory frameworks, are considered. Distinctive features of contemporary external relations are added to the necessity and urgency of developing technological sovereignty. The main functional requirements for a software and hardware platform for developing modern automated control systems (ACS) for mechanical engineering applications, as well as the required capabilities of an engineering center for solving applied problems of overcoming import dependence, are described. The components of the production of capital goods (engineering) and its role in the product life cycle are shown.

Results. The selection of a pilot engineering object comprising a sectional glass-forming machine, along with a software-hardware complex including elements of industrial electronics and ACS, is justified. The main functional elements of the ACS and their interrelations are shown.

Conclusions. The results confirm the necessity of achieving complete import substitution for the creation of digital products. Prospects for cooperation with interested organizations are shown.

Keywords: product life cycle, technological sovereignty, import substitution, mechanical engineering, software and hardware platform, glass-forming machine, automated control system, reverse engineering

• Submitted: 24.04.2023 • Revised: 08.06.2023 • Accepted: 21.06.2023

For citation: Mikhailidi D.Kh., Ragutkin A.V., Skobelev D.O., Sukhaterin A.B. Organization of an engineering center for industrial import substitution. *Russ. Technol. J.* 2023;11(4):105–115. <https://doi.org/10.32362/2500-316X-2023-11-4-105-115>

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

С точки зрения классической экономики затраты ресурсов на импортозамещение нельзя считать высокоэффективными. Объясняется это тем, что заказчик процесса будет вынужден тратить средства для повторного освоения уже существующего способа производства.

К причинам, побуждающим Российскую Федерацию к политике импортозамещения в 2023 г., стоит отнести не столько фундаментальные, проистекающие из возможности увеличить внутренний экономический потенциал, сколько ситуационные, вызванные санкциями, создающими экзистенциальную угрозу государству и его экономике [1]. Сложившаяся среда враждебных или приостановленных отношений со вчерашними экономическими партнерами диктует необходимость и срочность локализации полного спектра продукции. При этом критически важные виды производственной деятельности требуется ориентировать исключительно на внутренний рынок, воссоздавая производственные цепочки внутри страны [2].

В 2022 г. был принят целый блок мер экстренного реагирования на санкционную политику, который противоречит концепции импортозамещения, но позволяет избежать рисков лавинообразного падения производства и потребления. Так, разрешение на параллельный импорт без разрешения правообладателей и обнуление импортных пошлин способствуют как раз развитию импортозависимости, но представляются временно необходимым решением, т.к. обеспечивают временной лаг не только для создания и потребления отечественной продукции, но и для улучшения свойств

производимых объектов по сравнению с исходными образцами. Так формулируется императив для создания технологического суверенитета.

СОЗДАНИЕ СРЕДСТВ ПРОИЗВОДСТВА КАК ОСНОВА ТЕХНОЛОГИЧЕСКОГО СУВЕРЕНИТЕТА СТРАНЫ

Чаще всего страна специализируется на производстве тех продуктов, где имеет преимущества в обладании факторами производства [3]. Однако в теории сравнительных издержек [4, 5] считается, что страна не может обладать конкурентными преимуществами по всему спектру производимых локально, но имеющихся на глобальных рынках товаров. В условиях санкций потребность в собственном производстве увеличивается вне зависимости от наличия конкурентных преимуществ.

Анализ схемы индустриального жизненного цикла (рис. 1) показывает, что существует сектор производства средств производства или машиностроения, который имеет влияние на всю цепь трансформаций (переделов) вещества – от добычи сырья до утилизации продукта. Очевидно, что требуется синхронизация многовекторного замещения – средств производства и предметов потребления. В конечном итоге потребность в потреблении должна обеспечить эффективность импортозамещения средств производства.

Упомянутый сектор состоит из двух индустриальных частей (рис. 2): физической (машины, станки) и цифровой (сбор данных, управление процессами). Цифровая часть реализуется в физической части посредством программно-аппаратных

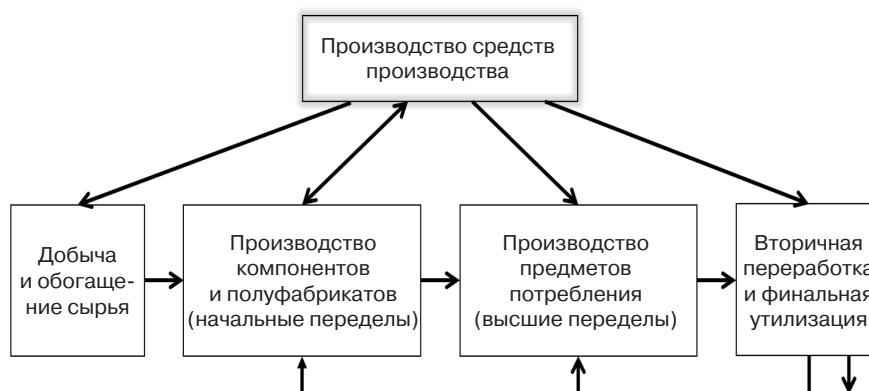


Рис. 1. Схема индустриального жизненного цикла продукции

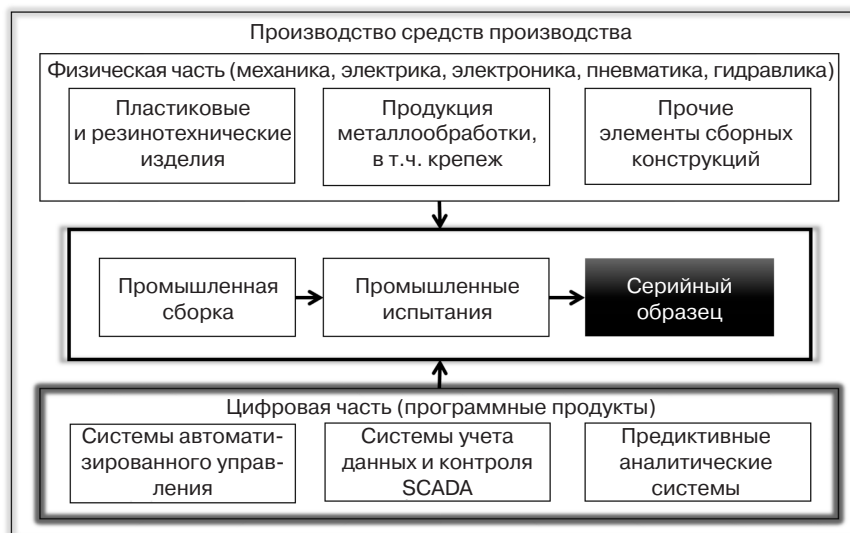


Рис. 2. Взаимодействие промышленных частей при производстве средств производства

платформ (ПАП), обеспечивающих надлежащую функциональность систем автоматического управления, опирающуюся на исполнительные устройства (клапаны, сервоприводы и др.). Для управления сложным производственным оборудованием используется система диспетчерского управления и сбора данных (supervisory control and data acquisition – SCADA), которая позволяет централизовать функции исполнительных устройств, а также обеспечивать производственную статистику для автоматизированных систем управления (АСУ) технологического процесса и для предиктивных аналитических систем.

Развитие машиностроения представляется целью национальной безопасности и основополагающим условием дальнейшего роста и развития страны. Зависимость от импорта 50–90%^{1, 2, 3, 4, 5} элементов и комплектующих оборудования, даже с учетом сокращения за последние 8 лет, выступает главным ограничителем развития внутренних производств и обеспечения экономического роста и, таким образом, определяет приоритеты импортозамещения, без

которых процесс будет происходить спонтанно, без должной эффективности [1].

Стадию проведения импортозамещения в 50–70 гг. XX века прошли большинство новых промышленных стран [6]. Существовало научное предположение о том, что экономическое положение стран-экспортеров сырьевых товаров (развивающихся государств) постепенно ухудшается, если они не прибегают к политике импортозамещения при индустриализации (гипотеза Пребиша – Зингера [7]). Южная Америка ориентировалась, в основном, на насыщение внутреннего рынка, Юго-Восточная Азия – на экспортные поставки. Экспортная стратегия, как показало время, оказалась более эффективной [8], но лишь при условии постоянного ослабления национальной валюты, что приводит к снижению добавочной стоимости изделия, высокой внутренней инфляции или ограничениям роста уровня жизни [9].

В России импортозамещение (там, где оно осуществлялось) опиралось на экспортную стратегию [10]. В ноябре 2015 г. Государственный совет РФ

¹ Приказ Минпромторга РФ от 07.07.2021 № 2486. [Order of the Ministry of Industry and Trade of the Russian Federation No. 2486 dated 07.07.2021.]. <https://rulings.ru/acts/Prikaz-Minpromtorga-Rossii-ot-07.07.2021-N-2486/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

² Приказ Минпромторга РФ от 02.08.2021 № 2913. [Order of the Ministry of Industry and Trade of the Russian Federation No. 2913 dated 02.08.2021.]. <https://rulings.ru/acts/Prikaz-Minpromtorga-Rossii-ot-02.08.2021-N-2913/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

³ Приказ Минпромторга РФ от 20.08.2021 № 3273. [Order of the Ministry of Industry and Trade of the Russian Federation No. 3273 dated 20.08.2021.]. <https://rulings.ru/acts/Prikaz-Minpromtorga-Rossii-ot-20.08.2021-N-3273/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

⁴ Приказ Минпромторга РФ от 30.07.2021 № 2882. [Order of the Ministry of Industry and Trade of the Russian Federation No. 2882 dated 30.07.2021.]. <https://rulings.ru/acts/Prikaz-Minpromtorga-Rossii-ot-30.07.2021-N-2882/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

⁵ Приказ Минпромторга РФ от 30.07.2021 № 2881. [Order of the Ministry of Industry and Trade of the Russian Federation No. 2881 dated 30.07.2021.]. <https://legalacts.ru/doc/prikaz-minpromtorga-rossii-ot-30072021-n-2881-ob-utverzhenii/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

принял базовые решения по данной проблематике. В рамках действий было предложено:

- обеспечить процесс замещения импорта финансовыми ресурсами, в основном за счет изменения налогообложения, а также за счет средств Фонда развития промышленности и Федеральной корпорации по развитию малого и среднего бизнеса, включая «проектное финансирование» и предоставление государственных гарантий;
- полностью ориентировать систему государственных закупок на приобретение отечественного оборудования, применяя специальные стимулы для загрузки российских производственных мощностей;
- создать систему централизованного управления импортозамещением и соответствующего контроля, которая позволяла бы при наличии российских аналогов блокировать покупку иностранных продуктов⁶.

Во исполнение этих задач приняты нормативно-правовые акты (НПА), в частности ПП РФ № 208⁷ и ПП РФ № 209⁸ от 18.02.2022 г. При всех очевидных достоинствах следует обратить внимание на их недостатки. Так, критерием результативности в ПП РФ № 209 является количество созданных комплектов конструкторской документации, что не может отражать экономическую результативность. Постановление Правительства РФ № 208 представляется более совершенным механизмом, однако там отсутствуют механизмы вертикальной интеграции и агрегации объектов разработки по различным признакам, а каждое комплектующее разрабатывается отдельными коллективами, что сильно повышает удельные затраты на разработку изделия в целом [11].

В России поддерживаются локальные разработки программного обеспечения и электроники. Кроме того, реализуются программы по импортозамещению в рамках национального проекта «Малое и среднее предпринимательство», разделение усилий

происходит по двум направлениям: развертывание серийного производства, производство уникального оборудования и создание новых технологий, облегчающих внедрение ноу-хау и наукоемких средств производства [12].

Целью принятых НПА является опережающее развертывание внутри страны производств, создающих продукцию, заменяющую импортные аналоги и превосходящую последние по потребительским характеристикам, освоение новых технологий (рис. 3) [13]. Для российской экономики это касается не только изделий, но и технологий и институтов (правил), импорт которых в значительной степени обусловил зависимость развития от глобальных центров инженерных разработок. Подобная стратегия развития навязывалась не только в промышленности, но и в образовании и науке, которые в своем развитии привязаны к внешним оценкам и стандартам, и результатом стала «модель неявных санкций» [1], реализованная задолго до современной конфронтационной эпохи.



Рис. 3. Эволюция технологий по О.С. Сухареву [14]

Сформированная привязанность к импортным технологиям и правилам является одним из главных препятствий к их замещению. Этот управленческий эффект связан с содержанием ядра технологий, следовательно, отечественным разработчикам

⁶ Комиссии Госсовета обсудили идею создания единого сервиса по импортозамещению. [State Council commissions discussed the idea of creating a unified service for import substitution.]. <https://tass.ru/ekonomika/14848255>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

⁷ О предоставлении субсидии из федерального бюджета автономной некоммерческой организации «Агентство по технологическому развитию» на поддержку проектов, предусматривающих разработку конструкторской документации на комплектующие изделия, необходимые для отраслей промышленности. Постановление Правительства РФ № 208 от 18.02.2022. [On the granting of subsidies from the federal budget to the autonomous non-profit organization “Agency for Technological Development” to support projects involving the development of design documentation for components required for industries. Decree of the Government of the Russian Federation No. 208 dated 18.02.2022.]. <http://static.government.ru/media/files/zISr7dzERAaYQY0N2HBwbN4FoBah9M6Y.pdf>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

⁸ О предоставлении грантов в форме субсидий из федерального бюджета на реализацию проектов по созданию и (или) развитию центров инженерных разработок на базе образовательных организаций высшего образования и научных организаций, реализующих проекты, связанные с разработкой комплектующих. Постановление Правительства РФ № 209 от 18.02.2022. [On providing grants in the form of subsidies from the federal budget for the implementation of projects to create and (or) develop engineering development centers on the basis of educational organizations of higher education and scientific organizations implementing projects related to the development of components. Decree of the Government of the Russian Federation No. 209 dated 18.02.2022.]. <http://government.ru/docs/all/139438/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

необходимо создать такое ядро и ввести в действие в условиях острой конкуренции с работающей, пусть и за счет параллельного импорта, инфраструктурой [1]. Обработывающие производства, как правило, стремятся исключить из повседневной деятельности риски, связанные с необходимостью повышения квалификации, проведения научно-исследовательских опытно-конструкторских разработок и других инноваций, требующих больших временных затрат. Именно это объясняет инерционную реакцию на введение санкций, поскольку проще найти параллельный импорт, чем заставить элементы сложной системы локализовать технологию и оборудование.

Это составляет фундаментальное ограничение в сфере импортозамещения. Замещение технологии дает возможность создать собственный продукт, но, если сохранятся действующие правила ориентации хозяйствования на импортный продукт, как на лучший по умолчанию, то это будет продолжать снижать конкурентный потенциал отдельной отрасли и экономики в целом. Поэтому центрами компетенций по преодолению импортозависимости становятся инжиниринговые центры – элементы инновационного развития [15].

Проекты импортозамещения в современных условиях должны быть нацелены не только на обновление основных фондов, на повторение уже достигнутого технологического уровня, а в значительной степени ориентироваться на создание основы опережающего развития и модернизации реального сектора экономики.

Для большинства направлений производства средств производства значительную сложность представляют вопросы создания систем автоматического и автоматизированного управления. За прошедшие десятилетия стало традицией ориентироваться на микроэлектронные и программные решения в этой области, производимые за рубежом, что привело к текущей неопределенности в поддержании работоспособности производственного оборудования. Недостаточная развитость российского парка промышленных контроллеров, периферии и средств разработки не позволяет быстро решать задачи по локализации машиностроения.

Особенно остро стоит проблема создания ПАП, представляющей собой совокупность аппаратных и программных средств, которые осуществляют контроль и управление производственными и технологическими процессами, поддерживают обратную связь и активно воздействуют на ход процесса при отклонении его от заявленных параметров, а также обеспечивают регулирование и оптимизацию управляемого процесса.

ПАП включает в себя разветвленную структуру с распределенным вводом-выводом и централизованной обработкой сигналов и должна обеспечивать физическую взаимосвязь исполнительных механизмов и датчиков технологического оборудования. Сигнальные линии в зависимости от типа подключены к модулям ввода-вывода аналоговых и дискретных сигналов, соединенных промышленной шиной с программируемым логическим контроллером (ПЛК). Для повышения отказоустойчивости могут быть использованы 2 промышленных модуля ПЛК, поддерживающие горячее резервирование. В случае обрыва кабеля или выхода из строя одного из устройств удаленного ввода-вывода должно требоваться минимальное время восстановления работоспособности всей системы. По промышленным протоколам, например, Modbus, Industrial Ethernet и т.п. ПЛК обеспечивает информационное взаимодействие с центральной системой управления (возможно, через комплекс противоаварийной защиты).

Для программирования ПЛК используются интегрированные средства разработки приложений (например, *Codesys*, *ISAGRAfi* и др.), поддерживающие все 5 языков программирования стандарта МЭК 61131-3⁹ (LD, FBD, IL, ST, SFC). Таким образом, через промышленную сеть ПЛК отвечает за получение в реальном времени информации с датчиков, преобразование ее и обмен с другими компонентами системы автоматизации, а также для управления исполнительными механизмами. Далее через магистральную сеть происходит обмен данными между контроллерами и станциями оператора. Операторский уровень включает серверы и пользовательские автоматизированные рабочие места для проведения мониторинга оперативного технологического и производственного процесса и, при необходимости, передачи команд на изменение параметров.

В машиностроении ПАП реализуется в виде разработки и внедрения прикладных программно-аппаратных комплексов (ПАК) с различной степенью универсализации. РТУ МИРЭА и Научно-исследовательский институт «Центр экологической и промышленной политики» начинают совместный проект по отработке функционала такой платформы на примере инжиниринга АСУ стеклоформирующей машины (СФМ) (рис. 4). В отличие от физической части оборудования, где существуют объективные

⁹ Национальный стандарт Российской Федерации. *Контроллеры программируемые. Ч. 3. Языки программирования*. [National Standard of the Russian Federation. *Programmable Controllers. Part 3. Programming languages*.]. <https://docs.cntd.ru/document/1200135008>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

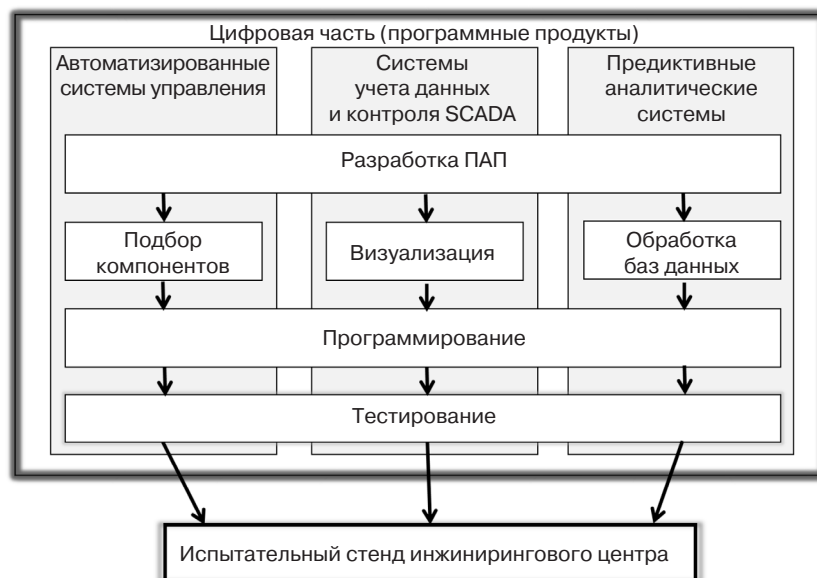


Рис. 4. Компетенции инжинирингового центра

ограничения для достижения 100%-ной независимости (например, натуральный каучук – составной компонент многих резинотехнических изделий – не может быть произведен в России по климатическим причинам), потенциал импортозамещения цифровой части практически не имеет ограничений и признается приоритетным в разработке по причинам производственной и информационной безопасности, соответствия требованиям наилучших доступных технологий.

На основе научного опыта, тесных связей с отраслевыми предприятиями, прежде всего в области соответствия наилучшим доступным технологиям, выбран объект инжиниринга. В сотрудничестве с одним из крупнейших предприятий стекольной промышленности стала доступна восьмисекционная СФМ, для которой разрабатывается программа реверс-инжиниринга (рис. 5). Наша потенциально совместная разработка АСУ СФМ относится к импортозамещению в области машиностроения (производство



Рис. 5. Иерархическая схема реверс-инжиниринга

средств производства) и находится в приоритетном перечне промышленной политики, поскольку машиностроение входит в основополагающие условия для роста и развития экономики страны.

Под АСУ обычно понимается целостное решение, обеспечивающее автоматизацию технологических процессов. Понятие «автоматизированный», в отличие от понятия «автоматический», подчеркивает необходимость участия человека в отдельных операциях, как в целях сохранения контроля над процессом, так и в связи со сложностью или нецелесообразностью автоматизации отдельных операций. АСУ обеспечивает безопасность операций за счет высоконадежных средств сигнализации, блокировок и защит с минимальным периодом реагирования, точного выполнения требований технологического регламента, исключения ошибочных действий оперативного персонала при ведении технологического процесса, обеспечения надежной работы оборудования и предотвращения аварийных ситуаций.

К уровню АСУ относится разработка алгоритма управления, реализующего сбор информации от первичных датчиков и исполнительных устройств, обработка полученной информации, автоматическое регулирование (поддержание технологических параметров на заданном значении), ручное и/или автоматическое управление электроприводами и технологическим оборудованием, передача информации от контроллеров и систем автоматики на верхний уровень. Для информационной связи всех подсистем используются промышленные сети.

В рамках работ по импортозамещению основной целью является разработка системы АСУ (программная прикладная часть), внедрение технических средств автоматизации для существующего в Российской Федерации парка секционных СФМ на российской элементной базе и использование прикладных и системных продуктов из реестра российского программного обеспечения¹⁰. Следующим шагом становится адаптация технологического оборудования машины, включающего пульта управления, датчики параметров сред, электромеханические клапаны, серводвигатели, шаговые двигатели, насосы и т.д., с постепенным переходом к использованию появляющейся российской продукции. Результатом разработки системы будет ПАК СФМ, который объединит прикладное программное обеспечение, обеспечивающее алгоритмы безаварийного управления исполнительными механизмами и контроля параметров машины. Программное обеспечение будет реализовано на ПЛК управления секциями (секционные контроллеры), устройствах

задания последовательностей синхроимпульсов центральным контроллером, управления скоростью и положением частотных преобразователей посредством SCADA (рис. 6).

SCADA является программным инструментом, цель которого состоит в автоматизации контроля над технологическими процессами. Корректировка параметров осуществляется в режиме реального времени. Оператор получает исчерпывающую и достоверную информацию об объектах и с помощью эффективных средств выполняет необходимые действия. В промышленности востребованы функционалы мониторинга, управления, архивирования полученных данных, оповещения, ведения отчетности. Одним из ключевых инструментов программного комплекса является иерархия уровней доступа. Наиболее распространенными SCADA на нашем рынке были продукты компаний Siemens¹¹, Aveva¹², GE Digital¹³ и др. В настоящее время улучшаются функциональные возможности, в т.ч. поддержка объектно-ориентированных моделей производства, российских аналогов, реализованных на платформе операционной системы Astra Linux¹⁴.

Этап прототипирования и испытания алгоритмов АСУ быстрее и дешевле всего проходит в ходе опытных (полунатурных) испытаний алгоритмов системы на электромеханическом стенде одной секции с элементами цифрового двойника, виртуально имитирующего работу других секций.

В развитие систем, созданных на основе объекта сравнения, планируется разработка версий ПАК, адаптированных под разные варианты исполнения СФМ, с различной степенью локализации производства элементов физической части машины. Решение задач требует объединения усилий заинтересованных сторон. На этапе разработки ПАП финансирование проводится за счет собственных средств. Текущая стоимость подобных систем занимает значительную долю в стоимости оборудования, поэтому замещающий продукт обладает весомым коммерческим потенциалом.

Важное значение придается возможности практического обучения на базе Инжинирингового центра как студентов российских ВУЗов, так и персонала предприятий-заказчиков. Для сокращения времени ознакомления интерфейс SCADA проектируется по образу того, который уже используется на действующих СФМ.

¹¹ <https://www.siemens.com/global/en.html>. Дата обращения 04.04.2023. / Accessed April 04, 2023.

¹² <https://www.aveva.com/>. Дата обращения 04.04.2023. / Accessed April 04, 2023.

¹³ <https://www.ge.com/digital/>. Дата обращения 04.04.2023. / Accessed April 04, 2023.

¹⁴ <https://astralinux.ru/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

¹⁰ <https://reestr.digital.gov.ru/>. Дата обращения 04.04.2023. / Accessed April 04, 2023 (in Russ.).

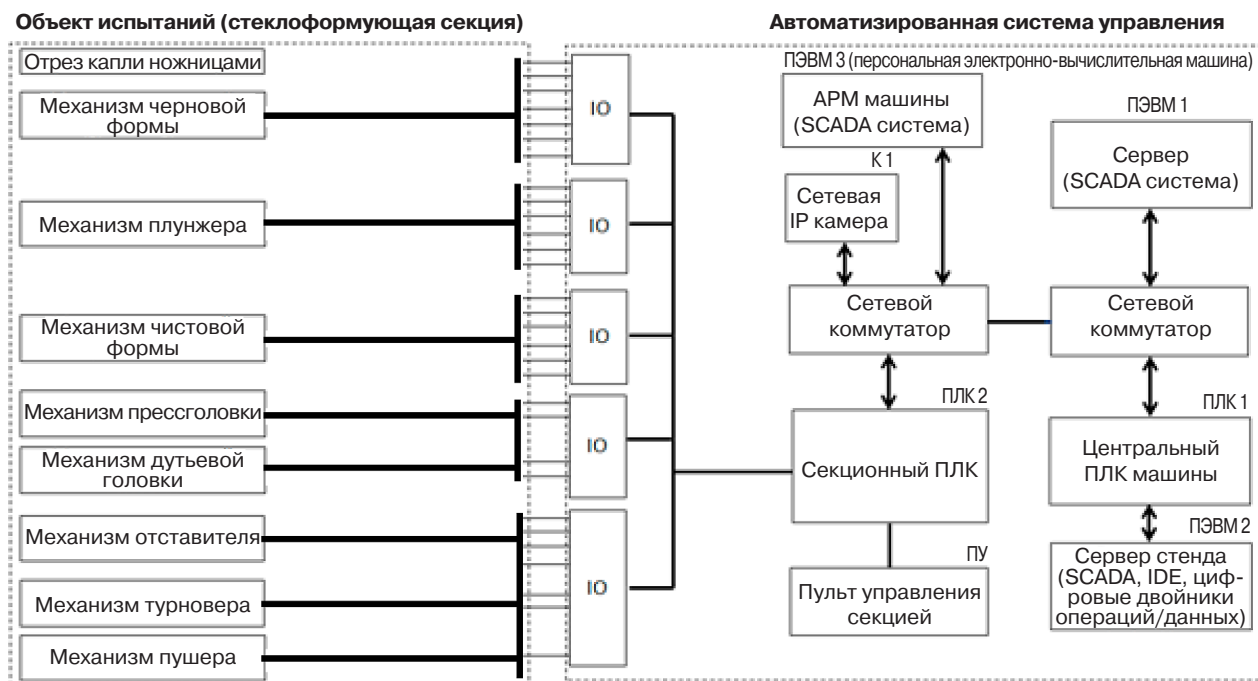


Рис. 6. Функциональная схема по разработке стенда АСУ. IO (input/output) – ввод/вывод; АРМ – автоматизированное рабочее место, IDE (integrated development environment) – интегрированная среда разработки

ЗАКЛЮЧЕНИЕ

По результатам исследований проблематики преодоления импортозависимости и создания технологического суверенитета авторы пришли к следующим выводам:

1. Разработка платформы АСУ СФМ относится к импортозамещению в области машиностроения и находится в приоритетном перечне промышленной политики России.
2. Параллельный импорт – это временное смягчение, но не решение проблемы. Критически важные составляющие, к которым относится производство электронных компонентов и программного обеспечения, возможно и необходимо замыкать на внутренний рынок, воссоздавая производственные цепочки внутри страны.
3. Технология – это не только оборудование, но и носители знания. Следует уделять ключевую роль подготовке кадров. Инжиниринговые центры становятся элементами инновационного развития и имеют потенциал дополнительного образования для студентов и отраслевых специалистов.
4. В основе технологического суверенитета должны лежать технологии и продукты, которые аналогичны или превосходят лучшие мировые образцы и практики.
5. Разработчики технологий и оборудования должны сотрудничать и координировать свои усилия

для максимальной эффективности как инноваций, так и обратных разработок (реверс-инжиниринга). Поскольку речь идет о производстве средств производства, а не о продукции массового потребления, считаем нецелесообразным параллельные разработки одних и тех же элементов (физических и цифровых) разными организациями. Объединение производственных заказов и их долгосрочное планирование позволят снизить себестоимость производственных машин, что скажется на экономике продукции массового спроса.

6. Разработка и внедрение АСУ технологических процессов имеют весомые перспективы монетизации, но начальная стадия исследований в отсутствии инновационных инвесторов, как правило, проводится за счет самофинансирования. Существующие методы государственной поддержки недостаточно согласуются между собой и характеризуются отсутствием планирования производств и формирования централизованных заказов, которые способны повлиять на стоимость элемента и изделия.
7. Требуется повысить статус проектов по импортозамещению и, по возможности, централизовать их проведение и финансирование, выделив в отдельный национальный проект, целью которого являются фиксация достижений отдельных отраслей и их масштабирование на индустрию в целом.

Вклад авторов

Д.О. Скобелев, А.В. Рагуткин – идея исследования, постановка задачи исследования, итоговое редактирование текста статьи.

Д.Х. Михайлиди (75%), А.Б. Сухатерин (25%) – выполнение рутинной работы по систематизации материала, анализ результатов исследования и подготовка данных, написание текста статьи.

Authors' contributions

D.O. Skobelev, A.V. Ragutkin – the research concept, statement of the research problem, and final editing the text of the article.

D.Kh. Mikhailidi (75%), A.B. Sukhaterin (25%) – performing routine work on the systematization of the material, analysis of the research results and data preparation, and writing the text of the article.

СПИСОК ЛИТЕРАТУРЫ

1. Сухарев О.С. Государственное управление импортозамещением: преодоление ограничений. *Управленец*. 2023;14(1):33–46. <https://doi.org/10.29141/2218-5003-2023-14-1-3>
2. Никитин Г.С., Скобелев Д.О. Эффективность государственных и корпоративных инвестиций в развитие реального сектора экономики. *Вестник Нижегородского университета им. Н.И. Лобачевского. Серия: Социальные науки*. 2022;4(68):32–41. URL: [http://www.unn.ru/pages/e-library/vestnik_soc/18115942_2022_-4\(68\)_unicode/4.pdf](http://www.unn.ru/pages/e-library/vestnik_soc/18115942_2022_-4(68)_unicode/4.pdf)
3. Улин Б. *Межрегиональная и международная торговля*: пер. с англ. М.: Дело; 2004. 416 с.
4. Портер М. *Международная конкуренция. Конкурентные преимущества стран*. М.: Интеллектуальная Литература; 2021. 948 с.
5. Авдокушин Е.Ф. *Международные экономические отношения*. М.: Юрист; 1999. С. 311–315. URL: <https://uchebnik.biz/book/111-mezhdunarodnye-yekonomicheskie-otnosheniya/> (дата обращения 04.04.2023).
6. Zheng X., Yu H., Yang L. Technology imports, independent innovation, and China's green economic efficiency: an analysis based on spatial and mediating effect. *Environ. Sci. Pollut. Res.* 2022;29(24):36170–36188. <https://doi.org/10.1007/s11356-021-17499-y>
7. Arezki R., Hadri K., Loungani P., Rao Y. Testing the Prebisch-Singer Hypothesis since 1650: Evidence from Panel Techniques that Allow for Multiple Breaks. *J. Int. Money & Finance*. 2014;42(180):208–223. <https://doi.org/10.1016/j.jimonfin.2013.08.012>
8. Нуреев Р.М. *Экономика развития: модели становления и модернизации рыночной экономики*. М.: Норма; 2008. 367 с.
9. Sachs J., Williamson J. External Debt and Macroeconomic Performance in Latin America and East Asia. *Brookings Papers on Economic Activity*. 1985;1985(2):523–573. <https://doi.org/10.2307/2534445>
10. Уханова Р.М., Райская М.В. Инновационное импортозамещение как основное направление технологической модернизации на предприятиях нефтегазового комплекса Республики Татарстан. *Вестник экономики, права и социологии*. 2015;4:134–136. URL: <http://www.vestnykeps.ru/0415/28.pdf>
11. Фальцман В. Импортозамещение в ТЭК и ОПК. *Вопросы экономики*. 2015;1:116–124. <https://doi.org/10.32609/0042-8736-2015-1-116-124>
12. Панышин И.В., Ярьес О.Б. Ресурсозамещение и импортозамещение при решении задач обеспечения экономической стабильности и стимулирования экономического роста региона. *Современные проблемы науки и образования*. 2014;6:511. URL: <https://science-education.ru/en/article/view?id=16092>

REFERENCES

1. Sukharev O.S. Import substitution policy: Breaking the limits. *Upravlenets = Manager*. 2023;14(1):33–46 (in Russ.). <https://doi.org/10.29141/2218-5003-2023-14-1-3>
2. Nikitin G.S., Skobelev D.O. Efficiency of state and corporate investments in the development of the real sector of the economy. *Vestnik Nizhegorodskogo universiteta im. N.I. Lobachevskogo. Seriya: Sotsial'nye nauki = Vestnik of Lobachevsky State University of Nizhni Novgorod. Series: Social Sciences*. 2022;4(68):32–41 (in Russ.). Available from URL: [http://www.unn.ru/pages/e-library/vestnik_soc/18115942_2022_-4\(68\)_unicode/4.pdf](http://www.unn.ru/pages/e-library/vestnik_soc/18115942_2022_-4(68)_unicode/4.pdf)
3. Ulin B. *Mezhregional'naya i mezhdunarodnaya trgovlya (Interregional and International Trade)*: transl. from Engl. Moscow: Delo; 2004. 416 p. (in Russ.). [Ohlin B. *Interregional and International Trade*. Cambridge (Massachusetts): Harvard University Press; 1967. 324 p.]
4. Porter M. *Mezhdunarodnaya konkurentsiya. Konkurentnye preimushchestva stran (International Competition. Competitive Advantages of Countries)*. Moscow: Intellektual'naya Literatura; 2021. 948 p. (in Russ.).
5. Avdokushin E.F. *Mezhdunarodnye ekonomicheskie otnosheniya (International Economic Relations)*. Moscow: Yurist; 1999. P. 311–315 (in Russ.). Available from URL: <https://uchebnik.biz/book/111-mezhdunarodnye-yekonomicheskie-otnosheniya/> (accessed April 04, 2023).
6. Zheng X., Yu H., Yang L. Technology imports, independent innovation, and China's green economic efficiency: an analysis based on spatial and mediating effect. *Environ. Sci. Pollut. Res.* 2022;29(24):36170–36188. <https://doi.org/10.1007/s11356-021-17499-y>
7. Arezki R., Hadri K., Loungani P., Rao Y. Testing the Prebisch-Singer Hypothesis since 1650: Evidence from Panel Techniques that Allow for Multiple Breaks. *J. Int. Money & Finance*. 2014;42(180):208–223. <https://doi.org/10.1016/j.jimonfin.2013.08.012>
8. Nureev R.M. *Ekonomika razvitiya: modeli stanovleniya i modernizatsii rynochnoi ekonomiki (Economics of Development: Models of Formation and Modernization of Market Economy)*. Moscow: Norma; 2008. 367 p. (in Russ.).
9. Sachs J., Williamson J. External Debt and Macroeconomic Performance in Latin America and East Asia. *Brookings Papers on Economic Activity*. 1985;1985(2):523–573. <https://doi.org/10.2307/2534445>
10. Ukhanova R.M., Raiskaya M.V. Innovative Import Substitution as the Main Direction of Technological Modernization in Oil and Gas Enterprises of the Republic of Tatarstan. *Vestnik ekonomiki, prava i sotsiologii = The Review of Economics, the Law and Sociology*. 2015;4:134–136 (in Russ.). Available from URL: <http://www.vestnykeps.ru/0415/28.pdf>

13. Волосатова А.А., Ученев А.А., Скобелев Д.О. Формирование концепции внедрения принципов зеленой экономики в евразийском экономическом союзе: роль гармонизации подходов к повышению ресурсной эффективности. *Вестник Евразийской науки*. 2022;14(4):14. URL: <https://esj.today/PDF/23ECVN422.pdf> (дата обращения 10.04.2023).
14. Сухарев О.С. *Экономическая теория эволюции институтов и технологий*. М.: Ленанд; 2019. 312 с.
15. Краснюк Л.В. Сущность и структурные элементы инновационного развития экономики. *π-Economy*. 2010;5(107):133–138. URL: <https://cyberleninka.ru/article/n/suschnost-i-strukturnye-elementy-innovatsionnogo-razvitiya-ekonomiki> (дата обращения 10.04.2023).
11. Faltsman V. Import Substitution in Energy and Military-Industrial Complexes. *Voprosy Ekonomiki*. 2015;1:116–124 (in Russ.). <https://doi.org/10.32609/0042-8736-2015-1-116-124>
12. Panshin I.V., Yares O.B. Resursozameschenie and import substitution in solving problems of economic stability and steam-lation region´s economic growth. *Sovremennye problemy nauki i obrazovaniya = Modern Problems of Science and Education*. 2014;6:511 (in Russ.). Available from URL: <https://science-education.ru/en/article/view?id=16092>
13. Volosatova A.A., Uchenov A.A., Skobelev D.O. Forming the concept of implementing green economy principles in the Eurasian Economic Union: the role of harmonizing resource efficiency approaches. *Vestnik Evraziiskoi nauki = The Eurasian Scientific J.* 2022;14(4):14 (in Russ.). Available from URL: <https://esj.today/PDF/23ECVN422.pdf> (accessed April 10, 2023).
14. Sukharev O.S. *Ekonomicheskaya teoriya evolyutsii institutov i tekhnologii (Economic Theory of the Evolution of Institutions and Technologies)*. Moscow: Lenand; 2019. 312 p. (in Russ.).
15. Krasnyuk L.V. Essence and the structural elements of the innovative development of economy. *π-Economy*. 2010;5(107):133–138 (in Russ.). Available from URL: <https://cyberleninka.ru/article/n/suschnost-i-strukturnye-elementy-innovatsionnogo-razvitiya-ekonomiki> (accessed April 10, 2023).

Об авторах

Михайлиди Дмитрий Христофорович, к.э.н., научный сотрудник отдела методологии ресурсосбережения, ФГАУ «Научно-исследовательский институт «Центр экологической и промышленной политики» (115054, Россия, Москва, Стремянный переулок, д. 38). E-mail: d.Mikhailidi@eipc.center. SPIN-код РИНЦ 6831-8043, <https://orcid.org/0009-0005-6491-0710>

Рагуткин Александр Викторович, к.т.н., проректор по вопросам инновационного развития, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: ragutkin@mirea.ru. Scopus Author ID 56871217700, ResearcherID AAE-4437-2022, SPIN-код РИНЦ 7531-7376, <https://orcid.org/0000-0001-8256-1941>

Скобелев Дмитрий Олегович, д.э.н., директор, ФГАУ «Научно-исследовательский институт «Центр экологической и промышленной политики» (115054, Россия, Москва, Стремянный переулок, д. 38); заведующий кафедрой наилучших доступных технологий и регуляторных практик Института технологий управления ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: skobelev@mirea.ru. Scopus Author ID 57205144505, SPIN-код РИНЦ 7830-9773, <https://orcid.org/0000-0002-8067-7016>

Сухатерин Алексей Борисович, преподаватель, кафедра промышленной информатики Института искусственного интеллекта, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: suhaterin@mirea.ru. <https://orcid.org/0009-0004-8997-4342>

About the authors

Dmitry Kh. Mikhailidi, Cand. Sci. (Econ.), Researcher, Department of Resource Saving Methodology, Environmental Industrial Policy Center (38, Stremyanniy per., Moscow, 115054 Russia). E-mail: d.Mikhailidi@eipc.center. RSCI SPIN-code 6831-8043, <https://orcid.org/0009-0005-6491-0710>

Alexander V. Ragutkin, Cand. Sci. (Eng.), Vice-Rector for Innovative Development, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: ragutkin@mirea.ru. Scopus Author ID 56871217700, ResearcherID AAE-4437-2022, RSCI SPIN-code 7531-7376, <https://orcid.org/0000-0001-8256-1941>

Dmitry O. Skobelev, Dr. Sci. (Econ.), Director, Environmental Industrial Policy Center (38, Stremyanniy per., Moscow, 115054 Russia); Head of the Department of Best Available Technologies and Regulatory Practices, Institute of Management Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: skobelev@mirea.ru. Scopus Author ID 57205144505, RSCI SPIN-code 7830-9773, <https://orcid.org/0000-0002-8067-7016>

Alexey B. Suhaterin, Lecturer, Department of Industrial Informatics, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: suhaterin@mirea.ru. <https://orcid.org/0009-0004-8997-4342>

Отпечатано в «МИРЭА – Российский технологический университет».

119454, РФ, г. Москва, пр-т Вернадского, д. 78.

Подписано в печать 31.07.2023 г.

Формат 60 × 90/8. Печать цифровая.

Уч.-изд. листов 14.5.

Тираж 100 экз. Заказ № 424.

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 July 31, 2023.

Format 60 × 90/8. Digital print.

C.p.l. 14.5.

100 copies. Order No. 424.

Subscription to the *Russian Technological
Journal* printed version can be made through
the *Kniga-Servis* Agency, www.akc.ru.

Subscription index: **79641**.

