

Учредитель: Федеральное государственное автономное
образовательное учреждение высшего образования
«Пермский государственный национальный исследовательский университет»

Научный журнал "Вестник Пермского университета. Математика. Механика. Информатика" публикует оригинальные научные статьи, краткие сообщения и обзорные статьи по направлениям журнала на русском и английском языках, ранее не публиковавшиеся и не представленные для публикации в других изданиях. Все статьи проходят рецензирование членами редакционной коллегии журнала и внешними рецензентами.

РЕДАКЦИОННЫЙ СОВЕТ

БУЗМАКОВА М. М., к. ф.-м. н., доцент; Пермский государственный национальный исследовательский университет, Пермь, Россия;

МАЛАНИН В. В., д. т. н., профессор; Пермский государственный национальный исследовательский университет, Пермь, Россия;

БАРУЛИНА М. А., д. ф.-м. н.; Пермский государственный национальный исследовательский университет, Пермь, Россия;

НУРБАКОВА Д. М., PhD in Computer Science; Национальный институт прикладных наук Лиона (National Institute of Applied Sciences of Lyon), Лион, Франция.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

БУЗМАКОВА М. М., к.ф.-м.н., доцент, **главный редактор**; Пермский государственный национальный исследовательский университет, Пермь, Россия;

РУСАКОВА О. Л., к.ф.-м.н., доцент, **ответственный секретарь**; Пермский государственный национальный исследовательский университет, Пермь, Россия;

АЛЕКСЕЕВ А. О., к.э.н., доцент; Пермский национальный исследовательский политехнический университет, Пермь, Россия;

АПТУКОВ В. Н., д.т.н., профессор; Пермский государственный национальный исследовательский университет, Пермь, Россия;

АПУШКИНСКАЯ Д. Е., д.ф.-м.н.; Российский университет дружбы народов имени Патриса Лумумбы, Москва, Россия;

БАРУЛИНА М. А., д.ф.-м.н.; Пермский государственный национальный исследовательский университет, Пермь, Россия;

БИРЮКОВ А. Н., д.э.н., доцент; Стерлитамакский филиал Башкирского государственного университета, Стерлитамак, Россия;

БОРИСОВ В. В., д.т.н., профессор; Смоленский филиал национального исследовательского университета «МЭИ», Смоленск, Россия;

ВЕЧТОМОВ Е. М., д.ф.-м.н., профессор; Вятский государственный университет, Киров, Россия;

ГУСЕВ А. Л., д.т.н.; Пермский государственный национальный исследовательский университет, Пермь, Россия;

ДЕМИДОВ С. С., д.ф.-м.н.; Московский государственный университет имени М. В. Ломоносова, Москва, Россия;

ИВАНОВ Д. В., д.ф.-м.н., доцент; Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, Россия;

ИСАЕВ К. П., д.ф.-м.н.; Уфимский университет науки и технологий, Уфа, Россия;

ИШКИН Х. К., д.ф.-м.н.; доцент; Уфимский университет науки и технологий, Уфа, Россия;

КРИВОШЕЕВА О. А., д. ф.-м. н., доцент; Уфимский университет науки и технологий, Уфа, Россия;

КОССОВИЧ Л. Ю., д.ф.-м.н., профессор; Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, Россия;

КУЗНЕЦОВ А. Г., к.т.н.; Пермский государственный гуманитарно-педагогический университет, Пермь, Россия;

МАЛАНИН В. В., д.т.н., профессор; Пермский государственный национальный исследовательский университет, Пермь, Россия;

МАТВЕЕНКО В. П., д.т.н., профессор, академик РАН; Институт механики сплошных сред УрО РАН, Пермь, Россия;

МАХНЕВ А. А., д.ф.-м.н., профессор, чл.-корр. РАН; Институт математики и механики им. Н.Н. Красовского УрО РАН, Екатеринбург, Россия;

ПАЛКОВА И. В., д.ф.-м.н., доцент; Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, Россия;

ПЕСТРЕНИН В. М., к.ф.-м.н., доцент; Пермь, Россия;

ПОЛОСКОВ И. Е., д.ф.-м.н., доцент; Пермь, Россия;

РУСАКОВ С. В., д.ф.-м.н., профессор; Пермский государственный национальный исследовательский университет, Пермь, Россия;

САВЧИН В. М., д.ф.-м.н., профессор; Российский университет дружбы народов имени Патриса Лумумбы, Москва, Россия;

СЕЛЕТКОВ И. П., к.т.н.; Пермский государственный национальный исследовательский университет, Пермь, Россия;

ТРУСОВ П. В., д.ф.-м.н., профессор; Пермский национальный исследовательский политехнический университет, Пермь, Россия;

ХАЧАЙ М. Ю., д.ф.-м.н., профессор РАН, чл.-корр. РАН; Институт математики и механики им. Н.Н. Красовского УрО РАН, Екатеринбург, Россия;

ХЕННЕР Е. К., д.ф.-м.н., профессор, чл.-корр. РАО; Пермский государственный национальный исследовательский университет, Пермь, Россия;

ФАЗУЛЛИН З.Ю., д.ф.-м.н., доцент; Уфимский университет науки и технологий, Уфа, Россия;

ЧЕРНЫХ В. В., д.ф.-м.н., доцент; Сыктывкарский государственный университет имени Питирима Сорочкина, Сыктывкар, Россия;

ЧЕРНИКОВ А. В., к.т.н.; Пермский государственный национальный исследовательский университет, Пермь, Россия;

ЧИНЕНОВА В. Н., к.ф.-м.н.; Московский государственный университет имени М. В. Ломоносова, Москва, Россия;

ШАХМАМЕТОВА Г. Р., д.т.н., доцент; Уфимский университет науки и технологий, Уфа, Россия;

ШВАРЦ К. Г., д.ф.-м.н., профессор; Пермский государственный национальный исследовательский университет, Пермь, Россия;

ЮМАГУЛОВ М. Г., д.ф.-м.н., профессор; Уфимский университет науки и технологий, Уфа, Россия;

ЯСНИЦКИЙ Л. Н., д.т.н., профессор; Пермский филиал Национального исследовательского университета "Высшая школа экономики", Пермь, Россия.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации средства массовой информации ПИ No ФС 77 -66787 от 08.08.2016.

Журнал входит в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (приказ от 08.07.2024). Научные специальности: Вещественный, комплексный и функциональный анализ; Дифференциальные уравнения и математическая физика; Геометрия и топология; Математическая логика, алгебра, теория чисел и дискретная математика; Вычислительная математика; Теоретическая механика, динамика машин; Механика деформируемого твердого тела; Механика жидкости, газа и плазмы; Искусственный интеллект и машинное обучение; Математическое моделирование, численные методы и комплексы программ; Теоретическая информатика, кибернетика; Системный анализ управление и обработка информации, статистика; Вычислительные системы и их элементы; Компьютерное моделирование и автоматизация проектирования. Издание включено в национальную информационно-аналитическую систему «Российский индекс научного цитирования» (РИНЦ, eLIBRARY ID 28484).

Founder: Perm State University

The journal "Bulletin of Perm University. Mathematics. Mechanics. Computer Science" publishes Full Articles, Short Communications Article and Review paper/perspectives on the journal's directions, which have not been previously published. The papers have been written in Russian and English. The papers are reviewed by the journal editorial board members and external reviewers.

EDITORIAL COUNCIL

BUZMAKOVA M. M., Candidate of Science in Physics and Mathematics, Associate Professor; Perm State University, Perm, Russia;
MALANIN V. V., Doctor of Science in Engineering, Professor; Perm State University, Perm, Russia;
BARULINA M. A., Doctor of Science in Physics and Mathematics; Perm State University, Perm, Russia;
NURBAKOVA D. M., PhD in Computer Science; National Institute of Applied Sciences of Lyon, France.

EDITORIAL BOARD

BUZMAKOVA M. M., Candidate of Science in Physics and Mathematics, Associate Professor; **Editor-in-Chief**; Perm State University, Perm, Russia;
RUSAKOVA O. L., Candidate of Science in Physics and Mathematics, Associate Professor; **Executive Secretary**; Perm State University, Perm, Russia;
ALEKSEEV A. O., Candidate of Science in Economics, Associate Professor; Perm National Research Polytechnic University, Perm, Russia;
APTUKOV V. N., Doctor of Science in Engineering, Professor; Perm State University, Perm, Russia;
APUSHKINSKAYA D. E., Doctor of Science in Physics and Mathematics; Peoples' Friendship University of Russia named after Patrice Lumumba, Moscow, Russia;
BARULINA M. A., Doctor of Science in Physics and Mathematics; Perm State University, Perm, Russia;
BIRUKOV A. N., Doctor of Science in Economics, Associate Professor; Sterlitamak Branch of Bashkir State University, Sterlitamak, Russia;
BORISOV V. V., Doctor of Science in Engineering, Professor; Smolensk Branch of National Research University "MEI", Smolensk, Russia;
VECHTOMOV E. M., Doctor of Science in Physics and Mathematics, Professor; Vyatka State University, Kirov, Russia;
GUSEV A. L., Doctor of Science in Engineering; Perm State University, Perm, Russia;
DEMIDOV S. S., Doctor of Science in Physics and Mathematics; Lomonosov Moscow State University, Moscow, Russia;
IVANOV D. V., Doctor of Science in Physics and Mathematics, Associate Professor; N.G. Chernyshevsky Saratov National Research State University, Saratov, Russia;
ISAEV K. P., Doctor of Science in Physics and Mathematics; Ufa University of Science and Technology, Ufa, Russia;
ISHKIN H. K., Doctor of Science in Physics and Mathematics, Associate Professor; Ufa University of Science and Technology, Ufa, Russia;
KRIVOSHEEVA O. A., Doctor of Science in Physics and Mathematics, Associate Professor; Ufa University of Science and Technology, Ufa, Russia;
KOSSOVICH L. Yu., Doctor of Science in Physics and Mathematics, Professor; N.G. Chernyshevsky Saratov National Research State University, Saratov, Russia;
KUZNECOV A. G., Candidate of Science in Engineering; Perm State Humanitarian and Pedagogical University, Perm, Russia;
MALANIN V. V., Doctor of Science in Engineering, Professor; Perm State University, Perm, Russia;
MATVEENKO V. P., Doctor of Science in Engineering, Professor, Academician of RAS; Institute of Continuous Media Mechanics, UrD RAS, Perm, Russia;

MAHNEV A. A., Doctor of Science in Physics and Mathematics, Professor, member-corr. RAS; N.N. Krasovsky Institute of Mathematics and Mechanics, UrD RAS;
PAPKOVA I. V., Doctor of Science in Physics and Mathematics, Associate Professor; N.G. Chernyshevsky Saratov National Research State University, Saratov, Russia;
PESTRENIN V. M., Candidate of Science in Physics and Mathematics, Associate Professor; Perm, Russia;
POLOSKOV I. E., Doctor of Science in Physics and Mathematics, Associate Professor; Perm, Russia;
RUSAKOV S. V., Doctor of Science in Physics and Mathematics, Professor; Perm State University, Perm, Russia;
SAVCHIN V. M., Doctor of Science in Physics and Mathematics, Professor; Peoples' Friendship University of Russia named after Patrice Lumumba, Moscow, Russia;
SELETKOV I. P., Candidate of Science in Engineering; Perm State University, Perm, Russia;
TRUSOV P. V., Doctor of Science in Physics and Mathematics, Professor; Perm National Research Polytechnic University, Perm, Russia;
HACHAI M. Yu., Doctor of Science in Physics and Mathematics, Professor of RAS, member-corr. RAS; N.N. Krasovsky Institute of Mathematics and Mechanics, UrD RAS, Yekaterinburg, Russia;
HENNER E. K., Doctor of Science in Physics and Mathematics, Professor, member-corr. RAO; Perm State University, Perm, Russia;
FAZULLIN Z. Yu., Doctor of Science in Physics and Mathematics, Associate Professor; Ufa University of Science and Technology, Ufa, Russia;
CHERMNYH V. V., Doctor of Science in Physics and Mathematics, Associate Professor; Pitirim Sorokin Syktyvkar State University, Syktyvkar, Russia;
CHERNIKOV A. V., Candidate of Science in Engineering; Perm State University, Perm, Russia;
CHINENOVA V. N., Candidate of Science in Physics and Mathematics; Lomonosov Moscow State University, Moscow, Russia;
SHAHMAMETOVA G. R., Doctor of Science in Engineering, Associate Professor; Ufa University of Science and Technology, Ufa, Russia;
SHVARTS K. G., Doctor of Science in Physics and Mathematics, Professor; Perm State University, Perm, Russia;
UMAGULOV M. G., Doctor of Science in Physics and Mathematics, Professor; Ufa University of Science and Technology, Ufa, Russia;
YASNICKIY L. N., Doctor of Science in Engineering, Professor; Perm Branch of the National Research University Higher School of Economics, Perm, Russia.

The journal is registered by the Federal Service for Supervision of Communications, Information Technology and Mass Communications (Roskomnadzor). Certificate of registration of the mass media PI No. FS 77 -66787 dated 08.08.2016. The journal is included in the peer-reviewed scientific publications list for the main research results should be published (Candidate of Sciences and Doctor of Sciences, order dated 07/08/2024). Scientific specialties: Real, complex and functional analysis; Differential equations and mathematical physics; Geometry and topology; Mathematical logic, algebra, number theory and discrete mathematics; Computational Mathematics; Theoretical mechanics, dynamics of machines; Mechanics of deformable solids; Mechanics of liquid, gas and plasma; Artificial intelligence and machine learning; Mathematical modeling, numerical methods and software packages; Theoretical Computer science, cybernetics; System analysis, information management and processing, statistics; Computer systems and their elements; Computer modeling and design automation. The journal is included in the national information and analytical system "Russian Science Citation Index" (eLIBRARY ID 28484).

Содержание

Аналитический отчет о проведении II Всероссийской научно-практической конференции с международным участием «Перспективы и возможности использования цифровых технологий в науке, образовании и управлении».....	5
--	---

МАТЕМАТИКА

Кунгиров М. Н. О бифуркации циклов на бесконечности в системах с однородными нелинейностями	6
Пенский О. Г. Метеоролог и математик. К 70-летию со дня рождения профессора К. Г. Шварца	15

МЕХАНИКА

Власов А. Н., Ощепков А. Ю. Моделирование робастных алгоритмов управления подъемом и горизонтальной стабилизацией беспилотного трикоптера.....	21
Морозов И. А., Беляев А. Ю. Исследование микроструктуры и локальных механических свойств растянутых полиуретанов.....	36

КОМПЬЮТЕРНЫЕ НАУКИ И ИНФОРМАТИКА

Djeguede A. M. A. E. An Algorithm for the Initial Detection of Malicious Traffic Based on the Autoencoder Reconstruction Error and a Variational Model: the Influence of the Error Distribution Density on the Performance Indicators of the Models.....	47
Бочкарев В. А., Усынин А. А., Осипов А. Д., Аухадиев М. Р., Шаров Р. В., Барулина М. А. Модели сегментации сердечных артерий по коронарографическим снимкам.....	65
Караулов В. Г., Грузликов А. М. Решение задачи навигации АНПА на ОС реального времени.....	88
Радченко Е. Н., Исаева Е. В. Анализ подходов к автоматизации разметки паралингвистических характеристик в русскоязычных речевых данных	101
Самохина М. А., Самохин А. С. Решение задачи расстановки обнаружителей для охраны периметра градиентным методом.....	123
Правила подачи и оформления рукописей	137

C o n t e n t s

Analytical Report on the II All-Russian Scientific and Practical Conference with international participation «Prospects and possibilities of using digital technologies in science, education and management»	5
--	---

MATHEMATICS

Kungirov M. N. y. On Bifurcation of Cycles at Infinity in Systems with Homogeneous Nonlinearities	6
Pensky O. G. Meteorologist and Mathematician. On the 70 Anniversary of the Birth of Professor K.G. Schwartz.....	15

MECHANICS

Vlasov A. N., Oshchepkov A. Yu. Modeling of Robust Algorithms for Control of Lift and Horizontal Stabilization of an Unmanned Tricopter	21
Morozov I. A., Beliaev A. Yu. Study of microstructure and local mechanical properties of stretched polyurethanes	36

COMPUTER SCIENCE

Djeguede A. M. A. E. An Algorithm for the Initial Detection of Malicious Traffic Based on the Autoencoder Reconstruction Error and a Variational Model: the Influence of the Error Distribution Density on the Performance Indicators of the Models	47
Bochkarev V. A., Usynin A. A., Osipov A. D., Auhadiev M. R., Sharov R. V. and Barulina M. A. Models of Cardiac Artery Segmentation Based on Coronary Angiographic Images	65
Karaulov V. G., Gruzlikov A. M. Solving the AUV Navigation Problem on a Real-Time Operating System.....	88
Radchenko E. N., Isaeva E. V. Analysis of Approaches to Paralinguistic Feature Annotation Automation in Russian Speech.....	101
Samokhina M. A., Samokhin A. S. Solution of the Sensors Placement Problem for Perimeter Defense Using the Gradient Method.....	123
Rules for submission and structure of manuscripts	137

**Аналитический отчет о проведении
II Всероссийской научно-практической конференции
с международным участием
«Перспективы и возможности использования цифровых технологий в науке,
образовании и управлении»**

Конференция состоялась в Астраханском государственном университете имени В.Н. Татищева 20–21 мая 2025 г. в очно-заочном формате. Мероприятие объединило более 50 участников, включая представителей из 15 вузов и организаций России и зарубежных стран, таких как Замбия, Танзания, Таджикистан, Казахстан, Китай и другие. Конференция продемонстрировала междисциплинарный подход к обсуждению актуальных вопросов цифровизации.



Широкий международный состав участников подчеркнул глобальный характер цифровых технологий и науки. Среди участников были представители академических кругов, бизнеса (например, компания Astraway) и государственных структур (Министерство Астраханской области). Доклады охватили широкий спектр тем, от теоретических исследований до практических кейсов, что свидетельствует о высокой актуальности выбранной тематики.

Основные темы и доклады:
робототехника в науке и образовании (Султанова Б.К., Казахстан),
модели распространения эпидемий (Тарасевич Ю.Ю., АГУ),
достоверность как сервис информационной безопасности (Ажмухамедов И.М., АГУ),
цифровизация туристической индустрии (Колесников В.Д. и Жудиниллсон А.М., Astraway).



Основными направлениями конференции стали информационная безопасность; цифровые технологии в науке и образовании; программное обеспечение и цифровая трансформация. По итогам конференции готовится к выпуску сборник материалов (РИНЦ), часть докладов отобраны для публикации в научном журнале «Вестник Пермского университета. Математика. Механика. Информатика». Конференция успешно выполнила свои цели, предоставив платформу для обмена знаниями и опытом в области цифровых технологий. Участники отметили высокий уровень организации и актуальность обсуждаемых вопросов.

Проведение подобных мероприятий способствует интеграции науки, образования и бизнеса, что особенно важно в эпоху глобальной цифровизации.

МАТЕМАТИКА

Научная статья

УДК 517.938

DOI: 10.17072/1993-0550-2025-2-6-14

<https://elibrary.ru/rgkmb>



О бифуркации циклов на бесконечности в системах с однородными нелинейностями

Мамирбой Норбек угли Кунгиров

Уфимский университет науки и технологий, г. Уфа, Россия

mamur.qongirov@mail.ru

Аннотация. В настоящей статье изучается задача о бифуркации циклов на бесконечности в динамических системах с однородной нелинейностью четного порядка. Предлагаются достаточные признаки такой бифуркации, определяемые как главной линейной частью, так и характеристиками нелинейностей. Получены асимптотические формулы, позволяющие описать эволюцию возникающих циклов при изменении параметров системы. В качестве приложения рассматривается задача о бифуркации циклов на бесконечности в модифицированной модели Ресслера.

Ключевые слова: бифуркация циклов на бесконечности; устойчивость; бифуркация Андронова–Хопфа; асимптотические формулы

Для цитирования: Кунгиров М. Н. О бифуркации циклов на бесконечности в системах с однородными нелинейностями // Вестник Пермского университета. Математика. Механика. Информатика. 2025. Вып. 2(69). С. 6–14. DOI: 10.17072/1993-0550-2025-2-6-14. <https://elibrary.ru/rgkmb>

Благодарности: работа выполнена в рамках государственного задания, соглашение № 075-03-2024-123/1 от 15.02.2024, проект № 324-21.

Статья поступила в редакцию 30.11.2024; одобрена после рецензирования 19.05.2025; принята к публикации 11.07.2025.

MATHEMATICS

Research article

On Bifurcation of Cycles at Infinity in Systems with Homogeneous Nonlinearities

Mamirboy. N. Kungirov

Ufa University of Science and Technology, Ufa, Russia

mamur.qongirov@mail.ru



Эта работа © 2025 Кунгиров М. Н. распространяется по лицензии CC BY 4.0. Чтобы ознакомиться с условиями этой лицензии, посетите сайт <https://creativecommons.org/licenses/by/4.0/>.

Abstract. In this article, the problem of bifurcation of cycles at infinity in dynamical systems with homogeneous nonlinearities of even order is studied. Sufficient conditions for such bifurcation are proposed, determined both by the principal linear part and by the characteristics of the nonlinearities. Asymptotic formulas are obtained, which allow describing the evolution of the emerging cycles as the system parameters change. As an application, the problem of bifurcation of cycles at infinity in a modified Rössler model is considered.

Keywords: *bifurcation of cycles at infinity; stability; Andronov–Hopf bifurcation; asymptotic formulas*

For citation: Kungirov, M. N. y. (2025), "On Bifurcation of Cycles at Infinity in Systems with Homogeneous Nonlinearities", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 6-14. (In Russ.). DOI: 10.17072/1993-0550-2025-2-6-14. <https://elibrary.ru/rgkmbq>

Acknowledgments: this work was supported by the state research program, agreement №. 075-03-2024-123/1 dated February 15, 2024, Project №. 324-21.

The article was submitted 30.11.2024; approved after reviewing 19.05.2025; accepted for publication 11.07.2025

1. Введение и постановка задачи

Рассматривается зависящее от вещественного параметра α нелинейное дифференциальное уравнение

$$\frac{dx}{dt} = f(x, \alpha), \quad x \in R^N, \quad N \geq 2, \quad (1)$$

в котором функция $f(x, \alpha)$ является гладкой по x и α . Пусть $f(x_0, \alpha_0) = 0$, т. е. уравнение (1) при $\alpha = \alpha_0$ имеет точку равновесия $x = x_0$, а матрица Якоби $A(\alpha) = f'_x(x_0, \alpha)$ при $\alpha = \alpha_0$ имеет простые собственные значения $\pm i\omega_0$, $\omega_0 > 0$. Остальные собственные значения матрицы $A(\alpha_0)$ имеют ненулевые вещественные части. Тогда α_0 является точкой бифуркации Андронова–Хопфа системы (1) в окрестности точки равновесия x_0 . Другими словами, система (1) при α , близких к α_0 , имеет нестационарные периодические решения $x = x(t, \alpha)$ периода $T = T(\alpha)$ такое, что:

1. Функции $x(t, \alpha)$ и $T(\alpha)$ непрерывно дифференцируемы по α ;
2. $T(\alpha) \rightarrow T_0$ при $\alpha \rightarrow \alpha_0$, $\max_t \|x(t, \alpha) - x_0\| \rightarrow 0$ при $\alpha \rightarrow \alpha_0$.

Здесь $T_0 = 2\pi/\omega_0$. $\|\cdot\|$ – евклидова норма в пространстве R^N .

Вместе с тем здесь возможны и другие бифуркации, в частности, бифуркация циклов на бесконечности. Говорят, что $\alpha = \alpha_0$ является точкой бифуркации Андронова–Хопфа системы (1) на бесконечности, если система (1) при α , близком к α_0 , имеет нестационарное периодическое решение $x = x(t, \alpha)$ периода $T = T(\alpha)$ такое, что:

1. Функции $x(t, \alpha)$ и $T(\alpha)$ непрерывно дифференцируемы по α ;
2. $T(\alpha) \rightarrow T_0$ при $\alpha \rightarrow \alpha_0$, $\max_t \|x(t, \alpha)\| \rightarrow \infty$ при $\alpha \rightarrow \alpha_0$.

Замечание 1. Понятие бифуркации Андронова–Хопфа на бесконечности аналогично понятию классической бифуркации Андронова–Хопфа. В первом случае периодические решения системы возникают в окрестности бесконечности, во втором – в окрестности точки равновесия. Вместе с тем эти виды бифуркаций не могут быть сведены друг к другу простой заменой координат типа $y = x/\|x\|^2$. Отметим также, что в обоих видах бифуркаций период T_n колебаний стремится к числу $T_0 = 2\pi/\omega_0$, т.е. к периоду свобод-

ных колебаний линеаризованной системы $x' = A(\alpha_0)x$. Это означает, что скорость движения по циклу в окрестности бесконечности сколь угодно большая, а в окрестности точки равновесия – сколь угодно мала.

Понятие бифуркации Андронова-Хопфа на бесконечности было введено в [1] для случая общих уравнений в R^N , там же сформулированы условия возникновения бифуркации на бесконечности, определяемые главной линейной частью. Теория бифуркации на бесконечности разрабатывалась в работах многих авторов. В [2] вместо непрерывности функциональной нелинейности по фазовым переменным предполагается лишь ее кусочная непрерывность. В [3] и многих других работах С. Рыбицкого (S. Rybicki) задачи о бифуркациях Андронова-Хопфа на бесконечности исследуются в предположении гамильтоновости изучаемых систем. В [4] существенно используется двумерность изучаемых систем. В [5] изучается вырожденный случай бифуркации Андронова-Хопфа на бесконечности.

Важное место в современной нелинейной динамике занимает исследование систем, содержащих однородные нелинейности, в частности, изучение в таких системах сценариев бифуркаций и перехода к хаосу. Укажем некоторые работы в этом направлении, прилегающие к постановкам настоящей статьи. Ряд моделей со степенными нелинейностями приведен в работе [11], в которой основное внимание уделяется вопросам возникновения хаотических режимов. В [8] приводится ряд формул для исследования задач о бифуркации в системах с однородными нелинейностями. В [12] изучаются вопросы о существовании предельных циклов в трехмерных системах с квадратичными нелинейностями.

В настоящей работе исследуется задача о бифуркации Андронова-Хопфа на бесконечности для систем с однородными нелинейностями. Рассматривается система вида:

$$\frac{dy}{dt} = A_0 y + \alpha f(y, \alpha), \quad y \in R^N, \quad N \geq 2, \quad (2)$$

в которой A_0 – квадратная матрица порядка N , функция $f(y, \alpha)$ имеет вид

$$f(y, \alpha) = A_1 y + a_{2q}(y, \alpha),$$

где A_1 – квадратная матрица порядка N , $a_{2q}(y, \alpha)$ является однородным по y вектор-полиномом четного порядка $2q$; здесь q – натуральное число. Предполагается, что матрица A_0 имеет простые собственные значения $\pm i\omega_0$, $\omega_0 > 0$, и не имеет других чисто мнимых собственных значений. Основной в настоящей работе является задача получения достаточных условий бифуркации циклов на бесконечности в системе (2), а также построение асимптотических формул для исследования такой бифуркации.

Настоящая работа продолжает исследования, начатые в [6] и [7]. В свою очередь, указанные статьи базируются на работе [8], в которой предложены новые формулы в задаче о бифуркации Андронова-Хопфа. Эти формулы позволяют провести детальное исследование ряда приложений, в частности, задач о бифуркации в системах с однородными нелинейностями, задач о вынужденной синхронизации автоколебаний под действием внешнего вынуждающего периодического воздействия [9].

2. Основные утверждения

Так как матрица A_0 имеет простые собственные значения $\pm i\omega_0$, $\omega_0 > 0$, то найдутся ненулевые векторы $e, g, \tilde{e}, \tilde{g} \in R^N$ такие, что выполняются равенства

$$A_0(e + ig) = i\omega_0(e + ig), \quad A_0^T(\tilde{e} + i\tilde{g}) = -i\omega_0(\tilde{e} + i\tilde{g}).$$

Здесь A_0^T – транспонированная матрица. Известно (см. [8]), что векторы $e, g, \tilde{e}, \tilde{g} \in R^N$ можно нормировать в соответствии с равенствами

$$(e, \tilde{e}) = (g, \tilde{g}) = 1, (e, \tilde{g}) = (g, \tilde{e}) = 0.$$

Положим

$$\gamma_1 = (A_1 e, \tilde{e}) + (A_1 g, \tilde{g}), \quad \gamma_2 = (A_1 e, \tilde{g}) + (A_1 g, \tilde{e}). \quad (3)$$

Предполагается, что $\gamma_1 \neq 0$.

Следуя [8], определим числа

$$\alpha_2 = -\frac{\omega_0}{\pi\gamma_1} (b_3, \tilde{e}), \quad T_2 = \frac{1}{\omega_0} \left[(b_3, \tilde{g}) + \frac{\gamma_2}{\gamma_1} (b_3, \tilde{e}) \right], \quad (4)$$

где

$$b_3 = T_0 \int_0^1 e^{(1-t)T_0 A_0} f_3(t) dt, \quad f_3(t) = F_2(t) \int_0^t e^{(1-s)T_0 A_0} a_{2q}(e(s), 0) ds, \\ F_2(t) = T_0 a'_{2qx}(e(t), 0) e^{T_0 A_0 t}.$$

Здесь $a'_{2qx}(e(t), 0)$ – матрица Якоби вектор-функции $a_{2q}(x, 0)$, $e(t) = e \cdot \cos(2\pi t) - g \cdot \sin(2\pi t)$. В формулах (3) и (4) используются обозначения, соответствующие их аналогам из статьи [8].

Теорема 1. Пусть матрица A_0 имеет простые собственные значения $\pm i\omega_0$, $\omega_0 > 0$, и не имеет других чисто мнимых собственных значений. Пусть $\gamma_1 \neq 0$, $\alpha_2 \neq 0$. Тогда значение $\alpha = 0$ является точкой бифуркации циклов системы (2) на бесконечности.

Таким образом, в условиях теоремы 1 существуют предельные значения $\alpha \rightarrow \alpha_0$ и $T(\alpha) \rightarrow T_0$, при которых для α , близких к α_0 , система (2) обладает нестационарным периодическим решением $x = x(t, \alpha)$ с периодом $T = T(\alpha)$, удовлетворяющим условию $\max_t \|x(t, \alpha)\| \rightarrow \infty$ при $\alpha \rightarrow \alpha_0$.

Бифурцирующие решения системы (2) и значения их периода часто более удобно изучать не в виде прямой зависимости от параметра α , т.е. в виде $y = y(t, \alpha)$ и $T = T(\alpha)$, которая может оказаться многозначной, а в параметрической форме:

$$y = y(t, \varepsilon), \quad \alpha = \alpha(\varepsilon), \quad T = T(\varepsilon), \quad (5)$$

где ε – вспомогательный малый параметр так, что $\alpha(0) = 0$ и $T(0) = T_0$.

Теорема 2. В условиях теоремы 1 существуют число $\varepsilon_0 > 0$ и определенные при $0 \leq \varepsilon < \varepsilon_0$ функции

$$\alpha(\varepsilon) = \alpha_2 \varepsilon^{2q} + O(\varepsilon^{2q+1}), \quad T(\varepsilon) = T_0 + T_2 \varepsilon^{2q} + O(\varepsilon^{2q+1}), \quad (6)$$

такие, что система (2) при $\alpha = \alpha(\varepsilon)$ имеет периодическое решение $y = y(t, \varepsilon)$ периода $T(\varepsilon)$, при этом

$$y(0, \varepsilon) = \alpha_2^{\frac{1}{1-2q}} \cdot \varepsilon^{\frac{1}{1-2q}} e + o\left(\varepsilon^{\frac{1}{1-2q}}\right). \quad (7)$$

Из теоремы 2 следует, что циклы больших амплитуд в системе (2) возникают при $\alpha > 0$ ($\alpha < 0$), если $\alpha_2 > 0$ ($\alpha_2 < 0$).

Теорема 3. Пусть в условиях теоремы 1 все отличные от $\pm\omega_0 i$ собственные значения матрицы A_0 имеют отрицательные вещественные части. Тогда, если $(b_3, \tilde{e}) < 0$ ($(b_3, \tilde{e}) > 0$), то циклы $y(t, \varepsilon)$ больших амплитуд системы (2) асимптотически орбитально устойчивы (неустойчивы).

Доказательство теорем 1 и 2

Приведем сначала вспомогательные утверждения. С этой целью рассмотрим систему

$$x' = (A_0 + \alpha A_1)x + a_{2q}(x, \alpha), \quad x \in R^N, N \geq 2. \quad (8)$$

Лемма 1. Пусть матрица A_0 имеет простые собственные значения $\pm \omega_0 i$, $\omega_0 > 0$ и не имеет других чисто мнимых собственных значений. Пусть $\gamma_1 \neq 0$. Тогда $\alpha = 0$ является точкой бифуркации Андронова–Хопфа системы (8) в окрестности точки равновесия $x = 0$. При этом существуют число $\varepsilon_0 > 0$ и определенные при $0 \leq \varepsilon < \varepsilon_0$ функции (6) такие, что система (8) при $\alpha = \alpha(\varepsilon)$ имеет периодическое решение $x = x(t, \varepsilon)$ периода $T(\varepsilon)$, для которого вектор $\tilde{x}_0(\varepsilon) = x(0, \varepsilon)$ представим в виде $\tilde{x}_0(\varepsilon) = \varepsilon e + O(\varepsilon^{2q})$.

Лемму 1 можно рассматривать как развитие аналогичного утверждения, доказанного в [8, теорема 2].

Доказательство леммы -

Сначала отметим, что функция $x = x(t)$ тогда и только тогда будет T – периодическим решением системы (8), когда функция $y(t) = x(tT)$ будет 1 – периодическим решением системы

$$y' = T(A_0 + \alpha A_1)y + Ta_{2q}(y, \alpha), \quad y \in R^N. \quad (9)$$

Далее функция $y = y(t)$ тогда и только тогда будет 1 – периодическим решением системы (9), когда вектор $u = y(0) \in R^N$ будет решением операторного уравнения

$$u = B(T, \alpha)u + b(u, T, \alpha), \quad u \in R^N, \quad (10)$$

где

$$B(T, \alpha) = e^{TA(\alpha)}, \quad b(u, T, \alpha) = T \int_0^1 e^{(1-s)TA(\alpha)} a_{2q}(y(s), \alpha) ds;$$

здесь $A(\alpha) = A_0 + \alpha A_1$, $y(t)$ – решение задачи Коши для уравнения (9) при начальном условии $y(0) = u$.

Вектор-функция $b(u, T, \alpha)$ представима в виде

$$b(u, T, \alpha) = b_{2q}(u, T, \alpha) + b_{3q}(u, T, \alpha) + b_{4q}(u, T, \alpha),$$

где

$$b_{2q}(u, T, \alpha) = T \int_0^1 e^{(1-s)TA(\alpha)} a_{2q}(e^{TA(\alpha)s} u, \alpha) ds,$$

$$b_{3q}(u, T, \alpha) = T^2 \int_0^1 e^{(1-s)TA(\alpha)} a'_{2qx}(e^{TA(\alpha)s} u, \alpha) \left(\int_0^s e^{(s-\tau)TA(\alpha)} a_{2q}(e^{TA(\alpha)\tau} u, \alpha) d\tau \right) ds,$$

а нелинейность $b_{4q}(u, T, \alpha)$ удовлетворяет соотношению $b_{4q}(u, T, \alpha) = O(\|u\|^{4q})$ при $u \rightarrow 0$ равномерно по T и α ; здесь $a'_{2qx}(x, \alpha)$ – матрица Якоби вектор-функции $a_{2q}(x, \alpha)$.

Определим зависящие от вспомогательного параметра $\varepsilon > 0$ функционалы

$$\alpha(u, \varepsilon) = \alpha_0 + \frac{1}{\varepsilon} [(u, \tilde{e}) - \varepsilon], \quad T(u, \varepsilon) = T_0 + \frac{1}{\varepsilon} (u, \tilde{g}) \quad (11)$$

и подставим их соответственно вместо α и T в уравнение (10). В результате получим уравнение

$$F(u, \varepsilon) \equiv G(u, \varepsilon) + W(u, \varepsilon) = 0, \quad u \in R^N, \quad (12)$$

где

$$G(u, \varepsilon) = u - B[T(u, \varepsilon), \alpha(u, \varepsilon)]u, \quad W(u, \varepsilon) = -b[u, T(u, \varepsilon), \alpha(u, \varepsilon)].$$

Проверка показывает, что для уравнения (12) в шаре $S(\varepsilon)$ радиуса $r = \varepsilon/2$ с центром в точке εe при малых ε выполнены все условия сходимости метода Ньютона-Канторовича с возмущениями [13]. Поэтому уравнение (12) при малых $\varepsilon > 0$ имеет ненулевое решение $u^*(\varepsilon) \in S(\varepsilon)$, которое может быть построено с помощью итераций.

$$u_{n+1}(\varepsilon) = u_n(\varepsilon) - G_0 F(u_n(\varepsilon), \varepsilon), \quad n = 0, 1, 2, \dots,$$

где F – оператор (12), $G_0 = (G')^{-1} : R^N \rightarrow R^N$, $u_0(\varepsilon) = \varepsilon e$. Здесь G' – производная оператора $G(u, \varepsilon)$ в точке $u = \varepsilon e$; по построению эта производная не зависит от ε и является обратимой. Подставляя вектор $u^*(\varepsilon)$ в формулы (11), получаем утверждение леммы 1.

Лемма 2. При $\alpha \neq 0$ замена $x = \alpha^{1/2q-1}u$ переводит уравнение (2) в уравнение (8). Обратная замена переводит уравнение (8) в уравнение (2).

Справедливость этого утверждения устанавливается простым подсчетом.

Вернемся к доказательству теорем 1 и 2. Из лемм 1 и 2 следует, что уравнение (2) при $\alpha = \alpha(\varepsilon)$ имеет $T(\varepsilon)$ – периодическое решение $y = y(t, \varepsilon)$, при этом вектор $\tilde{y}_0(\varepsilon) = y(0, \varepsilon)$ представим в виде

$$y(0, \varepsilon) = (\alpha_2 \varepsilon^{2q} + O(\varepsilon^{2q+1}))^{\frac{1}{1-2q}} (\varepsilon e + O(\varepsilon^{2q})) = \alpha_2^{\frac{1}{1-2q}} \cdot \varepsilon^{\frac{1}{1-2q}} e + o\left(\varepsilon^{\frac{1}{1-2q}}\right).$$

Так как $2q - 1 > 0$, то $\|y(0, \varepsilon)\| \rightarrow \infty$ при $\varepsilon \rightarrow 0$. Таким образом $\max_t \|y(t, \varepsilon)\| \rightarrow \infty$ при $\varepsilon \rightarrow 0$, что завершает доказательство теорем 1 и 2.

Доказательство теоремы 3

По построению, бифурцирующие решения $y(t, \varepsilon)$ системы (2) устойчивы тогда и только тогда, когда устойчивы бифурцирующие решения $x(t, \varepsilon)$ системы (8). В свою очередь, при малых ε свойства устойчивости решений $x(t, \varepsilon)$ системы (8) определяются знаком ляпуновской величины L_1 в задаче о бифуркации Андронова–Хопфа, которая, в соответствии с работой [14], определяется равенством $L_1 = (b_3, \tilde{e})$. Отсюда и из [14] следует, что если $L_1 < 0$, то при малых ε решения $y(t, \varepsilon)$ системы (2) орбитально асимптотически устойчивы, а если $L_1 > 0$, то неустойчивы. Теорема 3 доказана.

3. Пример: модель Рёсслера

В качестве иллюстрации рассмотрим модель Рёсслера, описываемую системой дифференциальных уравнений с квадратичной нелинейностью

$$\begin{aligned} x' &= -y - z, \\ y' &= x, \\ z' &= \alpha(y - y^2) - bz. \end{aligned} \quad (13)$$

В этой модели α и b – неотрицательные параметры. Модель Рёсслера (см., например, [11]) демонстрирует богатое бифуркационное и хаотическое поведение при различных значениях α и b .

Пусть для определенности $b = 1$. Покажем, что значение $\alpha = 0$ является точкой бифуркации Андронова–Хопфа системы (13) в окрестности нулевой точки равновесия $x =$

$y = z = 0$ и точкой бифуркации на бесконечности. Для этого воспользуемся теоремами 1–3.

Система (13) имеет вид (2):

$$u' = A_0 u + \alpha f(u), \quad u \in R^3,$$

где

$$A_0 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 0 & -b \end{bmatrix}, \quad f(u) = A_1 u + a_2(u), \quad A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad a_2(u) = \begin{bmatrix} 0 \\ 0 \\ -u_2^2 \end{bmatrix}.$$

При $\alpha = 0$ и $b = 1$ матрица A_0 имеет собственные значения $\lambda_{1,2} = \pm i\omega_0$ и $\lambda_3 = -b = -1$, где $\omega_0 = 1$. Несложно убедиться в том, что в качестве векторов $e, g, \tilde{e}, \tilde{g}$ можно взять векторы

$$e = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad g = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \tilde{e} = \begin{bmatrix} 0 \\ 1 \\ 1/2 \end{bmatrix}, \quad \tilde{g} = \begin{bmatrix} 1 \\ 0 \\ -1/2 \end{bmatrix}.$$

Подсчет показывает, что числа (3) и (4) здесь равны: $\gamma_1 = 1/2, \gamma_2 = 0$. Поэтому, в соответствии с теоремой 1, значение $\alpha = 0$ является для системы (13) точкой бифуркации циклов на бесконечности. В то же время в силу леммы 1 значение $\alpha = 0$ является и точкой бифуркации Андронова–Хопфа в окрестности нулевой точки равновесия системы (13).

Вычислим числа α_2 и T_2 по формуле (4): $\alpha_2 = -49105,58$ и $T_2 = 0$. Поскольку $\alpha_2 < 0$, то в соответствии с теоремой 2 циклы больших амплитуд в системе (13) возникают при $\alpha < 0$. Наконец, вычисления показывают, что здесь $(b_3, \tilde{e}) = 77134,86 > 0$. Отсюда и из теоремы 3 получим, что циклы больших амплитуд системы (13) неустойчивы.

Список источников

1. Красносельский А. М., Красносельский М. А. Циклы больших амплитуд в автономных системах с гистерезисом // Доклады Академии наук. Российская академия наук. 1985. Т. 283, № 1. С. 23–26.
2. He X. Hopf bifurcation at infinity with discontinuous nonlinearities // J. Austr. Math. Soc. Ser. B. 1991. Vol. 33, № 2. P. 133–148. URL: <https://doi.org/10.1017/S0334270000006950> (дата обращения: 02.10.2024).
3. Fura A., Rybicki S. Periodic solutions of second order Hamiltonian systems bifurcating from infinity // Ann. Inst. H. Poincaré Anal. Non Linéaire. 2007. Vol. 24, № 3. P. 471–490. DOI: 10.1016/J.ANIHPC.2006.03.003 (дата обращения: 02.10.2024).
4. Sabatini M. Hopf bifurcation from infinity // Rendiconti del Seminario Matematico della Università di Padova. 1987. Vol. 78. P. 237–253. URL: https://www.numdam.org/item/RSMUP_1987__78__237_0 (дата обращения: 02.10.2024).
5. Красносельский А. М. Вырожденный случай бифуркации Андронова–Хопфа на бесконечности // Автоматика и телемеханика. 2010. № 11. С. 55–68.
6. Юмагулов М. Г., Ибрагимова Л. С., Кунгиров М. Н. Бифуркации периодических решений в двухпараметрических автономных системах // Вестник Башкирского университета. 2023. Т. 28, № 2. С. 154–157.
7. Kungirov M. N. Bifurcation of periodic oscillations arising from a closed phase curve in systems with odd nonlinearities // Lobachevskii journal of mathematics. 2024. Vol. 45,

- по. 6. Р. 2739–2745. URL: <https://doi.org/10.1134/S1995080224603254> (дата обращения: 02.10.2024).
8. Юмагулов М. Г., Ибрагимова Л. С., Имангулова Э. С. Главные асимптотики в задаче о бифуркации Андронова–Хопфа и их приложения // Дифференциальные уравнения. 2017. Т. 53, № 12. С. 1627–1643.
 9. Имангулова Э. С. Синхронизация периодических колебаний в задаче о бифуркации Андронова–Хопфа // Вестник Башкирского университета. 2017. Т. 22, № 2. С. 292–296.
 10. Популяционная динамика: модели и методы / Юмагулов М. Г., Ибрагимова Л. С., Белова А. С. Уфа: РИЦ УУНиТ, 2022. 85 с.
 11. Элегантный хаос: алгебраически простые хаотические потоки / Спротт Д. К. М. Ижевск: РХД; ИКИ, 2012. 328 с.
 12. Musafirov E., Grin A., Pranevich A., Munteanu F., Sterbeti C. 3D quadratic ODE systems with an infinite number of limit cycles // ITM Web of Conferences. 2022. Vol. 49. P. 1–6.
 13. Приближённое решение операторных уравнений. / Красносельский М.А., Вайникко Г.М., Забрейко П.П., Стеценко В.Я. М.: Наука. 1969. 456 с. URL: <http://elib.bsu.by/handle/123456789/12803>. (дата обращения: 02.10.2024).
 14. Гусарова Н. И., Муртазина С. А., Фазлытдинов М. Ф., Юмагулов М. Г. Операторные методы вычисления ляпуновских величин в задачах о локальных бифуркациях динамических систем // Уфимский математический журнал. 2018. Т 10, № 1. С. 25–49.

References

1. Krasnosel'skii, A. M. and Krasnosel'skii, M. A. (1985), "Large amplitude cycles in autonomous systems with hysteresis", *Doklady Akademii Nauk. Rossiyskaya Akademiya Nauk*, vol. 283, no. 1, pp. 23–26.
2. He, X. (1991), "Hopf bifurcation at infinity with discontinuous nonlinearities", *The ANZIAM Journal*, vol. 33, issue 2, pp. 133–148.
3. Fura, J. and Rybicki, S. (2007), "Periodic solutions of second order Hamiltonian systems bifurcating from infinity", *In Annales de l'IHP Analyse non linéaire*, vol. 24, no. 3, pp. 471–490.
4. Sabatini, M. (1987), "Hopf bifurcation from infinity", *Rendiconti del Seminario Matematico della Università di Padova*, vol. 3, no. 1, pp. 1–17.
5. Krasnosel'skii, A. M. (2010), "A degenerate case of Andronov–Hopf bifurcation at infinity", *Automation and Remote Control*, no. 11, pp. 55–68.
6. Yumagulov, M. G., Ibragimova, L. S. and Kungirov, M. N. (2023), "Bifurcations of periodic solutions in two-parameter autonomous systems", *Vestnik Bashkirskogo universiteta*, vol. 28, no. 2, pp. 154–157.
7. Kungirov, M. N. (2024), "Bifurcation of periodic oscillations arising from a closed phase curve in systems with odd nonlinearities", *Lobachevskii Journal of Mathematics*, vol. 45, no. 6, pp. 2739–2745.

8. Yumagulov, M. G., Ibragimova, L. S. and Imangulova, E. S. (2017), "Principal asymptotics in the problem on the Andronov–Hopf bifurcation and their applications", *Differential Equations*, vol. 53, no. 12, pp. 1627-1643.
9. Imangulova, E. S. (2017), "Synchronization of periodic oscillations in the Andronov–Hopf bifurcation problem", *Vestnik Bashkirskogo universiteta*, vol. 22, no. 2, pp. 292–296.
10. Yumagulov, M. G., Ibragimova, L. S. and Belova, A. S. (2022), *Populyacionnaya dinamika: modeli i metody* [Population Dynamics: Models and Methods], Ufa University of Science and Technology, Ufa, Russia.
11. Sprott, J. K. (2012), *Elegantnyi kaos: algebraichesky prostye haoticheskie potoki* [Elegant Chaos: Algebraically Simple Chaotic Flows], Izhevsk, Russia.
12. Musafirov, E., Grin, A., Pranevich, A., Munteanu, F. and Șterbeți, C. (2022), "3D Quadratic ODE systems with an infinite number of limit cycles", *In ITM Web of Conferences*, vol. 49, p. 02006.
13. Krasnosel'skiy, M. A., Vainikko, G. M., Zabreiko, P. P. and Stetsenko, V. Ya. (1969), *Priblizhennoe reshenie operatornykh uravnenij* [Approximate Solution of Operator Equations], Nauka, Moscow, Russia.
14. Gusarova, N. I., Murtazina, S. A., Fazlytdinov, M. F. and Yumagulov, M. G. (2018), "Operator methods for calculating Lyapunov values in problems on local bifurcations of dynamical systems", *Ufimskii Matematicheskii Zhurnal*, vol. 10, no. 1, pp. 25-49.

Информация об авторе:

М. Н. Кунгиров – аспирант, Уфимский университет науки и технологий (450008, Россия, г. Уфа, ул. Заки Валиди, 32).

Information about the author:

M. N. Kungirov – Postgraduate student, Ufa University of Science and Technology (32, Zaki Validi St., Ufa, Russia, 450008).

Обзорная статья

УДК 001.32

ОИ: 10.17072/1993-0550-2025-2-15-20

<https://elibrary.ru/uuygmm>



Метеоролог и математик. К 70-летию со дня рождения профессора К. Г. Шварца

Олег Геннадьевич Пенский

г. Пермь, Россия

ogpensky@mail.ru

Аннотация. Представлена краткая биография доктора физико-математических наук, члена диссертационного совета ПГНИУ по географическим наукам, профессора Константина Григорьевича Шварца, приведены основные этапы его научной и популяризаторской деятельности в Пермском государственном национальном исследовательском университете.

Ключевые слова: *Пермский государственный университет; биография; наука; юбилей; К. Г. Шварц*

Для цитирования: Пенский О. Г. Метеоролог и математик. К 70-летию профессора К. Г. Шварца // Вестник Пермского университета. Математика. Механика. Информатика. 2025. Вып. 2(69). С. 15-20. DOI: 10.17072/1993-0550-2025-2-15-20. <https://elibrary.ru/uuygmm>

Статья поступила в редакцию 08.04.2025; одобрена после рецензирования 13.05.2025; принята к публикации 09.07.2025.

Review article

Meteorologist and Mathematician. On the 70 Anniversary of the Birth of Professor K. G. Schwartz

Oleg G. Pensky

Perm, Russia

ogpensky@mail.ru

Abstract. A brief biography of Professor Konstantin Grigorievich Schwartz, Doctor of Physico-Mathematical Sciences, member of the PSU Dissertation Council for Geographical Sciences, is presented, and the main stages of his scientific and popularization activities at Perm State University are presented.

Keywords: *Perm State University; biography; anniversary; science; C. G. Schwartz*

For citation: Pensky, O. G. (2025), "Meteorologist and Mathematician. On the 70 Anniversary of the Birth of Professor K. G. Schwartz", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 15-20. (In Russ.). DOI: 10.17072/1993-0550-2025-2-15-20. <https://elibrary.ru/uuygmm>



Эта работа © 2025 Пенский О. Г. распространяется по лицензии CC BY 4.0.. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>.

The article was submitted 08.04.2025; approved after reviewing 13.05.2025; accepted for publication 09.07.2025.



В 2025 г. профессор физико-математического института ПГНИУ Константин Григорьевич Шварц отметил свой 70-летний юбилей, а в 2026 г. исполняется 50 лет его трудовой деятельности в Пермском государственном университете [1]. В связи с этими датами мы решили написать статью о его научной и трудовой деятельности.

Константин Григорьевич Шварц родился в 1955 г. в г. Горький (ныне – Нижний Новгород) в семье офицера. В 1977 г. с отличием окончил механико-математический факультет Пермского государственного университета по специальности "Математика". В студенческие годы выступал

за сборную университета по шахматам, за активную научную деятельность получил почетный диплом СНО (студенческого научного общества), дающее преимущественное право поступления в аспирантуру. Занимался функционально-дифференциальными уравнениями, что в будущем ему пригодилось, однако, несмотря на наличие диплома СНО, в аспирантуру принят не был.

В настоящее время Константин Григорьевич – доктор физико-математических наук, с 2001 года – профессор кафедры прикладной математики и информатики механико-математического факультета ПГНИУ, с 2004 по 2018 годы – по совместительству заведовал кафедрой математики и естественнонаучных дисциплин в Пермском филиале РАНХиГС [2].

По окончании университета Константин Григорьевич работал программистом в вычислительном центре Пермского государственного университета им. А.М. Горького. В те годы студентов и сотрудников университета осенью отправляли в помощь колхозникам для проведения уборочных работ. Молодой программист вычислительного центра ПГУ К.Г. Шварц неизменно назначался руководителем сельскохозяйственных студенческих отрядов, работавших на уборке урожая в селе Большая Соснова Пермской области. Константин Григорьевич прекрасно играл на гитаре, и его сольные концерты вечерами у костра с исполнением песен популярных в Советском Союзе бардов, снискали огромное уважение среди студентов-слушателей.

В вычислительном центре работал в группе, которая разрабатывала АСНИ (автоматизированную систему управления научными экспериментами) в содружестве с МИРЭА (Московским институтом радиоэлектроники и автоматики). Этот коллектив

воспитал очень хороших для того времени программистов. В дальнейшем из них выросли три кандидата наук и один доктор физико-математических наук.



ВЦ ПГУ, группа АСНИ, машинный зал ЭВМ СМ-4 май 1982 г.



Стажировка DAAD, Потсдамский университет, Германия Сентябрь–декабрь 1996 г.



Приглашенный профессор, Лильский университет, Франция, март 2004 г.



Международная конференция BIFD2015, Высшая школа промышленной физики и химии им. Ж. Кюри в Париже (ESPSI), Люксембургский дворец, Париж, Франция, 15–17 июля 2015 г.



Кафедра математики и естественнонаучных дисциплин, Пермский филиал УрАГС. 2005 г.



Кафедра прикладной математики и информатики ПГНИУ. 2008 г.

В 1988 г. Константин Григорьевич поступил в заочную аспирантуру ИМСС УрО АН СССР. Его научным руководителем стал д.ф.-м.н., профессор Валерий Дмитриевич Зимин. В 1988–1989 гг. К. Г. Шварц помимо обучения в аспирантуре выполнял обязанности младшего научного сотрудника, а в 1989–1990 гг. – научного сотрудника научно-исследовательского сектора ПГУ. Под руководством В.Д. Зимины занимался численным моделированием тропических циклонов (тайфунов), работая в рамках хоздоговора между ИМСС и ПГУ с Институтом космических исследований АН СССР. Эта деятельность совпадала с темой его диссертации. В 1990 г. он досрочно защитил кандидатскую диссертацию на тему "Численное моделирование крупномасштабных вихревых процессов в тонком слое жидкости", и с 1990 г. начал работу ассистентом кафедры прикладной математики и информатики Пермского государственного университета, став в 1993 г. старшим преподавателем, в 1994 г. был принят на должность доцента, а в 1996 г. ему было присвоено звание доцента.

В 1996 г. К.Г. Шварц проходил трехмесячную стажировку ДААД в лаборатории нелинейной динамики Потсдамского университета (Германия).

В 2000 г. Константин Григорьевич защитил докторскую диссертацию на тему "Адвективные течения во вращающемся слое жидкости или газа", научным консультантом диссертационного исследования стал д.ф.-м.н., г. н. с. ИМСС УрО РАН Сергей Николаевич Аристов, с которым К.Г. Шварц был очень дружен.

В 2002 г. молодой доктор наук работал приглашенным профессором на факультете механики технического университета Ильменау (Германия), в 2004 г. – приглашенным профессором в лаборатории механики университета науки и технологии г. Лиль (Франция), в 2012 г. прошел научную стажировку в технологическом институте Технион в г. Хайфа (Израиль), а в 2014 г. – в Ариэльском университете (Израиль).

Область научных интересов К.Г. Шварца – нелинейные проблемы геофизической гидродинамики, численное и математическое исследование адвективных течений во вращающихся средах, моделирование переноса примесей в атмосфере. Константин Григорьевич неоднократно был руководителем грантов РФФИ. Он подготовил одного кандидата наук. В настоящее время у него несколько аспирантов. Практические приложения исследований Константина Григорьевича Шварца связаны с изучением процессов зарождения тайфунов, прогнозированием распространения вредных примесей в атмосфере и больших водоемах, в том числе океанах и морях; с построением оценок последствий техногенных катастроф методами математического моделирования. В последнее время занялся изучением воздействия плесени, пыльцы растений, распространяемых в воздушной среде, на здоровье людей в крупных агломерациях. Его исследования проводятся совместно с преподавателями кафедры метеорологии и охраны атмосферы географического факультета и кафедры генетики растений биологического факультета, они входят в направление научной деятельности Пермского государственного национального исследовательского университета по теме "Природоохрана и природопользование". Профессор К.Г. Шварц с 2002 г. является членом диссертационного совета по географическим наукам, функционирующим в нашем университете, в совете он выполняет обязанности эксперта по метеорологии [3].

Вместе с тем Константин Григорьевич уделяет внимание популяризации результатов научных исследований, которыми он занимается [4]. Константин Григорьевич Шварц – автор более 240 публикаций, среди которых, в частности, 28 статей в изданиях из списка Scopus и Web of Science, 59 статей в журналах из списка ВАК. Он опубликовал три монографии "Вихревые течения адвективной природы во вращающемся слое жидкости" (2006), "Вихревые течения в тонких слоях жидкости" (2011) в соавторстве с профессором С.Н. Аристовым и совместно с доцентом В.А. Шкляевым "Математическое моделирование мезомасштабных и крупномасштабных процессов переноса примеси в бароклинной атмосфере" (2015). Имеет шесть учебных и учебно-методических пособий. Член редколлегии журнала Вестник Пермского университета. Математика. Механика. Информатика и редколлегии научного журнала "ИТ АРКТИКА", издаваемым государственным автономным учреждением Республики Коми "Центр информационных технологий" (ГАУ РК "ЦИТ"). Участник более 20 международных научных конференций, проводимых в СССР, России, Великобритании, Германии, Израиле, Испании, США, Франции.

Константин Григорьевич неоднократно становился лауреатом научно-исследовательской деятельности ПГНИУ и лучшим ученым механико-математического факультета, он награжден медалью механико-математического факультета имени Л. Эйлера "За заслуги" (2006), медалью ПГНИУ им. А. Фридмана "За фундаментальные работы и перспективные исследования" (2021), является Почетным работником высшего профессионального образования РФ (2012).

Коллеги поздравляют Константина Григорьевича Шварца с юбилеем и желают ему дальнейшей успешной работы в науке.

Список источников

1. *Пенский О. Г.* Два юбилея ученого (к 60-летию со дня рождения К. Г. Шварца) // Вестник Пермского университета. Математика. Механика. Информатика. Вып. 1(32). 2016. С. 52–54. EDN: WFEFSZ
2. *Константин Григорьевич ШВАРЦ*: Персональная страница преподавателя. URL: <http://www.psu.ru/personalnye-stranitsy-prepodavatelej/sh/konstantin-grigorevich-shvarts> (дата обращения: 30.04.2025).
3. *Состав* диссертационного совета ПГНИУ, географические науки. URL: <http://www.psu.ru/nauka/dissertatsionnye-sovety/geograficheskie-nauki> (дата обращения: 30.04.2025).
4. *Шварц К. Г.* Гидродинамика природных явлений // Цифровой университет. Передача 14. 2010. <http://www.psu.ru/news/3146> (дата обращения: 30.04.2025).

References

1. Penskiy, O. G. (2016), "Dva yubileya uchenogo (k 60-letiyu so dnya rozhdeniya K. G. Shvarca)" [Two anniversaries of the scientist (dedicated to the 60th anniversary of the birth of K. G. Schwartz)], *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 1(32), pp. 52–54.

2. Personal Page of Konstantin Grigorievich Shvartz (accessed April 30, 2025), available at: <http://www.psu.ru/personalnye-stranitsy-prepodavatelej/sh/konstantin-grigorevich-shvarts>
3. Dissertation Councils in Geographical Sciences (accessed April 30, 2025), available at: <http://www.psu.ru/nauka/dissertatsionnye-sovety/geograficheskie-nauki>
4. Shvarts K. G. (2010), "Gidrodinamika prirodnikh yavleniy", *Tsifrovoy universitet, Transmission 14*. Available at: <http://www.psu.ru/news/3146> (Accessed April 30, 2025).

Информация об авторе:

О. Г. Пенский – доктор технических наук, профессор.

Information about the author:

O. G. Pensky – Doctor of Technical Sciences, Professor.

МЕХАНИКА

Научная статья

УДК 004.94, 533.65, 681.515

DOI: 10.17072/1993-0550-2025-2-21-35

<https://elibrary.ru/thonoj>



Моделирование робастных алгоритмов управления подъемом и горизонтальной стабилизацией беспилотного трикоптера

Александр Николаевич Власов¹, Александр Юрьевич Ощепков²

^{1,2}Пермский государственный национальный исследовательский университет, г. Пермь, Россия

¹anvlasov99@yandex.ru

²aos57@mail.ru

Аннотация. Рассматривается задача управления подъемом и горизонтальной ориентацией беспилотного трикоптера с тяговыми двигателями, расположенными в одной плоскости, в вершинах правильного треугольника. Построена математическая модель динамики трикоптера с учетом наличия нагрузки с переменной массой и координатой приложения опрокидывающей силы. Представлены алгоритм управления для стабилизации по высоте и ориентации с учетом углов крена и тангажа, основанный на ПИД-регуляторе и алгоритм управления тягой двигателей трикоптера для стабилизации его только по высоте без учета отклонения ориентации, использующий быстрый алгоритм 2-го порядка. Построена компьютерная модель системы управления трикоптером в пакете Simulink. С помощью модели исследована работа быстрого алгоритма для управления взлетом и посадкой летательного аппарата.

Ключевые слова: моделирование процессов управления; робастные алгоритмы; система стабилизации; беспилотные летательные аппараты; трикоптеры

Для цитирования: Власов А. Н., Ощепков А. Ю. Моделирование робастных алгоритмов управления подъемом и горизонтальной стабилизацией беспилотного трикоптера // Вестник Пермского университета. Математика. Механика. Информатика. 2025. Вып. 2(69). С. 21–35. DOI: 10.17072/1993-0550-2025-2-21-35. <https://elibrary.ru/thonoj>

Статья поступила в редакцию 21.02.2025; одобрена после рецензирования 27.05.2025; принята к публикации 09.07.2025.

MECHANICS

Research article

Modeling of Robust Algorithms for Control of Lift and Horizontal Stabilization of an Unmanned Tricopter

Alexander N. Vlasov¹, Alexander Yu. Oshchepkov²

^{1,2}Perm State University, Perm, Russia

¹anvlasov99@yandex.ru

²aos57@mail.ru



Эта работа © 2025 Власов А. Н., Ощепков А. Ю. распространяется по лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>.

Abstract. The problem of control of the lift and horizontal orientation of an unmanned tri-copter with traction motors located in the same plane, at the vertices of a regular triangle, is considered. A mathematical model of the tricopter dynamics is constructed taking into account the presence of a load with a variable mass and the coordinate of application of the overturning force. A control algorithm for stabilization in height and orientation taking into account the angles of roll and pitch, based on a PID controller, and an algorithm for controlling the tricopter engine thrust for stabilizing it only in height without taking into account the orientation deviation, using a fast algorithm of the 2nd order are presented. A computer model of the tricopter control system in the Simulink package is constructed. Using the model, the work of the fast algorithm for controlling the takeoff and landing of the aircraft is studied.

Keywords: *modeling of control processes; robust algorithms; stabilization system; unmanned aerial vehicles; tricopters*

For citation: Vlasov, A. N. and Oshchepkov, A. Yu. (2025), "Modeling of Robust Algorithms for Control of Lift and Horizontal Stabilization of an Unmanned Tricopter", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 21–35. (In Russ.). DOI: 10.17072/1993-0550-2025-2-21-35. <https://elibrary.ru/thonoj>

The article was submitted 21.02.2025; approved after reviewing 27.05.2025; accepted for publication 09.07.2025.

Введение

Мультикоптеры относятся к беспилотным летательным аппаратам (БПЛА) с двигателями роторного типа, обеспечивающими вертикальный взлет и посадку, а также полет по маршруту и управление углами тангажа, крена и рысканья для обеспечения навигационного полетного задания. Обычно используются многороторные БПЛА с тремя двигателями (трикоптеры), четырьмя (квадрокоптеры) и шестью двигателями (гексакоптеры).

Для обеспечения заданного вектора тяги необходимо три управляющих элемента, поэтому квадрокоптеры и гексакоптеры обладают избыточным количеством органов управления по сравнению с числом контролируемых параметров. В работе [1] отмечается, что такая избыточность обеспечивает отказобезопасность системы управления. Однако повышение количества двигателей, пусть и увеличивает подъемные возможности и стабильность полета, вместе с тем оно увеличивает энергопотребление и массу аппарата. Кроме того, мультикоптер с меньшим количеством двигателей обладает большей маневренностью. Таким образом, конструкция трикоптера представляется оптимальной и весьма перспективной, хотя разработка его является еще малоизученной [2].

В данной работе исследуются возможности применения традиционных и новых алгоритмов для наискорейшего управления высотой полета трикоптера с помощью двигателей вертикальной тяги, а также рассматриваются процессы управления углами крена и тангажа при несимметричном изменении массы платформы, например при сбрасывании с нее какого-либо груза. Работоспособные алгоритмы в дальнейшем могут быть реализованы в бортовых системах управления с помощью методов модельно-ориентированного проектирования [3, 4].

1. Постановка задачи

1.1. Математическая модель трикоптера с двигателями вертикальной тяги

Исследование работы алгоритмов управления осуществляется посредством компьютерного моделирования на основе специально разработанной математической модели трикоптера.

Трикоптер будем представлять в виде симметричной плоской платформы, на которой размещены три реверсивных двигателя вертикальной тяги, при этом работа двигателей не вызывает реактивного вращения платформы в плоскости трикоптера. Двигатели находятся в вершинах правильного треугольника со стороной $2a$ в точках 1, 2 и 3 (рис. 1). Массу платформы с двигателями обозначим через M_0 и будем считать, что она равномерно распределена по этим точкам. Начало осей собственной системы координат (ССК) модели выберем в центре треугольника, направление осей X и Y указано на рис. 1.

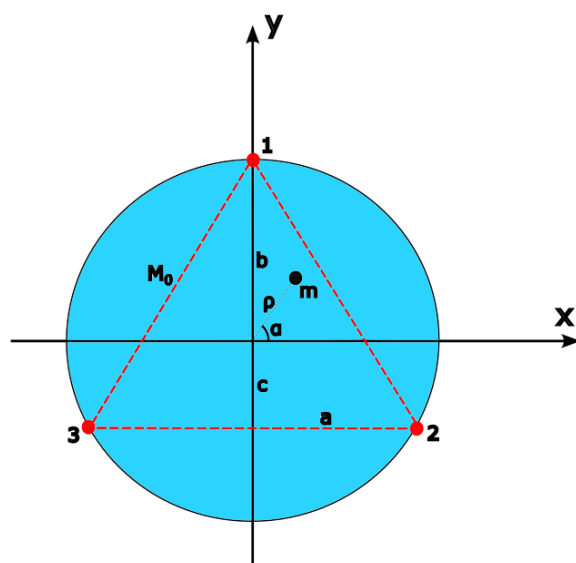


Рис. 1. Геометрическая модель трикоптера, собственная система координат (ССК)

На платформе находится груз массы m в точке с полярными координатами (ρ, α) в ССК, полная масса объекта

$$M = M_0 + m. \quad (1)$$

Задача состоит в том, чтобы поднять трикоптер вдоль вертикальной оси Z системы координат, связанной с землёй, на заданную высоту h и удерживать платформу в горизонтальном положении как при наличии груза, так и при его удалении (сбрасывании). В другой постановке задачи: трикоптер поднимается на заданную высоту без груза, затем груз прикрепляется к указанной точке, платформа должна сохранять горизонтальное положение. Высота подъема предполагается незначительной, поэтому ускорение силы тяжести g вдоль вертикальной оси считается постоянным.

При известных параметрах объекта задача выбора тяговых усилий для обеспечения горизонтального равновесия платформы в вышеописанной простой модели имеет аналитическое решение. Если начало координаты Z выбрать на поверхности земли, а ось направить вертикально вверх, то **при отсутствии груза** ($m = 0$) уравнение верти-

кального движения примет вид (здесь и далее аэродинамическим трением пренебрегаем):

$$M_0 \ddot{z} = -M_0 g + F_1 + F_2 + F_3, \quad z(0) = 0, \quad \dot{z}(0) = 0. \quad (2)$$

Здесь F_1, F_2, F_3 – силы тяги каждого из двигателей. Для обеспечения горизонтального положения платформы достаточно выполнение условия

$$F_1 = F_2 = F_3 = \frac{1}{3} F_T, \quad (3)$$

где F_T – суммарная тяга двигателей для обеспечения необходимого ускорения. В равновесии (по высоте)

$$F_T = M_0 g. \quad (4)$$

При наличии груза ($m \neq 0$) возникает момент сил, приводящий к опрокидыванию платформы. Угол поворота вокруг оси Y обозначим φ (угол крена), угол поворота вокруг оси X – ϑ (угол тангажа) (рис. 2).

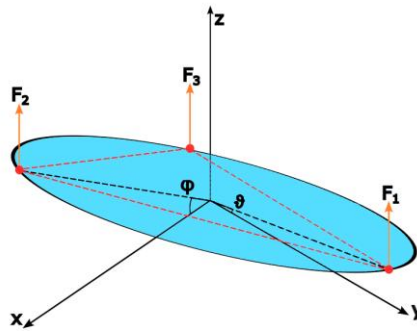


Рис. 2. Углы крена и тангажа. Оси X, Y лежат в горизонтальной плоскости, оси ССК трикоптера направлены по соответствующим пунктирным линиям

Будем считать углы ϑ, φ малыми, так что в линейном приближении по этим углам уравнение вращательного движения в ССК запишется в виде:

$$I_X \ddot{\vartheta} = -c(F_2 + F_3) + bF_1 - mg\rho \sin \alpha; \quad (5)$$

$$I_Y \ddot{\varphi} = a(F_2 - F_3) + mg\rho \cos \alpha,$$

где моменты инерции I_X, I_Y выражаются через геометрические характеристики платформы и координаты груза следующим образом:

$$\begin{aligned} I_X &= \frac{1}{3} M_0 (2c^2 + b^2) + m(\rho \sin \alpha)^2; \\ I_Y &= \frac{1}{3} M_0 (2a^2) + m(\rho \cos \alpha)^2. \end{aligned} \quad (6)$$

Уравнение движения вдоль Z в рассматриваемом случае сохраняет вид (2), где M_0 следует заменить на M в соответствии с (1). Полная сила тяги для сохранения заданной высоты при этом будет равна

$$F_T = (M_0 + m)g, \quad (7)$$

однако ее распределение по двигателям, обеспечивающее горизонтальную ориентацию платформы, будет отличаться от (3).

Запишем требуемые силы тяги для каждого двигателя в виде

$$F_i = \frac{1}{3}F_T + U_i, \quad i=1,2,3. \quad (8)$$

Введем обозначения для матриц-столбцов размерности 3×1 :

$$\bar{F} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}, \quad \bar{U} = \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix}, \quad (9)$$

после чего выражение (8) может быть записано в матричной форме

$$\bar{F} = \frac{1}{3}F_T + \bar{U}. \quad (10)$$

Поскольку суммарная тяга двигателей должна соответствовать выражению (7), компоненты матрицы \bar{U} должны удовлетворять условию

$$\sum_{i=1}^3 U_i = 0. \quad (11)$$

Решение уравнений (5) для состояния равновесия при условии (11) приводит к следующим выражениям для компонентов матрицы \bar{U} :

$$\begin{aligned} U_i &= \frac{1}{3}mgu_i, \quad i=1,2,3; \\ u_1 &= \sqrt{3} \frac{\rho}{a} \sin \alpha, \\ u_2 &= -\frac{1}{2} \frac{\rho}{a} (\sqrt{3} \sin \alpha + 3 \cos \alpha), \\ u_3 &= -\frac{1}{2} \frac{\rho}{a} (\sqrt{3} \sin \alpha - 3 \cos \alpha). \end{aligned} \quad (12)$$

1.2. Задача управления трикоптером

Система управления с обратными связями «не знает» параметров объекта управления. Ее задача состоит в том, чтобы обеспечить требуемую суммарную силу тяги и распределить ее по двигателям в автоматическом режиме. Для этого на трикоптер уста-

наливаются три датчика: датчик высоты, измеряющий координату z , датчик угла крена φ и датчик угла тангажа ϑ . Три независимых блока управления вырабатывают сигналы управления $U(z), U(\varphi), U(\vartheta)$ в соответствии с заложенными в них алгоритмами. Цели управления на всем протяжении полета:

$$z \rightarrow h, \varphi \rightarrow 0, \vartheta \rightarrow 0. \quad (13)$$

Сигналы управления передаются в блок расчета управляющих воздействий, который определяет суммарную тягу двигателей и компоненты матрицы \bar{U} по формулам, полученным с учетом симметрии объекта и условия (11):

$$\begin{aligned} F_T &= U(z), \quad U_1 = U(\vartheta), \\ U_2 &= U(\varphi) - \frac{1}{2}U(\vartheta), \\ U_3 &= -U(\varphi) - \frac{1}{2}U(\vartheta). \end{aligned} \quad (14)$$

Блок-схема системы управления приведена на рис. 3. Математическая модель объекта управления в схеме описывается уравнениями движения (2) и (5) с учетом выражений (6).

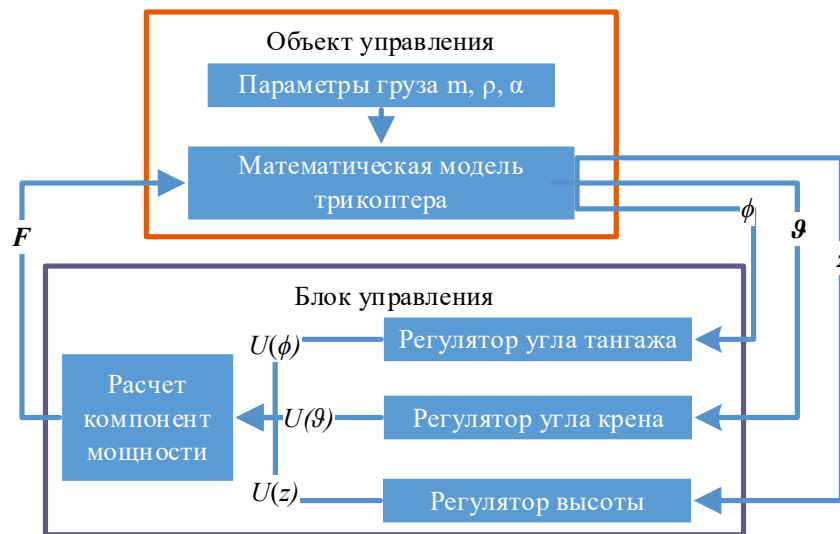


Рис. 3. Блок-схема системы управления высотой трикоптера h и углами крена ϑ и тангажа φ

2. Моделирование работы системы управления подъемом

2.1. Дискретная модель системы

Рассмотрим моделирование задачи подъема и стабилизации высоты при отсутствии груза и его результаты. Для синтеза регуляторов проведена дискретизация уравнения (2) для тела массы M по схеме Эйлера [7] с шагом дискретизации Δt . Дискретная система, эквивалентная непрерывной, имеет вид:

$$\begin{aligned} X_1[k+1] &= X_1[k] + \Delta t \cdot X_2[k], \\ X_2[k+1] &= X_2[k] - \Delta t \cdot g + \frac{\Delta t}{M} F_T[k]. \end{aligned} \quad (15)$$

Здесь введены обозначения

$$\begin{aligned} X_1 &= z - h, \quad X_2 = \dot{z}; \quad X_1(0) = -h, \quad X_2(0) = 0. \\ X(k \cdot \Delta t) &\equiv X[k], \quad k = 0, 1, 2, \dots \end{aligned} \quad (16)$$

Цель управления:

$$X_1[k] = 0, \quad X_2[k] = 0, \quad k > k_*. \quad (17)$$

2.2. S-модель системы управления

Блок-схема, изображенная на рис. 3, и модель системы, рассчитывающая динамику трикоптера и влияние нагрузки, реализована в пакете имитационного моделирования MATLAB+Simulink (рис. 4).

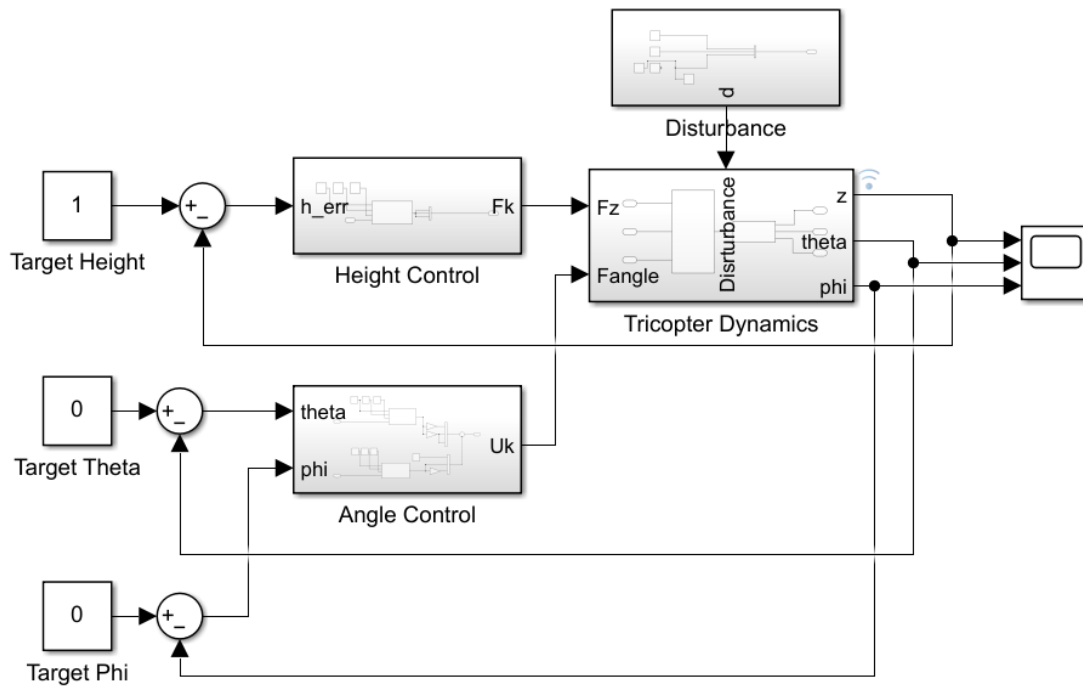


Рис. 4. S-модель, реализующая процесс управления высотой трикоптера h и углами крена ϑ и тангажа φ

Блоки Height Control и Angle Control отвечают за управление высотой и углами крена и тангажа соответственно. Также в них происходит распределение компонент мощности по двигателям. Блок Tricopter Dynamics включает в себя модель динамики трикоптера (уравнения движения 5, 7 и вычисление моментов инерции по уравнениям 6). Блок Disturbance отвечает за моделирование параметров нагрузки, меняющихся во времени.

2.3. Моделирование управления подъемом трикоптера с использованием ПИД-регулятора

ПИД-регуляторы – алгоритм, наиболее часто используемый в системах управления БПЛА из-за его простоты в реализации и настройке. Для моделирования был использован дискретный ПИД-регулятор [5], коэффициенты которого были подобраны с помощью программного пакета Response Optimizer. Дискретный ПИД-закон описывается выражением

$$F_T[k] = K_p X_1[k] + K_I \sum_{j=0}^k X_1[j] + K_D (X_1[k] - X_1[k-1]) \quad (18)$$

На рисунке 5 представлены результаты моделирования подъема трикоптера без дополнительной нагрузки с использованием ПИД-регуляторов.

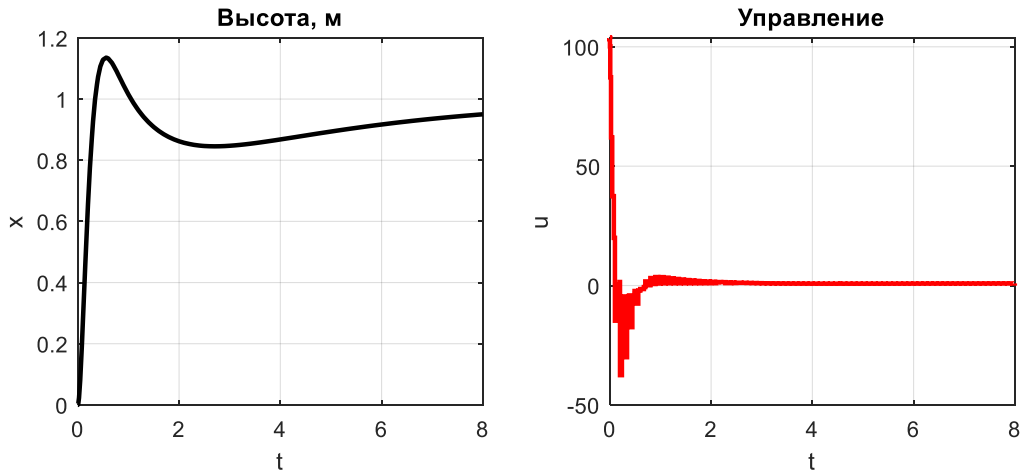


Рис. 5. Работа ПИД-регулятора

2.4. Моделирование управления подъемом трикоптера с использованием быстрого алгоритма

ПИД-регулятор, хоть и является стандартным законом управления для БПЛА, требует подстройки трех коэффициентов, что не всегда тривиально и требует применения специальных методик. В этом разделе рассмотрен алгоритм из класса быстрых алгоритмов [8], имеющий меньшее количество параметров и позволяющий в теории достичь цели управления за меньшее время. Этот алгоритм решает задачу терминального управления, которой, по сути, и является задача вывода трикоптера на заданную высоту [6].

Цель (17) для системы управления (15) с начальными условиями (16) может быть достигнута с помощью быстрого алгоритма для дискретных систем второго порядка [6]:

$$\begin{aligned} F_T[0] &= Mg; \\ F_T[k] &= Mg - \frac{1}{2} \left(1 - (-1)^k \right) \frac{M}{\Delta t^2} X_1[k] - \frac{M}{\Delta t} X_2[k], \\ k &= 1, 2, \dots \end{aligned} \quad (19)$$

Здесь учтена специфика задачи, заключающаяся в том, что воздействие $F_T[0] = Mg$ обеспечивается не работой двигателей, а реакцией стартовой опоры летательного аппарата, управление начинается с шага $k = 1$.

Управляющие воздействия, рассчитанные по формуле (19), и подставляемые последовательно в дискретную систему уравнений (15), принимают значения

$$\begin{aligned} F_T[1] &= Mg \cdot \left(1 + \frac{h}{g} \cdot \frac{1}{\Delta t^2}\right), F_T[2] = 0, F_T[3] = M \cdot \frac{h}{\Delta t^2}, \\ F_T[4] &= Mg + Mg \left(1 - \frac{h}{g} \cdot \frac{1}{\Delta t^2}\right), \\ F_T[k] &= Mg, X_1[k] = 0, X_2[k] = 0, k \geq 5. \end{aligned} \quad (20)$$

Как видим, быстрый алгоритм для дискретной системы достигает цели управления после 4-го шага, т.е. $k_* = 4$.

Поведение непрерывной системы приближается к ее дискретному аналогу при малых шагах дискретизации. Однако в рассматриваемом нами случае этот шаг не может быть сделан сколь угодно малым. Если двигатели могут обеспечить максимальную суммарную тягу

$$F_T^{\max} = K \cdot Mg, \quad K > 1, \quad (21)$$

то из выражения для $F_T[1]$ в (20) следует ограничение на минимальное значение шага дискретизации:

$$\Delta t > \sqrt{\frac{h}{g} \cdot \frac{1}{K-1}}, \quad K > 1. \quad (22)$$

С другой стороны, увеличение шага дискретизации ведет к отклонению поведения непрерывной системы от ее дискретного аналога, вследствие чего увеличивается количество шагов, необходимых для достижения цели, и снижается качество работы системы управления [8]. Таким образом, возникает задача определения области робастности дискретного алгоритма (19) применительно к непрерывному объекту (2). Задача решается с помощью моделирования системы управления в Simulink.

В S-модели дискретизация осуществляется с помощью блока Zero-Order Hold.

Рассмотрим работу быстрого алгоритма при разных значениях шага дискретизации. При $\Delta t = 0,1$ (рис. 6) трикоптер быстро выходит на заданную высоту, однако условие (22) при этом не выполняется. При увеличении Δt (рис 7, 8) время выхода на уставку увеличивается, а максимальное значение управляющего воздействия уменьшается. В реальной системе кроме условия (22) необходимо также учитывать инерционность двигателей, которые не могут мгновенно изменять тягу. Очевидно, что с увеличением шага дискретизации влияние инерционности на качество управления уменьшается.

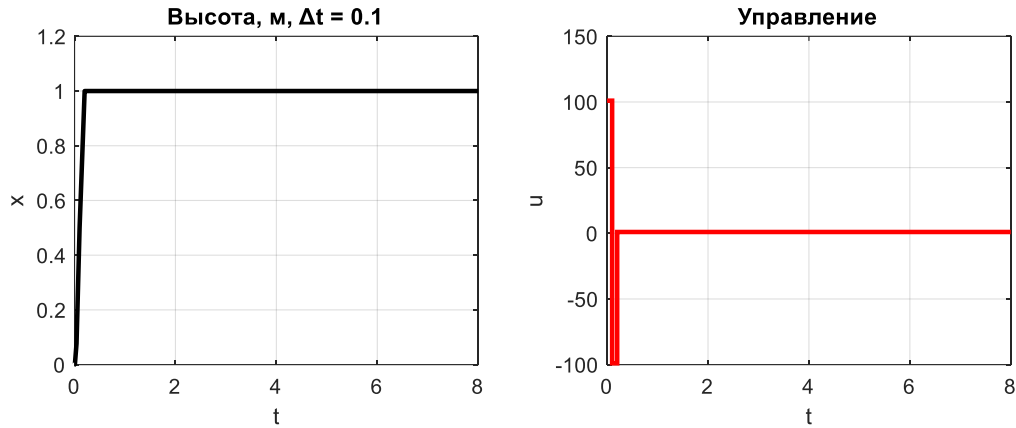


Рис. 6. Работа быстрого алгоритма при $\Delta t = 0,1$

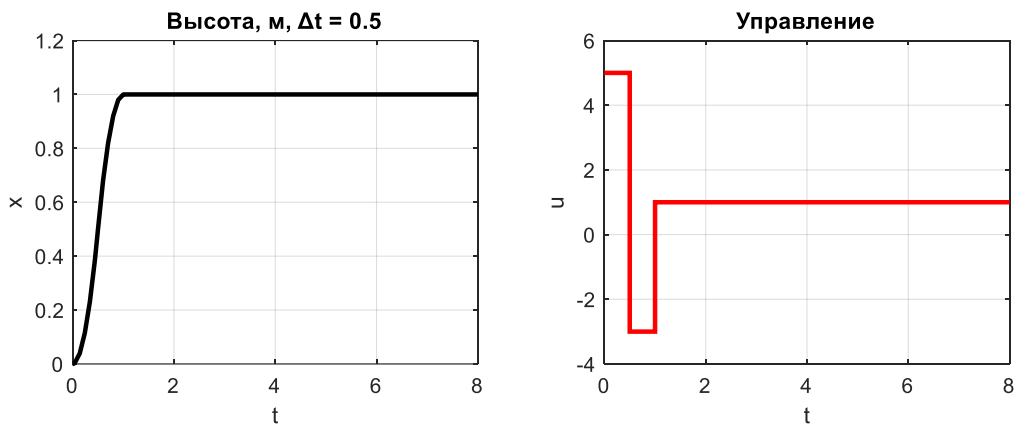


Рис. 7. Работа быстрого алгоритма при $\Delta t = 0,5$

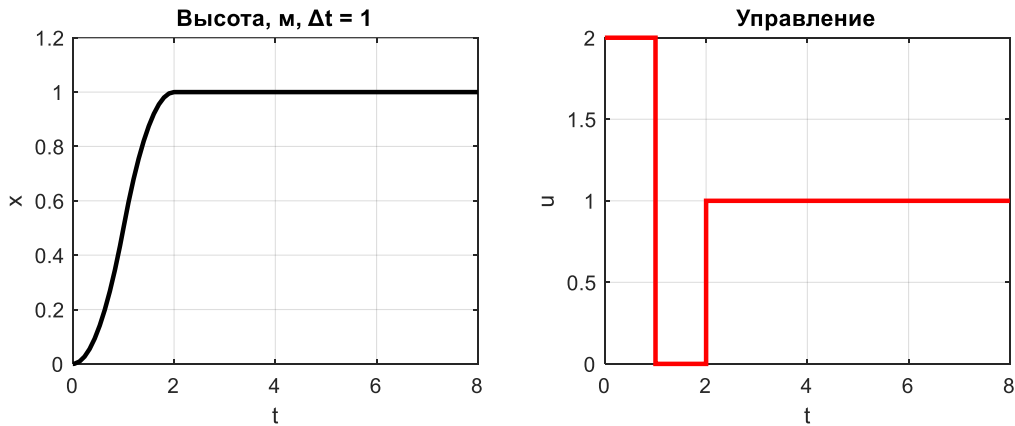


Рис. 8. Работа быстрого алгоритма при $\Delta t = 1$

При моделировании работы системы с учетом (22) и разными шагами дискретизации получены следующие результаты. Максимальная тяга двигателей (22) взята равной $4Mg$.

При малых значениях Δt (рис. 9) учет ограничения приводит к частым переключениям между максимальной и минимальной тягой, что ухудшает качество управления и является физически сложно реализуемым. При увеличении Δt (рис 10, 11) качество управления улучшается, однако время выхода на уставку остается равным времени выхода при Δt , при котором максимальная тяга не превышает ограничение.

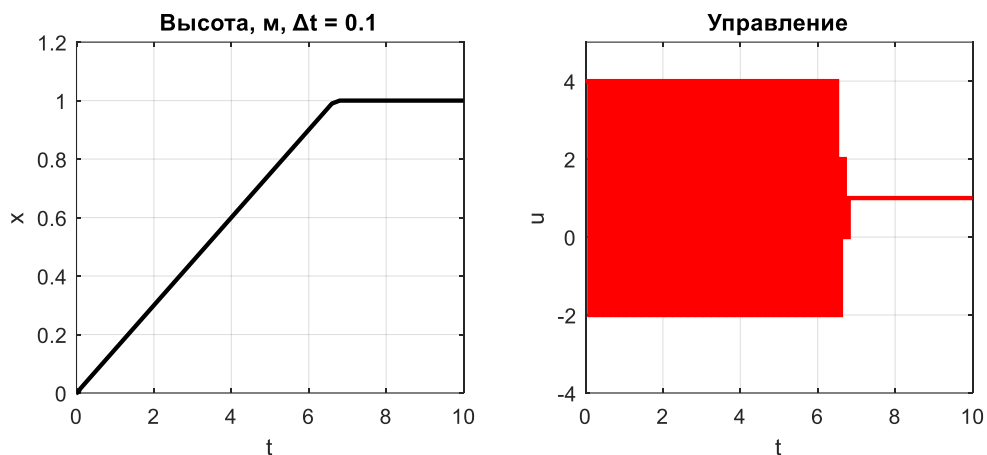


Рис. 9. Работа быстрого алгоритма с ограничением тяги при $\Delta t = 0,1$

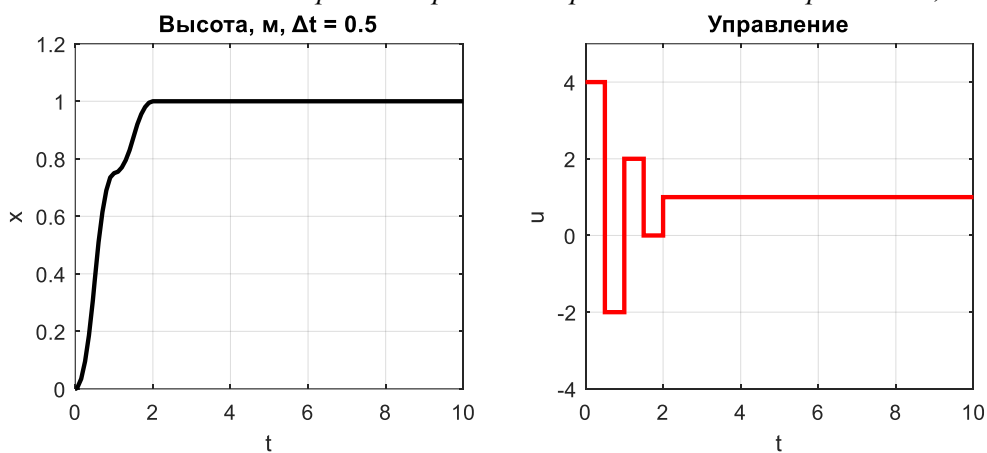


Рис. 10. Работа быстрого алгоритма с ограничением тяги при $\Delta t = 0,5$

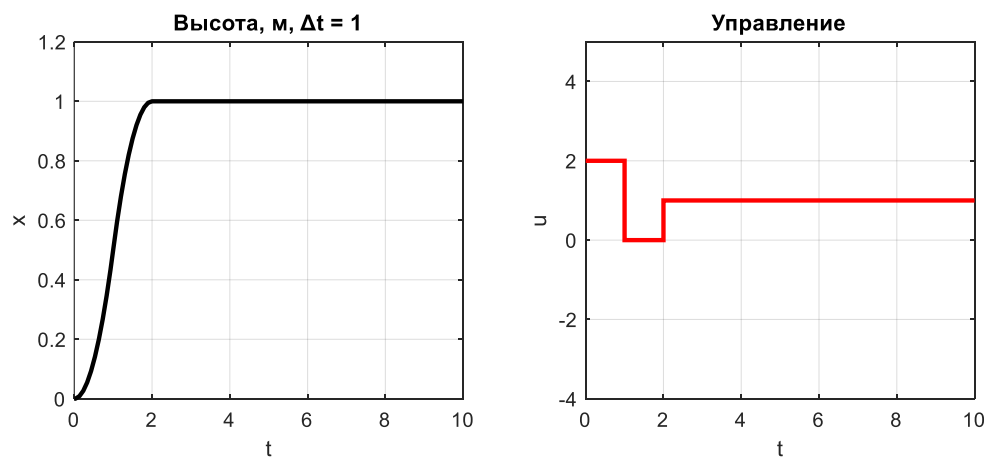


Рис. 11. Работа быстрого алгоритма с ограничением тяги при $\Delta t = 1$

3. Стабилизация высоты и положения трикоптера при нагрузке

В реальных условиях БПЛА поднимают в воздух полезный груз, параметры которого могут изменяться: аппарат может поднимать или сбрасывать его, груз может менять положение, например, при раскачивании. Для исследования поведения системы управления в подобных сценариях были рассмотрены три случая: груз с фиксирован-

ными параметрами, расположенный со смещением относительно центра масс трикоптера, груз, меняющий положение, и груз, меняющий массу.

На рисунке 12 представлены результаты моделирования подъема трикоптера с использованием ПИД-регуляторов. Параметры нагрузки постоянны:

$$m = 0,1 \text{ кг}; \alpha = 22,5^\circ; \rho = 0,1 \text{ м}$$

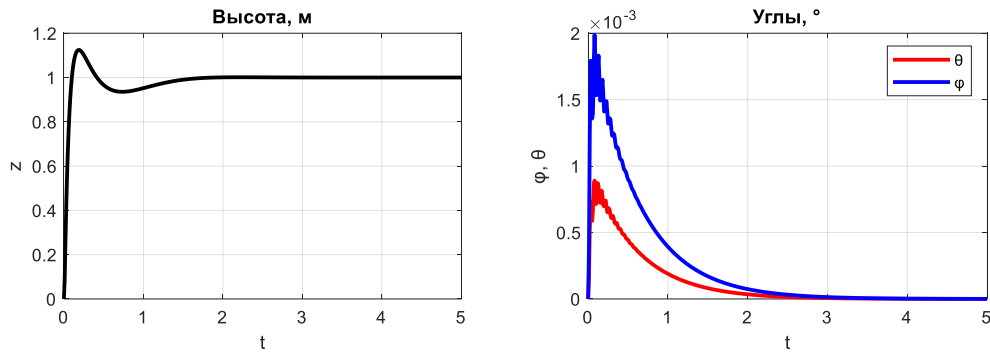


Рис. 12. Результаты моделирования подъема при постоянной асимметричной нагрузке

На рисунке 13 представлены результаты моделирования подъема при нагрузке, меняющей положение. Параметры нагрузки:

$$m = 0,1 \text{ кг}; \alpha = 22,5^\circ; \rho = \begin{cases} 0,1 \text{ м}, & t < 2 \\ 0,2 \text{ м}, & t \geq 2 \end{cases}$$

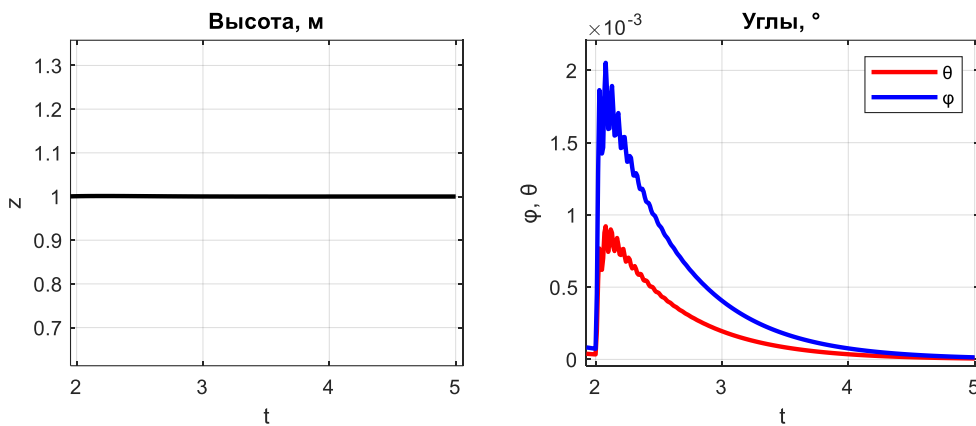


Рис. 13. Результаты моделирования подъема при изменяющемся положении нагрузки

На рисунке 14 представлены результаты моделирования подъема при изменении массы, имитирующем сброс нагрузки. Параметры нагрузки:

$$\rho = 0,1 \text{ м}; \alpha = 22,5^\circ; m = \begin{cases} 0,2 \text{ кг}, & t < 2 \\ 0 \text{ кг}, & t \geq 2 \end{cases}$$

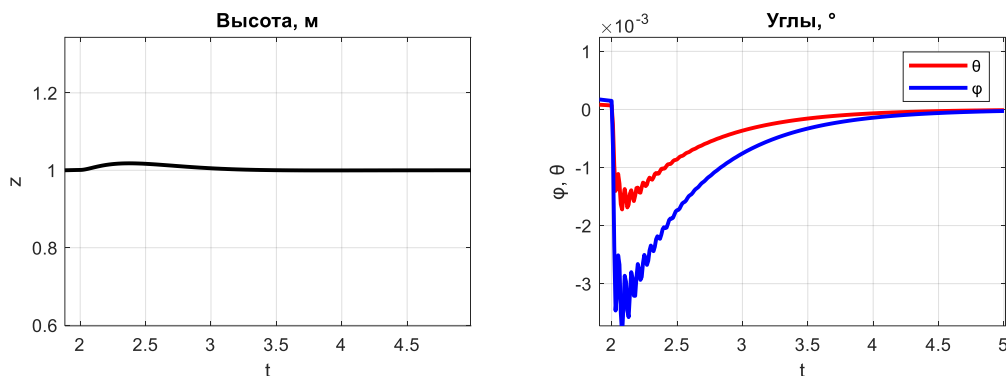


Рис. 14. Результаты моделирования при изменяющейся массе нагрузки

Заключение

Рассмотрена задача управления подъемом и горизонтальной стабилизации беспилотного трикоптера с центрально-симметричным расположением двигателей. В математической модели трикоптера учтено наличие нагрузки с параметрами, меняющимися со временем.

Представлен алгоритм управления подъемом и горизонтальной стабилизацией, основанный на дискретном ПИД-регуляторе. Разработан алгоритм управления подъемом, основанный на быстром алгоритме 2-го порядка. Представлена S-модель управления, построенная в пакете MATLAB/Simulink, и результаты моделирования алгоритмов управления.

Для быстрого алгоритма исследовано поведение системы при разных шагах дискретизации. Определено, что алгоритм является наиболее работоспособным при учете физического ограничения на максимальную тягу при выборе шага дискретизации и продемонстрирована работа алгоритма при разных его значениях, как без учета ограничения тяги, так и с учетом. Для большей робастности системы можно рекомендовать переключаться на другой алгоритм управления, например ПИД, после выхода на уставку.

Дальнейшим направлением исследований может стать изучение робастности быстрого алгоритма для случаев, когда параметры системы не определены точно. Также возможно учитывать при моделировании большее количество физических свойств трикоптера, таких, как инерционность тяговых двигателей, нелинейности при больших отклонениях по горизонтали и т. д.

Список литературы

1. *Формирование структуры и разработка алгоритмов системы автоматического управления БЛА конвертопланного типа* / Аполлонов Д. В., Артамонов М. А., Бибикина К. И., Ермаков М. В., Петрова А. А., Шибаев В. М. // XIV Всероссийское совещание по проблемам управления "ВСПУ 2024". Россия, Москва, ИПУ РАН 17-20 июня 2024 г. Электронный препринт. С. 1494–1497.
2. *Алгоритмы управления полётом беспилотных летательных аппаратов типа трикоптера с поворотными винтами* / Нгуен З. Х., Путов В. В., Шелудько В. Н. // XIV Всероссийское совещание по проблемам управления "ВСПУ 2024". Россия, Москва, ИПУ РАН 17-20 июня 2024 г. Электронный препринт. С. 1498–1502.
3. *Ощепков А. Ю.* Компьютерное моделирование робастных и адаптивных алгоритмов в технологии модельно-ориентированного проектирования / Ощепков А. Ю., Власов А. Н. // Вестник Пермского университета. Математика. Механика.

Информатика. 2024. № 64. С. 43–52. DOI: 10.17072/1993-0550-2024-1-43-52
EDN: PLYDTM

4. Деменков Н. П. Модельно-ориентированное проектирование систем управления / Деменков Н. П. // Промышленные АСУ и контроллеры. 2008. № 10. С. 66–69. EDN: KAXWLN
5. Дорф Р. Современные системы управления. / Дорф Р., Бишоп Р. М.: Лаборатория Базовых Знаний, 2002. С. 779–780.
6. Мирошник И. В. Теория автоматического управления. Линейные системы / Мирошник И.В. СПб.: Питер, 2005. 336 с. ISBN: 5-469-00350-7 EDN: QMOQGZ
7. Андриевский Б. Р. Избранные главы теории автоматического управления с примерами на языке MATLAB / Б. Р. Андриевский, А. Л. Фрадков. СПб.: Наука, 1999. 467 с.
8. Ощепков А. Ю. Универсальные быстрые алгоритмы управления для объектов с инерцией / Ощепков А. Ю., Исаков К. А. // Приборы и системы. Управление, контроль, диагностика. 2009. № 3. С. 10–16. EDN: KPZDNX

References

1. Apollonov, D. V., Artamonov, M. A., Bibikova, K. I., Ermakov, M. V., Petrova A. A. and Shibaev V. M. (2024), "Formation of the structure and development of algorithms for the automatic control system of a convertiplane UAV", *XIV Vserossiyskoe soveshchanie po problemam upravleniya* [XIV All-Russian Conference on Management Problems], *XIV Vserossiyskoe soveshchanie po problemam upravleniya* [XIV All-Russian Conference on Management Problems], Moscow, Russia, 17-20 June 2024, pp. 1494-1497.
2. Nguen, Z. H., Putov, V. V. and Shelud'ko, V. N. (2024), "Flight control algorithms for unmanned aerial vehicles of the tricopter type with rotary propellers", *XIV Vserossiyskoe soveshchanie po problemam upravleniya* [XIV All-Russian Conference on Management Problems], *XIV Vserossiyskoe soveshchanie po problemam upravleniya* [XIV All-Russian Conference on Management Problems], Moscow, Russia, 17-20 June 2024, pp. 1498-1502.
3. Oshchepkov, A. Yu. and Vlasov, A. N. (2024), "Computer modeling of robust and adaptive algorithms in model-based design technology", *Vestnik Permskogo universiteta. Matematika. Mekhanika. Informatika*. vol 64, pp. 43-52.
4. Demenkov, N. P. (2008), "Model-Based Design of Control Systems", *Promyshlennye ASU i kontrollery*, vol. 10, pp. 66–69.
5. Dorf, R. and Bishop, R. (2002), *Sovremennye sistemy upravleniya* [Modern Control Systems], Translated by Kopylov B. I., Laboratoriya Bazovyh Znaniy, Moscow, Russia, pp 779-780.
6. Miroshnik, I. V. (2005), *Teoriya avtomaticheskogo upravleniya. Linejnye sistemy* [Automatic Control Systems. Linear Systems], Saint-Petersburg, Piter, Russia, p. 336.
7. Andrievskij, B. R and Fradkov, A. L. (1999), *Izbrannye glavy teorii avtomaticheskogo upravleniya s primerami na yazyke MATLAB* [Selected Chapters of Automatic Control Theory with examples in MATLAB language]. Saint-Petersburg, Nauka, Russia, p. 467.
8. Oshchepkov, A. Yu. and Isakov, K. A. (2009), "Universal Fast Control Algorithms for Objects with Inertia", *Pribory i sistemy. Upravlenie, kontrol', diagnostika*, vol. 3, pp. 10-16.

Информация об авторах:

А. Ю. Ощепков – кандидат физико-математических наук, доцент, профессор кафедры радиоэлектроники и защиты информации Пермского государственного национального исследовательского университета (614068, Россия, г. Пермь, ул. Букирева, 15), член международного общества по физике и управлению (International Physics And Control Society – IPACS), AuthorID 175649;

А. Н. Власов – аспирант физического факультета Пермского государственного национального исследовательского университета (614068, Россия, г. Пермь, ул. Букирева, 15).

Information about the authors:

A. Yu. Oshchepkov – Candidate of Physical and Mathematical Sciences, Associate Professor, Professor of the Department of Radio Electronics and Information Security, Perm State University (15 Bukireva St., Perm, 614068, Russia), member of IPACS, AuthorID 175649.

A. N. Vlasov – postgraduate student of the Department of Radio Electronics and Information Security, Perm State University (15 Bukireva St., Perm, 614068, Russia).

Научная статья

УДК 539.8; 53.086

DOI: 10.17072/1993-0550-2025-2-36-46

<https://elibrary.ru/rtdefj>



Исследование микроструктуры и локальных механических свойств растянутых полиуретанов

Илья Александрович Морозов¹, Антон Юрьевич Беляев²

¹Институт механики сплошных сред УрО РАН, Пермь, Россия

imorozov@icmm.ru

²Пермский государственный национальный исследовательский университет, Пермь, Россия

belyaev@icmm.ru

Аннотация. Эластичный полиуретан – синтетический эластомер, состоящий из жестких и мягких блоков, образующих в материале неоднородные надмолекулярные структуры. В работе методами атомно-силовой микроскопии исследована микроструктура и жесткость двух полиуретанов, отличающихся плотностью надмолекулярной жесткой фибриллярной сетки, вплетенной в более мягкую матрицу. Показана эволюция локальных свойств при одноосном растяжении, в том числе в предельно растянутом состоянии. Методом конечных элементов моделировали вдавливание индентора в растянутый упругий полимер; получена зависимость жесткости от кратности удлинения. Этот результат использован для сравнения с экспериментальными данными и оценки величины локального удлинения растянутого полиуретана. Полученные результаты объясняют и дополняют макроскопические механические свойства полимеров.

Ключевые слова: полиуретан; растяжение; атомно-силовая микроскопия; модуль упругости; индентирование; метод конечных элементов

Для цитирования: Морозов И. А., Беляев А. Ю. Исследование микроструктуры и локальных механических свойств растянутых полиуретанов // Вестник Пермского университета. Математика. Механика. Информатика. 2025. № 2(69). С. 36–46. DOI: 10.17072/1993-0550-2025-2-36-46. <https://elibrary.ru/rtdefj>

Благодарности: работа выполнена в рамках госбюджетной программы АААА-А20-120022590044-7.

Статья поступила в редакцию 16.01.2025; одобрена после рецензирования 08.06.2025; принята к публикации 09.07.2025.

Research article



Эта работа © 2025 Морозов И. А., Беляев А. Ю. распространяется по лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>.

Study of Stretched Polyurethanes Microstructure and Local Mechanical Properties

Илья. А. Морозов¹, Антон. Ю. Белиаев²

¹Institute of Continuous Media Mechanics, Perm, Russia

imorozov@icmm.ru

²Perm State University, Perm, Russia

belyaev@icmm.ru

Abstract. Elastic polyurethane is a synthetic elastomer consisting of hard and soft blocks forming inhomogeneous supramolecular structures in the material. In this work, the microstructure and stiffness of two polyurethanes differing in the density of supramolecular stiff fibrillar mesh woven into a softer matrix have been investigated by atomic force microscopy. The evolution of local properties under uniaxial tension, including the ultimate tensile state, is shown. The indentation of a stretched elastic polymer was modeled by the finite element method; the dependence of stiffness on the elongation ratio was obtained. This result was used for comparison with experimental data and estimation of local elongation of the stretched polyurethane. The obtained results explain and expand the macroscopic mechanical properties of the polymers.

Keywords: polyurethane; elongation; atomic force microscopy; elastic modulus; indentation; finite element method

For citation: Morozov, I. A. and Beliaev, A. Yu. (2025), "Study of microstructure and local mechanical properties of stretched polyurethanes", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 36–46. (In Russ.). DOI: 10.17072/1993-0550-2025-2-36-46. <https://elibrary.ru/rtdefj>

Acknowledgments: the work was carried out within the framework of the state budget program AAAA-A20-120022590044-7.

The article was submitted 16.01.2025; approved after reviewing 08.06.2025; accepted for publication 09.07.2025

Введение

Полиуретановые эластомеры представляют собой сегментированные полимеры, состоящие из мягких и жестких блоков [1]. Первичные жесткие блоки формируют в более мягкой матрице вторичные надмолекулярные структуры различной конфигурации (фибриллярные, цилиндрические, ленточные, сеточные или глобулярные). Свойства таких структур зависят от химического состава [2–5], времени либо температуры отверждения [6, 7], наличия в составе наполнителя [8–10], толщины материала [11], дополнительного термического [12, 13] или механического воздействия [14, 15].

К настоящему времени известно значительное количество работ, посвященных физико-механическим характеристикам полиуретановых эластомеров. Однако зачастую интерпретация результатов с точки зрения особенностей их строения и локальных механических свойств делается на основе косвенных данных (обычно — спектроскопические исследования). Учитывая сложное строение полиуретанов, это является явно недостаточным. Интерес представляет исследование эволюции надмолекулярной структуры полиуретана и изменение локальных механических свойств в результате одноосной деформации, в том числе в предельно растянутом состоянии (вершина надреза в растянутом полимере).

В работе синтезировали полиуретаны одной рецептуры, но с различной температурой отверждения. Изменение температуры отверждения позволяет получать полиуретаны с различной плотностью надмолекулярной сетки жестких фибрилл. Целью работы являлась оценка влияния одноосной деформации на локальную жесткость и структурные особенности материалов.

Материалы и методы

Полиуретаны изготавливали из коммерчески доступных преполимера и сшивающего агента. Компоненты разогревали при 80 °С и вакуумировали. Смесь отверждали в вакууме в течение 24 часов в виде пластин со свободной верхней поверхностью толщиной 2 мм. Температуру отверждения задавали 50 °С либо 90 °С. Ранее было установлено влияние температуры отверждения на структуру и механические свойства надмолекулярной структуры полимеров [16]. Полученные образцы обозначим Т50 и Т90.

Для оценки макроскопических механических свойств образцы подвергали одноосному растяжению до разрыва на испытательной машине *Zwick*. Скорость нагружения – 10% / мин.

В работе использовали атомно-силовой микроскоп *Ntegra Prima* в режиме наномеханического картирования. Исследовали образцы как в недеформированном, так и растянутом состоянии. В последнем случае образцы растягивали до нужной кратности удлинения λ , после выдержки в 10 минут чего приклеивали к деревянной подложке; $\lambda = l / l_0$, где l и l_0 – длина растянутого и нерастянутого материала. Интерес представляют материалы в критически растянутом состоянии. Для реализации этого случая на краю растянутого образца делается надрез, возникший дефект прорастает некоторое время вглубь материала, а затем останавливается. Материал в вершине такого надреза является предельно растянутым – любое увеличение растяжения образца влечет за собой дальнейшее прорастание дефекта.

Применяли зонды с калиброванными радиусами R острия и изгибной жесткостью k балки: *ScanAsyst-Air* (индентирование слабо растянутых материалов, $R \sim 3...5$ нм, $k \sim 0.4...0.5$ нН/нм) либо *CSG30* (индентирование сильно растянутых материалов, $R \sim 7$ нм, $k \sim 1.5$ нН/нм). Высокая скорость индентации (20 нм / мсек) позволяет получать данные с высоким разрешением в плоскости образца: массив кривых взаимодействия $F(u)$ – усилия от глубины вдавливания. Модуль E упругости в эксперименте вычисляли, используя модель Джонсона–Кендалла–Роберта (ДКР) – контактное упругое взаимодействие Герца с учетом адгезии между индентором и материалом. Формула Герца имеет вид: $F(u) = 4ER^{0.5}u^{1.5}/(3(1-\nu^2))$, коэффициент Пуассона ν полагали равным 0.5. Модель ДКР [17] предполагает решение двух уравнений и в данной работе не приводится.

В работе индентуем как недеформированные, так и растянутые материалы. В последнем случае следует говорить не об упругой константе материала, а о жесткости. Чтобы не усложнять обсуждение, в тексте будем употреблять обе формулировки.

С помощью метода конечных элементов (МКЭ) решали задачу о вдавливании индентора в растянутый материал (рис. 1а). Для описания материала использовали данные из эксперимента на одноосное растяжение (см. далее рис. 2), которые аппроксимировали гиперупругим потенциалом Арруды–Бойса (АБ) [18]; выражение для напряжений в случае одноосного растяжения несжимаемого материала:

$$\sigma = 2\mu \left(\lambda - \frac{1}{\lambda^2} \right) \sum_{i=1}^5 \frac{ic_i}{\lambda^{2i-2}} I^{i-1}, \text{ где } I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \text{ в случае одноосного растяжения } \lambda_1 = \lambda; \lambda_2 = \lambda_3 = \frac{1}{\sqrt{\lambda}}, \text{ } c_i - \text{коэффициенты ряда } c_1 = \frac{1}{2}; c_2 = \frac{1}{20}; c_3 = \frac{11}{1050}; c_4 =$$

$\frac{19}{7000}$; $c_5 = \frac{519}{673750}$; ... При подборе кривой, константы материала T50 получились равными: $\mu = 3.48e-4$, $\lambda_\mu = 3.54$; для T90: $\mu = 7.735e-4$, $\lambda_\mu = 3.676$. Экспериментальные кривые одноосного растяжения и соответствующие аппроксимации представлены на рис. 2. Отметим, что помимо потенциала АБ также была исследована возможность аппроксимации потенциалами Огдена и Джента, однако, именно АБ дал наилучшее совпадение для наших материалов, показавших в эксперименте (рис. 2) плавное упрочнение.

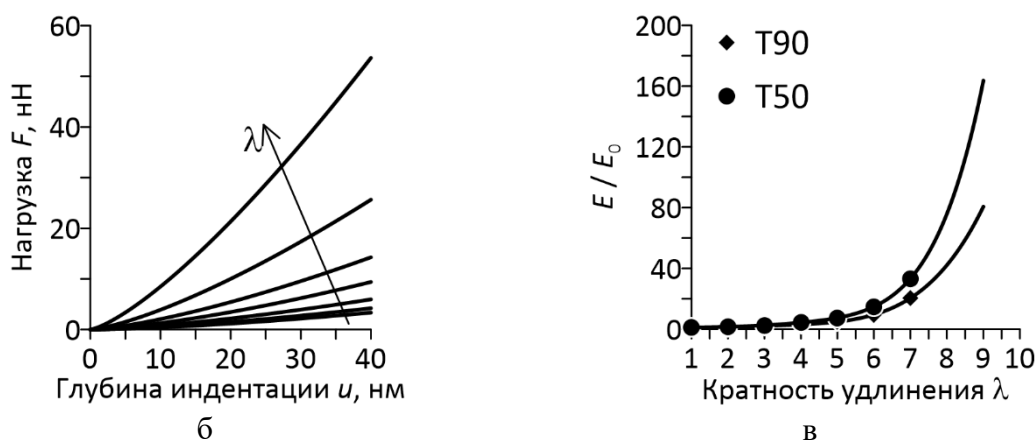
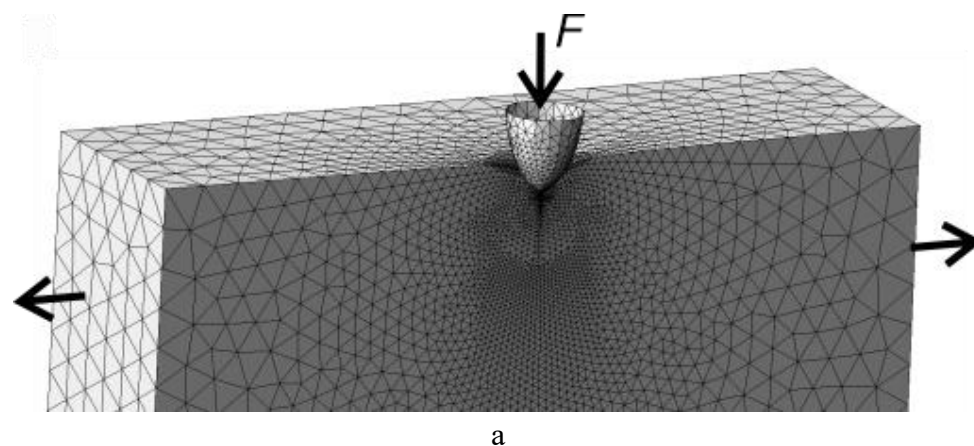


Рис. 1. Модель индентации растянутого материала (а), полученный силовой отклик для материала T90 (б) и нормированная жесткость E/E_0 (в): результаты МКЭ-вычислений (маркеры) и экстраполяция (линия) до больших кратностей удлинения

Контактная задача решалась в трехмерной постановке. На первом шаге материал растягивался до заданной кратности удлинения, на втором происходило индентирование абсолютно жестким параболическим индентором с радиусом острия 10 нм.

В результате был получен набор силовых кривых $F(u)$ в зависимости от кратности удлинения λ модельного материала (рис. 1б): $\lambda = 1 \dots 7$. Модельные силовые кривые аппроксимировали моделью Герца, определяя жесткость (модуль) E материала. Для сравнения экспериментальных данных с результатами конечно-элементного моделирования используем нормированную жесткость $S = E/E_0$, где E, E_0 – значения для растянутого и нерастянутого материалов. Нормированная жесткость отражает увеличение жесткости полимера в зависимости от кратности удлинения (рис. 1в).

Получив аналогичное отношение E/E_0 из эксперимента (для экспериментальных кривых использовали модель ДКР) и, сравнив с калибровочной кривой (рис. 1в), можно получить оценку локальной кратности удлинения поверхности.

Результаты и их обсуждение

По результатам испытаний образцов на одноосное механическое растяжение (рис. 2) можно установить, что материал T50 более жесткий и разрушается при меньшем растяжении, чем T90.

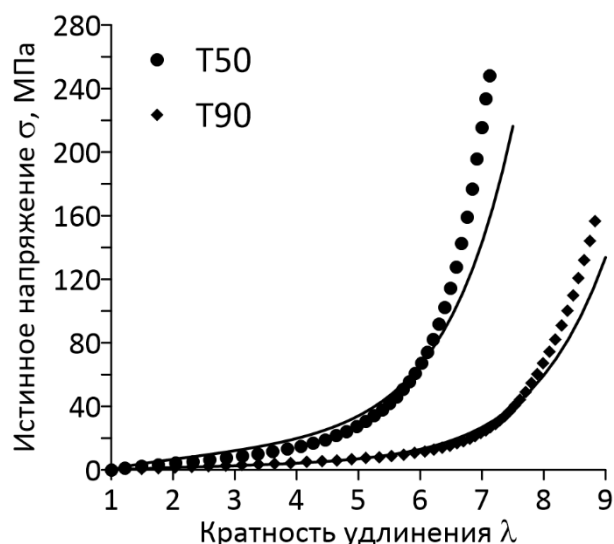


Рис. 2. Кривые одноосного растяжения до разрыва: экспериментальные данные (маркеры) и аппроксимация потенциалом Арруди–Бойса

Неоднородная структура синтезированных полиуретанов (рис. 3) представляет собой жесткие фибриллярные структуры (надмолекулярные образования с высокой концентрацией жестких блоков полиуретана), неоднородно переплетенные в более мягкой матрице (области высокой концентрации мягких блоков).

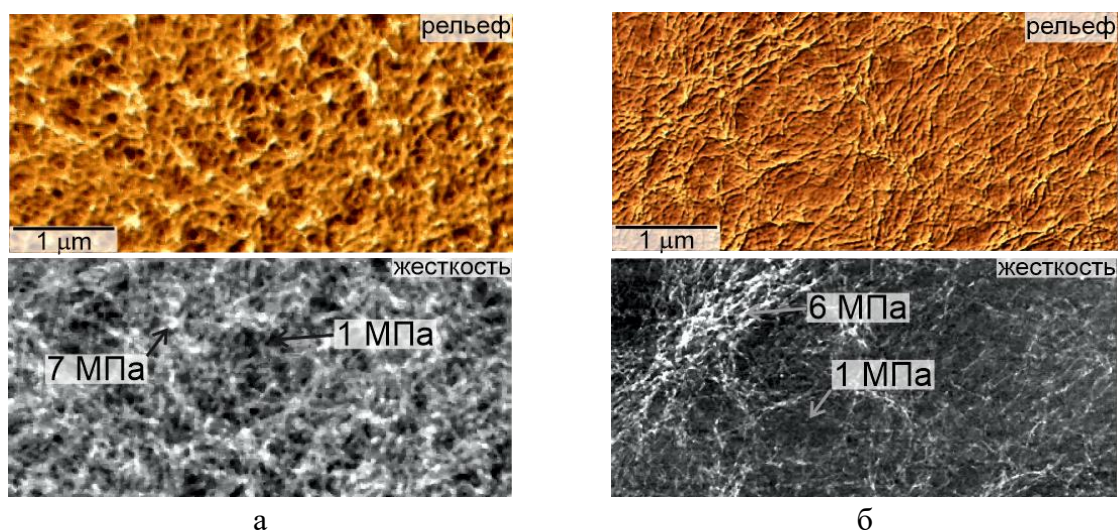


Рис. 3. АСМ-изображения рельефа поверхности и карты жесткости полиуретана T50 (а) и T90 (б)

Надмолекулярная структура материала, синтезированного при температуре 50 °С, представляет собой плотную сетку (рис. 3а) жестких фибрилл, переплетенных в более мягкой матрице. Структура материала Т90 (рис. 3б) более неоднородная: наблюдаются как отдельные протяженные фибриллы, так и их плотные агломераты. Длина фибрилл может достигать 1 мкм. Средняя ширина фибриллы 3...5 нм (точность измерений ограничена радиусом скругления зонда ~2 нм), что соответствует перпендикулярно упакованным жестким блокам полиуретана.

Модуль упругости жесткой фибриллярной структуры (рис. 3, нижний ряд) составил 7 МПа для материала Т50 или 6 МПа для материала Т90. Модуль упругости "матрицы" – 1 МПа. Сравнение локальных структурно-механических свойств (рис. 3) с испытаниями на одноосное растяжение (рис. 2) показало, что высокая плотность жестких надмолекулярных структур увеличивает напряжения в материале, а также снижает его разрывную деформацию, по сравнению с полиуретаном, синтезированным при большей температуре и обладающим менее плотной надмолекулярной сеткой.

Структура и жесткость поверхностей изменяются в процессе растяжения. При кратностях удлинения до 4 (рис 4а, б) жесткие блоки материала Т50 становятся более мелкозернистыми, ориентированных структур не наблюдается. То есть, происходит разделение плотной фибриллярной сетки на более мелкие фрагменты. При кратностях удлинения свыше 6 (рис. 4в, г) происходит слабо выраженная ориентация жестких структур; рельеф поверхности становится более неоднородным, но жесткость поверхности качественно не изменяется. Отметим, что согласно результатам инфракрасной спектроскопии [19] при кратностях удлинения $\lambda > 5$ происходит снижение интенсивности спектра жестких структур; это связывают с их перераспределением в матрице мягкой фазы.

Сравнение отношения локального модуля упругости растянутого полиуретана (отмечено на рис. 4в) к среднему модулю недеформированного материала (4 МПа, рис. 3) и сопоставление с калибровочной кривой (рис. 1в) позволило оценить локальную кратность удлинения. Результаты показали, что для данного материала локальная кратность удлинения примерно равна макроскопической.

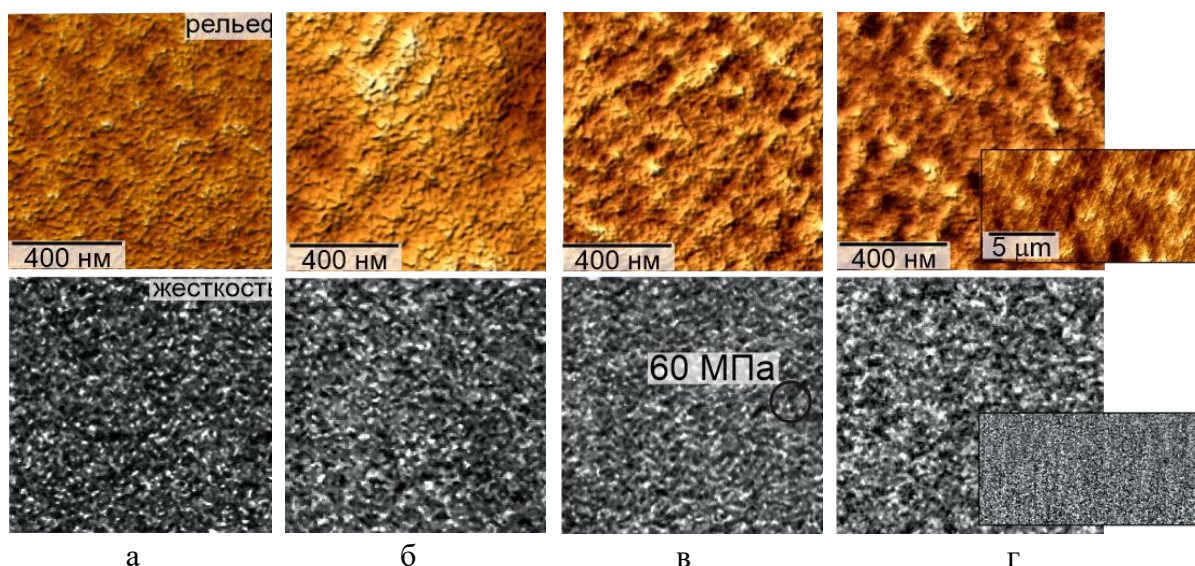


Рис. 4. Детальные АСМ-изображения рельефа (верхний ряд) и жесткости (нижний ряд) материала Т50 растянутого в 2 (а), 4 (б), 6 (в) и 8 (г). Для восьмикратного удлинения представлены также и обзорные изображения

Более существенные изменения происходят с микроструктурой материала Т90 (рис. 5). При растяжении до четырех раз (рис. 5а, б) сетка длинных фибриллярных структур ориентируется вдоль оси деформации. С ростом деформации жесткие фибриллы разрушаются и при шестикратном растяжении (рис. 5в) на поверхности видны лишь одиночные протяженные вытянутые жесткие блоки. При восьмикратном растяжении (рис. 5г) фибриллярные структуры полностью разрушены; структура поверхности напоминает растянутый в 6...8 раз материал Т50 (рис. 4в, г).

Сравнение отношения локального модуля упругости растянутого полиуретана Т90 (отдельные значения отмечены на рис. 5) к модулю недеформированного материала (рис. 3) и сопоставление с калибровочными кривыми (рис. 1в) позволило оценить локальную кратность удлинения. Так, при шестикратной макроскопической кратности удлинения, локальная λ фибриллярных структур составила 7...8. Более точная оценка затруднена в силу структурно-механических неоднородностей трехмерной структуры исходного полимера.

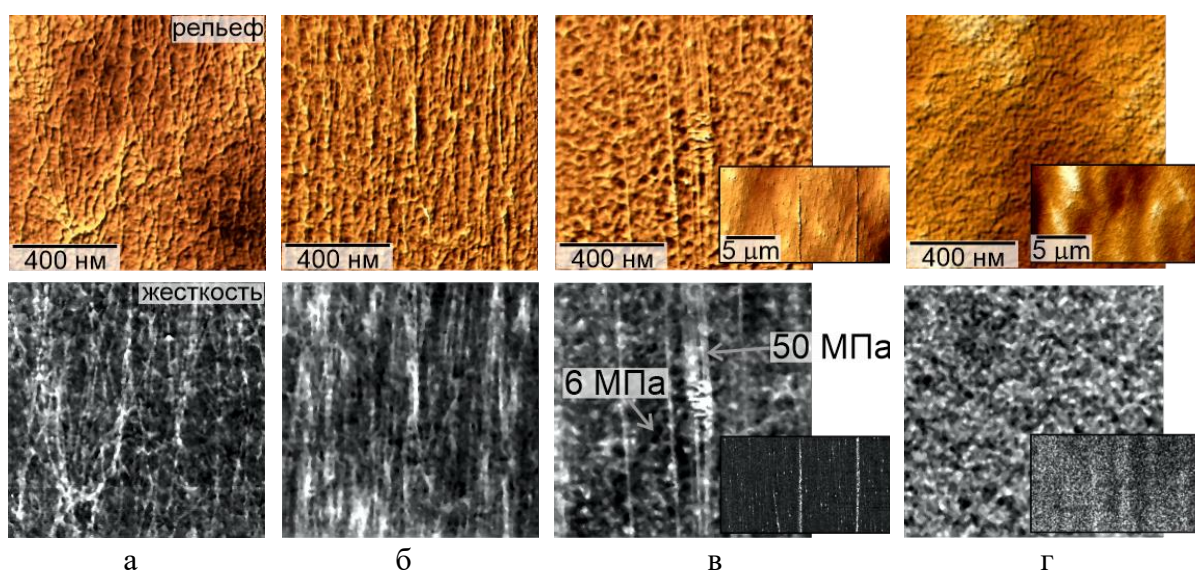


Рис. 5. Детальные АСМ-изображения рельефа (верхний ряд) и жесткости (нижний ряд) материала Т90, растянутого в 2 (а), 4 (б), 6 (в) и 8 (г). Для шести- и восьмикратного удлинений представлены обзорные изображения

Материал в вершине раскрытого надреза находится в предельно растянутом состоянии. Исследование такой поверхности (рис. 6а), показало, что от краев надрыва к его оси идут ориентированные структуры (рис. 6б, в). Образование сетки тяжей не связано с двухфазной структурой полиуретана, а обусловлено высокой вытяжкой полимера и локальными микроразрывами. Плотность такой ориентированной сетки тяжей выше в материале Т50 (рис. 6б), чем в Т90 (рис. 6в). Жесткость в окрестности оси разрыва максимальна, а затем резко снижается. Предельные значения жесткости для обоих материалов совпали. Ширина перенапряженной области максимальной жесткости отличается и составила 1 мкм (для Т50) или 2 мкм (для Т90). Это говорит о более плавном распределении нагрузки в материале Т90 за счет большей концентрации структур, насыщенных мягкими блоками.

Жесткость в вершине трещины (рис. 6) в 45 раз превышает значение для недеформированного полимера (за модуль недеформированного полимера принимаем среднее арифметическое модуля мягкой и жесткой фазы), что, согласно калибровочной кривой нормированной жесткости, соответствует 8...9-кратному локальному удлинению

полимера Т90 в вершине разрыва. Такая локальная кратность удлинения в трещине сопоставима с макроскопическим разрывным удлинением материала Т90; в то же время локальное предельное удлинение в материале Т50 на одну единицу превышает макроскопическое (рис. 2). Последнее можно объяснить тем, что в этом полиуретане имеются структурно-механические дефекты, являющиеся очагами развития разрушения. Такими дефектами могут быть области с высокой концентрацией жестких блоков.

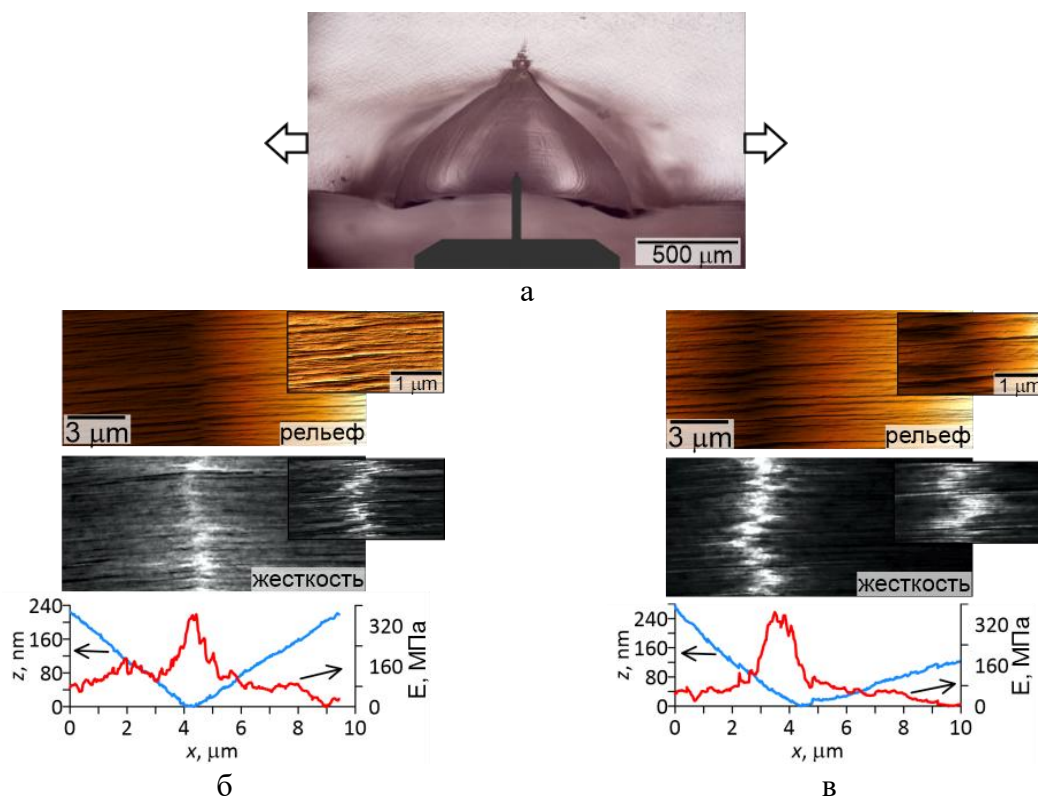


Рис. 6. Оптическое изображение зонда АСМ в раскрытом надрезе растянутого полиуретана (а); детальные и обзорные АСМ-изображения поверхностей Т50 (б) и Т90 (в) в окрестности вершины надреза – показаны рельеф, жесткость, а также их характерные профили

Заключение

Было проведено исследование структурно механических особенностей надмолекулярной структуры полиуретановых эластомеров, в результате чего были сделаны следующие выводы: варьирование температуры отверждения позволяет получать материалы с различной плотностью сетки более жестких фибриллярных структур, вплетенных в более мягкую матрицу. Материал Т50, синтезированный при температуре 50 °С, обладает более жесткой и однородной структурой фибрилл; полиуретан, отвержденный при 90°С (Т90), дает неоднородную надмолекулярную структуру протяженных фибрилл (до нескольких микрон) и агломератов.

Методом конечных элементов было исследовано индентирование растянутого эластомера. Получена зависимость нормированной жесткости поверхности от кратности удлинения. Эта калибровочная кривая использована при обработке экспериментальных данных. Экспериментальное исследование материалов в растянутом состоянии показало разделение фибриллярной структуры на отдельные жесткие и мягкие сегменты, а также образование единичных, ориентированных вдоль оси деформации, структур, локальная кратность удлинения в которых превышает макроскопическую кратность удлинения. В вершине разрыва локальная кратность удлинения достигает 8...9, что сопоставимо

(материал Т90) или превышает (для материала Т50) макроскопическое разрывное удлинение.

Список источников

1. Petrović Z. S., Ferguson J. Polyurethane elastomers // *Progress in Polymer Science*. 1991. Vol. 16. P. 695–836. DOI:10.1016/0079-6700(91)90011-9 EDN:XSOOJE.
2. Sheth J. P., Klinedinst D. B., Wilkes G. L., Yilgor I., Yilgor E. Role of chain symmetry and hydrogen bonding in segmented copolymers with monodisperse hard segments // *Polymer*. 2005. Vol. 46. P. 7317–7322. DOI:10.1016/j.polymer.2005.04.041 EDN:KGQHRX.
3. Klinedinst D. B., Yilgor I., Yilgor E., Zhang M., Wilkes G. L. The effect of varying soft and hard segment length on the structure-property relationships of segmented polyurethanes based on a linear symmetric diisocyanate, 1,4-butanediol and PTMO soft segments // *Polymer*. 2012. Vol. 53. P. 5358–5366. DOI:10.1016/j.polymer.2012.08.005.
4. Tocha E., Janik H., Debowski M., Vancso G. J. Morphology of polyurethanes revisited by complementary AFM and TEM // *Journal of Macromolecular Science, Part B*. 2002. Vol. 41. P. 1291–1304. DOI:10.1081/MB-120013098.
5. Yilgor I., Yilgor E., Wilkes G. L. Critical parameters in designing segmented polyurethanes and their effect on morphology and properties: A comprehensive review // *Polymer*. 2015. Vol. 58. P. A1–A36.
6. Kimball M. E., Fielding-Russell G. S. Effect of cure temperature on urethane networks // *Polymer*. 1977. Vol. 18. P. 777–780. DOI:10.1016/0032-3861(77)90180-X.
7. Cheng B.-X., Gao W.-C., Ren X.-M., Ouyang X.-Y., Zhao Y., Zhao H., Wu W., Huang C.-X., Liu Y., Liu X.-Y., Li H.-N., Li R. K. Y. A review of microphase separation of polyurethane: Characterization and applications // *Polymer Testing*. 2022. Vol. 107. P. 107489. DOI:10.1016/j.polymertesting.2022.107489 EDN: DCHRZG.
8. Zheng J., Ozisik R., Siegel R. W. Disruption of self-assembly and altered mechanical behavior in polyurethane/zinc oxide nanocomposites // *Polymer*. 2005. Vol. 46. P. 10873–10882. DOI:10.1016/j.polymer.2005.08.082 EDN: KGQCDR.
9. Zheng J., Ozisik R., Siegel R. W. Phase separation and mechanical responses of polyurethane nanocomposites // *Polymer*. 2006. Vol. 47. P. 7786–7794. DOI:10.1016/j.polymer.2006.08.068 EDN: KGQUDJ.
10. Larraza I., Alonso-Lerma B., Gonzalez K., Gabilondo N., Perez-Jimenez R., Corcuera M. A., Arbelaiz A., Eceiza A. Waterborne polyurethane and graphene/graphene oxide-based nanocomposites: Reinforcement and electrical conductivity // *Express Polymer Letters*. 2020. Vol. 14. P. 1018–1033. DOI:10.3144/expresspolymlett.2020.83 EDN:QLWLGR.
11. Kojio K., Kugumiya S., Uchiba Y., Nishino Y., Furukawa M. The Microphase-separated Structure of Polyurethane Bulk and Thin Films // *Polymer Journal*. 2009. Vol. 41 P. 118–124. DOI:10.1295/polymj.PJ2008186.
12. Takahashi A., Kita R., Kaibara M. Effects of thermal annealing of segmented-polyurethane on surface properties, structure and antithrombogenicity // *Journal of Materials Science: Materials in Medicine*. 2002. Vol. 13. P. 259–262. DOI:10.1023/A:1014054716444 EDN: AUXNEJ.
13. Li X., Lu Y., Wang H., Pösel E., Eling B., Men Y. Crystallization of hard segments in MDI/BD-based polyurethanes deformed at elevated temperature and their dependence on the MDI/BD content // *European Polymer Journal*. 2017. Vol. 97. P. 423–436.

14. Christenson E. M., Anderson J. M., Hiltner A., Baer E. Relationship between nanoscale deformation processes and elastic behavior of polyurethane elastomers // *Polymer*. 2005. Vol. 46. P. 11744–11754. DOI:10.1016/j.polymer.2005.08.083 EDN: KGQARP.
15. Scetta G., Euchler E., Ju J., Selles N., Heuillet P., Ciccotti M., Creton C. Self-Organization at the Crack Tip of Fatigue-Resistant Thermoplastic Polyurethane Elastomers // *Macromolecules*. 2021. Vol. 54. P. 8726–8737. DOI:10.1021/acs.macromol.1c00934 EDN: NRMZLQ.
16. Morozov I. A., Beliaev A. Yu., Scherban M. G. Effect of curing temperature on surface and subsurface properties of polyurethane elastomer // *Materials Physics and Mechanics*. 2023. Vol. 51. P. 107–118. DOI:10.18149/MPM.5162023_10 EDN: DWRZRT.
17. Johnson K. L., Kendall K., Roberts A. D. Surface energy and the contact of elastic solids // *Proc. R. Soc. Lond. A*. 1971. Vol. 324. P. 301–313. DOI:10.1098/rspa.1971.0141.
18. Arruda E. M., Boyce M. C. A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials // *Journal of the Mechanics and Physics of Solids*. 1993. V. 41. N. 2. P. 389–412. DOI:10.1016/0022-5096(93)90013-6.
19. Tereshatov V. V., Makarova M. A., Senichev V. Yu., Slobodinyuk A. I. Interrelationship between ultimate mechanical properties of variously structured polyurethanes and poly(urethane urea)s and stretching rate thereof // *Colloid Polym Sci*. 2012. Vol. 290. P. 641–651. DOI:10.1007/s00396-011-2585-7 EDN:PDTBQP.

References

1. Petrović, Z. S. and Ferguson, J. (1991), "Polyurethane elastomers", *Progress in Polymer Science*, vol. 16, pp. 695–836.
2. Sheth, J. P., Klinedinst, D. B., Wilkes, G. L., Yilgor, I. and Yilgor, E. (2005), "Role of chain symmetry and hydrogen bonding in segmented copolymers with monodisperse hard segments", *Polymer*, vol. 46, pp. 7317–7322.
3. Klinedinst, D. B., Yilgör, I., Yilgör, E., Zhang, M. and Wilkes, G. L. (2012), "The effect of varying soft and hard segment length on the structure–property relationships of segmented polyurethanes based on a linear symmetric diisocyanate, 1,4-butanediol and PTMO soft segments", *Polymer*, vol. 53, pp. 5358–5366.
4. Tocha, E., Janik, H., Debowski, M. and Vancso, G. J. (2002), "Morphology of polyurethanes revisited by complementary AFM and TEM", *Journal of Macromolecular Science, Part B*, vol. 41, pp. 1291–1304.
5. Yilgör, I., Yilgör, E. and Wilkes, G. L. (2015), "Critical parameters in designing segmented polyurethanes and their effect on morphology and properties: A comprehensive review", *Polymer*, vol. 58, pp. A1–A36.
6. Kimball, M. E. and Fielding-Russell, G. S. (1977), "Effect of cure temperature on urethane networks", *Polymer*, vol. 18, pp. 777–780.
7. Cheng, B.-X., Gao, W.-C., Ren, X.-M., Ouyang, X.-Y., Zhao, Y., Zhao, H., Wu, W., Huang, C.-X., Liu, Y., Liu, X.-Y., Li, H.-N. and Li, R. K. Y. (2022), "A review of microphase separation of polyurethane: Characterization and applications", *Polymer Testing*, vol. 107, pp. 107489.
8. Zheng, J., Ozisik, R. and Siegel, R. W. (2005), "Disruption of self-assembly and altered mechanical behavior in polyurethane/zinc oxide nanocomposites", *Polymer*, vol. 46, pp. 10873–10882.

9. Zheng, J., Ozisik, R. and Siegel, R. W. (2006) "Phase separation and mechanical responses of polyurethane nanocomposites", *Polymer*, vol. 47, pp. 7786–7794.
10. Larraza, I., Alonso-Lerma, B., Gonzalez, K., Gabilondo, N., Perez-Jimenez, R., Corcuera, M. A., Arbelaiz, A. and Eceiza, A. (2020), "Waterborne polyurethane and graphene/graphene oxide-based nanocomposites: Reinforcement and electrical conductivity", *Express Polymer Letters*, vol. 14, pp. 1018–1033.
11. Kojio, K., Kugumiya, S., Uchiba, Y., Nishino, Y. and Furukawa, M. (2009), "The Microphase-separated Structure of Polyurethane Bulk and Thin Films", *Polymer Journal*, vol. 41, pp. 118–124.
12. Takahashi, A., Kita, R. and Kaibara, M. (2002), "Effects of thermal annealing of segmented-polyurethane on surface properties, structure and antithrombogenicity", *Journal of Materials Science: Materials in Medicine*, vol. 13, pp. 259–262.
13. Li, X., Lu, Y., Wang, H., Pösel, E., Eling, B. and Men, Y. (2017), "Crystallization of hard segments in MDI/BD-based polyurethanes deformed at elevated temperature and their dependence on the MDI/BD content", *European Polymer Journal*, vol. 97, pp. 423–436.
14. Christenson, E. M., Anderson, J. M., Hiltner, A. and Baer, E. (2005), "Relationship between nanoscale deformation processes and elastic behavior of polyurethane elastomers", *Polymer*, vol. 46, pp. 11744–11754.
15. Scetta, G., Euchler, E., Ju, J., Selles, N., Heuillet, P., Ciccotti, M. and Creton, C. (2021), "Self-Organization at the Crack Tip of Fatigue-Resistant Thermoplastic Polyurethane Elastomers", *Macromolecules*, vol. 54, pp. 8726–8737.
16. Morozov, I. A., Beliaev, A. Yu. and Scherban, M. G. (2023), "Effect of curing temperature on surface and subsurface properties of polyurethane elastomer", *Materials Physics and Mechanics*, vol. 51, pp. 107–118.
17. Johnson, K. L., Kendall, K. and Roberts, A. D. (1971), "Surface energy and the contact of elastic solids", *Proc. R. Soc. Lond. A.*, vol. 324, pp. 301–313.
18. Arruda, E. M. and Boyce, M. C. (1993), "A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials", *Journal of the Mechanics and Physics of Solids*, vol. 41, pp. 389–412.
19. Tereshatov, V. V., Makarova, M. A., Senichev, V. Yu. and Slobodinyuk, A. I. (2012), "Interrelationship between ultimate mechanical properties of variously structured polyurethanes and poly(urethane urea)s and stretching rate thereof", *Colloid Polym Sci.*, vol. 290, pp. 641–651.

Информация об авторах:

И. А. Морозов – кандидат физико-математических наук, старший научный сотрудник, Институт механики сплошных сред УрО РАН (614013, Россия, Пермский край, г. Пермь, ул. Академика Королёва, д. 1), AuthorID 172118, ORCID: 0000-0001-6395-4301;

А. Ю. Беляев – младший научный сотрудник, Пермский государственный национальный исследовательский университет (614068, Россия, Пермский край, г. Пермь, ул. Букирева, 15), AuthorID 584106, ORCID 0000-0002-0966-5979.

Information about the authors:

I. A. Morozov – Candidate of Physics and Mathematics, Senior Researcher, Institute of Continuous Media Mechanics, Ural Branch of the Russian Academy of Sciences, (1 Akademika Koroleva str., Perm, Perm Region, 614013, Russia), AuthorID 172118, ORCID: 0000-0001-6395-4301;

A. Yu. Belyaev – Junior Researcher, Perm State University (15 Bukireva St., Perm, Perm Region, 614068, Russia), AuthorID 584106, ORCID 0000-0002-0966-5979.

КОМПЬЮТЕРНЫЕ НАУКИ И ИНФОРМАТИКА

Research article

УДК 004.021, 004.056

DOI: 10.17072/1993-0550-2025-2-47-64

<https://elibrary.ru/cvpcqc>**An Algorithm for the Initial Detection of Malicious Traffic Based on the Autoencoder Reconstruction Error and a Variational Model: the Influence of the Error Distribution Density on the Performance Indicators of the Models****Adeyemi Marc Aurele Emmanuel Djeguede**

RUDN University, Moscow, Russia

djeguede.marc@gmail.com

Abstract. The emergence of new sophisticated types of attacks forces the community of computer security researchers to constantly improve detection tools and response methods. The present study explores different factors of autoencoders and variational models that influence their effectiveness in identifying novel attack types and malicious network traffic. The general idea of the proposed algorithm is to construct a confidence interval for the reconstruction error of the training sample, based on which a decision is made on the maliciousness of a particular traffic. Additional emphasis was placed on selecting an appropriate error metric to minimize the overlap between the density distributions of reconstruction errors for normal and malicious traffic. In the study of the variational model, the effect of the t-distribution on the quality of detecting new types of attacks was investigated. The studies were conducted on the CIC-IDS2017 dataset of the Canadian Cybersecurity Institute, containing up to 14 types of traffic and attacks. The experimental results show that with a competent selection of the error measure and the threshold values of the confidence interval, our models outperform existing analogues in various performance indicators.

Keywords: *autoencoders; variational models; zero-day attacks detection; reconstruction error*

For citation: Djeguede, A. M. A. E. (2025), "An Algorithm for the Initial Detection of Malicious Traffic Based on the Autoencoder Reconstruction Error and a Variational Model: the Influence of the Error Distribution Density on the Performance Indicators of the Models", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no 2(69), pp. 47–64. DOI: 10.17072/1993-0550-2025-2-47-64. <https://elibrary.ru/cvpcqc>

The article was submitted 16.01.2025; approved after reviewing 05.05.2025; accepted for publication 11.07.2025.



Эта работа © 2025 Djeguede A. M. A. E. распространяется по лицензии CC BY 4.0. Чтобы ознакомиться с условиями этой лицензии, посетите сайт <https://creativecommons.org/licenses/by/4.0>.

Научная статья

Алгоритм первичного обнаружения вредоносного трафика на основе ошибки восстановления автокодировщика и вариационной модели: влияние плотности распределения ошибок на показатели эффективности моделей

Адейеми Марк Орель Эммануэль Джегеде

Российский университет дружбы народов, Москва, Россия
djeguede.marc@gmail.com

Аннотация. Появление новых изощрённых типов атак вынуждает сообщество исследователей в области компьютерной безопасности постоянно совершенствовать инструменты обнаружения и методы реагирования. В данном исследовании рассматриваются различные факторы, влияющие на эффективность автокодировщиков и вариационных моделей при выявлении новых типов атак и вредоносного сетевого трафика. Общая идея предлагаемого алгоритма заключается в построении доверительного интервала для ошибки восстановления обучающей выборки, на основе которого принимается решение о вредоносной природе конкретного трафика. Особое внимание было уделено выбору подходящей метрики ошибок, чтобы минимизировать перекрытие распределений плотности ошибок восстановления для обычного и вредоносного трафика. При изучении вариационной модели было исследовано влияние t -распределения на качество обнаружения новых типов атак. Исследования проводились на наборе данных CIC-IDS2017 Канадского института кибербезопасности, содержащем до 14 типов трафика и атак. Результаты эксперимента показывают, что при грамотном выборе меры погрешности и пороговых значений доверительного интервала наши модели превосходят существующие аналоги по различным показателям эффективности..

Ключевые слова: автокодировщики; вариационные модели; обнаружение атак нулевого дня; ошибка реконструкции

Для цитирования: Джегеде А. М. О. Э. Алгоритм первичного обнаружения вредоносного трафика на основе ошибки восстановления автокодировщика и вариационной модели: влияние плотности распределения ошибок на показатели эффективности моделей // Вестник Пермского университета. Математика. Механика. Информатика. 2025. № 2(69). С. 47–64. DOI: 10.17072/1993-0550-2025-2-47-64. <https://elibrary.ru/cvpcqc>

Статья поступила в редакцию 16.01.2025; одобрена после рецензирования 05.05.2025; принята к публикации 11.07.2025.

Introduction

The tasks of recognizing known types of attacks and classifying them are based on a priori knowledge about the characteristics of malicious traffic, but the formation and correct formalization of knowledge about new types of attacks is not always a trivial task. The general established concept of detecting new types of attacks is based on deviations searching from the specified characteristics of some normal traffic. A trivial application of this concept was to calculate statistical indicators for each traffic characteristic with the subsequent establishment of a confidence interval for these characteristics, based on which a decision is made about whether a particular traffic belongs to normal traffic or malicious. The rapid development of neural network architectures has given a new impetus to the improvement of tools and models for detecting new types of attacks. The family of autoencoder architectures, with its ability to recreate the input training sample, has found the most suitable application in the field of detecting new types of attacks (zero-day attacks). The main idea in this context is to train an autoencoder to reconstruct normal traffic and, based on the statistical indicators of the reconstruction error, to construct a confidence interval that allows maximizing the

model's performance when detecting various malicious traffic. In this regard, a number of questions arise that have not yet been fully studied by the scientific community. In this paper, the goal was to analyze how the reconstruction error metric and its distribution impact the performance of autoencoders, as well as how the distribution function of the parametrization coefficient ϵ affects the prediction quality of variational models.

Literature review

W. Qiu et al. [1] propose a hybrid intrusion detection system called DST-IDS. DST-IDS analyzes both individual data packets and entire flows, and then applies the Dempster-Schafer theory to predict network attacks. DST-IDS is designed to work in real time with heterogeneous data distribution, when the distribution of the training data set may differ from the distribution of data during implementation. This property significantly increases the applicability of DST-IDS. Y. Jingcheng et al. in [2] presented a new technique for detecting anomalies in traffic that can detect zero-day attacks. For this purpose, 2D images are formed from the traffic flow and fed to the input of an auto encoder based on convolutional neural networks to extract the feature vector. Then, based on the value of the distance function between the vectors of the normal class and the remaining classes, a conclusion is made about the presence of anomalies and zero-day attacks in the traffic. In [3], to solve the class imbalance problem associated with the small number of malicious traffic representatives, J. Zhang et al. proposed Siamese neural networks called MF2POSE. Using the same architecture for the twin networks, they perform two different tasks, multi-class classification for the main network and binary classification for the second network, while the authors applied the transfer learning mechanism from the network with the binary task to the network with the multi-class task. N. Gupta et al. [4] propose an intrusion detection system LIO-IDS based on a recurrent neural network with long short-term memory (LSTM) and an improved One-vs-One algorithm for processing malicious network traffic. LIO-IDS is a two-tier system, the first tier of which separates malicious traffic from normal network traffic using the LSTM classifier. The second tier uses ensemble algorithms and the one-versus-one (I-OVO) method to classify malicious traffic. In addition, to improve the detection performance of small malicious attacks, synthetic data was generated to replenish these classes. D. Olszewski et al. [5] conducted a study to identify the impact of the following algorithms - self-organizing map (SOM), t-distributed stochastic neighbor embedding (t-SNE) and adaptive neighborhood search visualizer (NeRV) on the quality of anomaly detection in IoT network traffic. The original contribution of this study is to conduct a comparative analysis of the t-SNE and NeRV dimensionality reduction approaches in both theoretical and empirical aspects.

To effectively identify malicious traffic, M. Gao et al. [6] developed an anomaly detection method based on the IoT traffic attribute graph. To ensure the applicability of the model in large-scale IoT networks, a graph node compression algorithm based on the Huffman coding algorithm was developed. Qian Ma et al. [7] applied a support vector machine with a kernel function for linearly inseparable samples to classify network traffic anomalies. The work consists of three phases. In the first phase, a new method for transforming input traffic data is proposed based on statistical transformations and k-gram method, which transform each URL into a feature vector with a fixed length. In the second stage, a classifier is trained using the obtained feature vectors as input and the hyperparameter of the support vector machine kernel is tuned using one-dimensional optimization. In the last stage, the classification of unknown URLs is carried out after pre-transforming them into feature vectors. X. Wang et al. [8] propose a lightweight spatio-temporal knowledge transfer (STKD) architecture. The main components of this architecture for anomaly detection include two neural networks MSSTRNet and LENet. MSSTRNet integrates one-dimensional convolutional (1D CNN) and recurrent (LSTM) neural networks as the main learning model with the subsequent objective of transferring its knowledge to LENet, a lightweight model. In order to improve the prediction quality of the model, a focal loss function was introduced instead of the traditional cross-entropy loss function.

Y. Wu et al. [9] developed a new zero-day attack detection system based on the Deep Q-Network (DQN) architecture. Functionally, the system consists of three components - a network intrusion detection system classifier, a sample selection strategy, and annotation. The DQN model plays a control role in detecting, selecting, and annotating packets corresponding to zero-day attacks. A bidirectional recurrent network (BiLSTM) is integrated into the DQN model structure to form a traffic selection policy. The selection is made based on the analysis of temporal correlation by the classifier using the euclidean distance as a measure to separate the selected samples. S. Shen et al. [10] proposed a DQN-based intrusion detection system using heuristic learning for IoT traffic in the scenario of insufficient training samples. The proposed system consists of two modules, a traffic processing module and a DQN-based module.

The network traffic processing module generates IoT traffic samples, splits them into subsets and then transmits them to the classifier and cybersecurity expertise center. The application of reinforcement learning in DQN allows its ability to detect malicious traffic to be gradually improved. In this regard, the reward function is designed to reward correctly classified malicious samples, so that the DQN network maximizes the value of the reward function to find the optimal strategy for training. Swathy Akshaya et al. in [11] explored the benefits of game theory in intrusion recognition and detection. This study considers zero-day traffic samples generated by generative adversarial networks, which pose a significant threat while being undetected by conventional classifiers. It emphasizes the use of adaptive game theory strategies while conducting experiments with a modified Bi-LSTM recurrent network and autoencoders. M. Soltani et al. [12] developed a deep learning-based novel system consisting of four stages. The first stage uses a deep neural network to identify unknown samples, i.e. potential new attacks, while simultaneously generating reports on different known attacks. Then, the new samples are divided into different groups using clustering algorithms by combining the deep model and clustering algorithms. These clusters are the key to creating a label for the unknown samples, thus reducing the time and effort of a team of experts. The next stage involves a team of experts to more formally explore the clusters and assign labels. The last stage uses the labeled groups to update the deep learning model.

September 2016 and March 2018 were marked by the outbreak of Mirai botnet attacks and memcached attacks. Such attacks are characterized by pre-collection of information about the victim devices and their infection. Given that none of the existing intrusion detection systems were able to detect and mitigate the impact, A. Blaise et al. [13] developed an early detection method for emerging botnets and unknown vulnerabilities, which consists of monitoring and collecting port-level statistics using a simple but effective time series anomaly detection algorithm based on the modified Z-score and the hypothesis of normal distribution of the series. Wenqi Chen, Zhiliang Wang, Liyuan Chang et al. [14] presented an anomaly detection model based on an ensemble algorithm called ADSIM.

The ADSIM model implements the unsupervised learning paradigm, that first iteratively and gradually maintains a distance matrix to record the similarity between classifiers during the training phase, and then uses hierarchical clustering to group similar classifiers. During the discovery phase, each cluster will be assigned a weight based on the consistency of the discovery results of the classifiers.

Materials

CIC-IDS-2017 (Canadian Institute for Cybersecurity Intrusion Detection System 2017) is an important resource developed by the Canadian Institute for Cybersecurity to support cybersecurity researchers, academics, and practitioners in their efforts to detect and prevent network attacks. This dataset includes a variety of network traffic types, including both normal and abnormal. Specifically, it covers malicious traffics such as.

PortScan – A port scan is a technique used to probe a server or host for open ports. It can also be used maliciously by attackers to gather information about a target system before launching an attack.

Botnet – A botnet attack is a type of cyberattack that involves a network of compromised devices (called bots or zombies) controlled by an attacker (called the botmaster or command-and-

control (CandC) server). These devices are typically infected with malware, allowing the attacker to remotely control them and use them for malicious purposes. Botnets are a significant threat to cybersecurity due to their ability to scale and execute large-scale attacks.

DoS – Denial-of-Service is an attack where a single source overwhelms a target (server, network, or service) with excessive traffic or malicious requests, rendering it unavailable to legitimate users.

DDoS – Distributed Denial-of-Service is a coordinated attack using multiple compromised devices (a botnet) to flood a target from numerous locations, amplifying impact and complexity.

Heartbleed – Heartbleed is a critical security vulnerability discovered in 2014 that affected the widely used OpenSSL cryptographic library. The Heartbleed bug allowed attackers to exploit a flaw in OpenSSL to steal sensitive data from servers, including private keys, passwords, and other confidential information.

XSS – Cross-Site Scripting (XSS) is a common web application vulnerability that allows attackers to inject malicious scripts (typically JavaScript) into web pages viewed by other users. This can lead to unauthorized actions, data theft, or defacement of websites. XSS attacks exploit the trust a user has for a particular website, making it a serious security threat.

SQL Injection – SQL Injection (SQLi) is a type of web security vulnerability that allows attackers to inject malicious SQL code into queries made by an application to its database. It occurs when user-supplied input is not properly validated, sanitized, or parameterized and can lead to unauthorized access, data theft, data manipulation, or even full control over the database.

Infiltration – It refers to the unauthorized entry into a computer network by malicious actors with the intent to compromise data, systems, or operations. It emphasizes stealth to avoid detection, allowing attackers to persist within the network and achieve objectives such as data theft, espionage, or disruption.

SSH and FTP Brute-force – A brute force attack is a trial-and-error method used by attackers to gain unauthorized access to systems, accounts, or encrypted data. The attacker systematically tries all possible combinations of passwords, encryption keys, or other credentials until the correct one is found.

CIC-IDS-2017 also contains over 80 features, allowing researchers to test their models on real-world scenarios. The data is presented in various formats, including PCAP (Packet Capture) and CSV (Comma-Separated Values), which provides flexibility in its use. The dataset contains metadata about network connections, such as timestamps, IP addresses, ports, and other parameters, which makes it convenient for analysis and training machine learning models. CIC-IDS-2017 is actively used to develop machine learning and deep learning algorithms for intrusion detection, as well as to improve existing security systems.

Methods. Mathematical model and architecture of auto encoder

The mathematical model of the auto encoder can be represented as an optimization of the reconstruction error of input vectors of the form $\phi, \psi = \operatorname{argmin}_{\phi, \psi} \|X - \phi \circ \psi(X)\|^2$, where ϕ - encoding function and ψ - decoding function such that $\phi: X \rightarrow F$ $\psi: F \rightarrow X'$. The reconstruction error can be calculated using one of the following formulas

$$\|x\|_2 = \sqrt{\sum (x' - x_i)^2} \text{ and } \|x\|_1 = \sum_{i=1}^n |x' - x_i|. \quad (1)$$

Variational autoencoder. Based on variational inference, the variational auto encoder belongs to a class of generative models that learn to map input vectors to a latent probability space and back. The VAE model answers the question of what distribution the latent space vectors Z should be generated from, so that the decoder can generate new vectors from the same distribution as X , but different from the existing vectors.

For any $x \in X$ we define $P(x) = \int P(x|z) \cdot P(z) \cdot dz$. To generate latent space vectors, it is necessary to maintain a distribution $P(Z|X)$. From the Bayesian formula we can write $p(Z|X = D) = \frac{P(X=D|Z) \cdot P(Z)}{P(X=D)}$, where $P(X = D) = \int_{z_0} \dots \int_{z_{D-1}} P(X, Z) dz_0 \dots dz_{D-1}$. Since it's impossible to find the closed distribution formula $P(Z|X = D)$ due to expression $P(X = D)$, in the future we will approximate it by distribution $q(z) = N(Z, \mu, \sigma)$ from the family of Gaussian distributions. To estimate the permissible error, replacing $P(Z|X = D)$ by $q(Z)$ we use the Kullback-Leibler divergence. This divergence usually allows us to measure the distance between two distributions. The further task is to minimize the distance between $P(Z|X = D)$ and $q(Z)$, which is expressed by the formula (1) $q = \arg \min_{q(Z) \in Q} KL(q(Z)||P(Z|X = D))$, where Q - family of Gaussian distributions. KL divergence can be defined as the mathematical expectation of the logarithm of the ratio of $q(Z)$ and $P(Z|X = D)$ and express through the formula $KL(q(Z)||P(Z|X)) = E_{Z \sim q(Z)} \left[\log \left(\frac{q(Z)}{p(Z|X=D)} \right) \right]$. Replacing $p(Z|D) = \frac{p(Z,D)}{p(D)}$ in the expression of KL-divergence and transforming it we obtain $KL(q(Z)||p(Z|X)) = -E_{Z \sim q(Z)} \left[\log \left(\frac{p(Z,D)}{q(Z)} \right) \right] + \log(p(D))$, where $L(q) = E_{Z \sim q(Z)} \left[\log \left(\frac{p(Z,D)}{q(Z)} \right) \right]$. We can then rewrite the divergence expression as $KL = -L(q) + \log(p(D))$, where $\log(p(D)) \leq 0$ and $KL \geq 0$, then $L(q) \leq 0$ and defines the evidence lower bound $\log(p(D))$ (ELBO). From all of the above, we can conclude that $L(q) = \log(p(D))$ then and only then when $KL(q(Z)||p(Z|X)) = 0$ and the problem of minimizing divergence $q = \arg \min_{q(Z) \in Q} KL(q(Z)||p(Z|X = D))$ equivalent to $q = \arg \max_{q(Z) \in Q} L(q)$. Variational autoencoder model is showed in fig. 1.

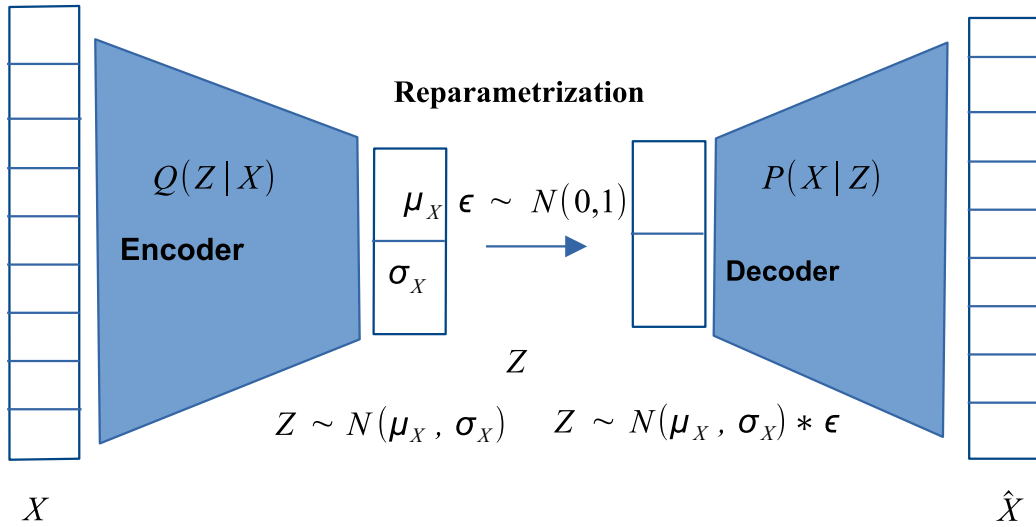


Fig. 1. Variational autoencoder model

Results and discussions

The models of autoencoder and variational autoencoder studied in this paper have the common basic architecture described in table 1. The main difference is in the latent space representation. For the autoencoder, the latent space is deterministic. The encoder maps the input to a fixed point in the latent space. There is no explicit structure or constraints on the latent space,

which can lead to irregular gaps or discontinuities. Variational Autoencoder (VAE) has a probabilistic latent space. The VAE outputs the parameters (mean and variance) of a probability distribution (e.g., Gaussian). The latent space is regularized to follow a specific distribution (e.g., standard normal distribution), ensuring a smooth and continuous latent space. This allows for meaningful interpolation and sampling in the latent space.

Table 1. *Autoencoder architecture*

Encoder	Decoder
512 - ReLu	4 - ReLu
256 - ReLu	8 - ReLu
128 - ReLu	16 - ReLu
64 - ReLu	32 - ReLu
32 - ReLu	64 - ReLu
16 - ReLu	128 - ReLu
8 - ReLu	256 - ReLu
4 - ReLu	512 - ReLu
2 - ReLu	

Since our two models belong to the class of "unsupervised machine learning", the research methodology will be based on finding the threshold value of the training sample reconstruction error at which it is possible to make a decision with a high degree of probability about the belonging of the input vector to a particular traffic class. Note that in the process of training our models, only normal traffic data was used, which allows us to further test the ability of these models to detect new unknown types of attacks and anomalies. As a loss function for the auto encoder, we use binary cross-entropy

$$L_{BCE}(Y, \hat{Y}) = -\frac{1}{N} \cdot \sum_{i=1}^N [y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)].$$

The proposed algorithm fig. 2 below is based on the autoencoder reconstruction error. The term reconstruction error refers to the distance between the original input vector and its image generated by the models under study. The main hypothesis on which the algorithm is built is that the reconstruction error for malicious traffic will be higher than the reconstruction error for normal traffic.

In this paper, we consider several distance measures - mean absolute error (MAE), mean square error (MSE), binary cross entropy (BCE) and harmonic mean error. Figure 3 shows the error panel of all above listed errors and for all traffic classes.

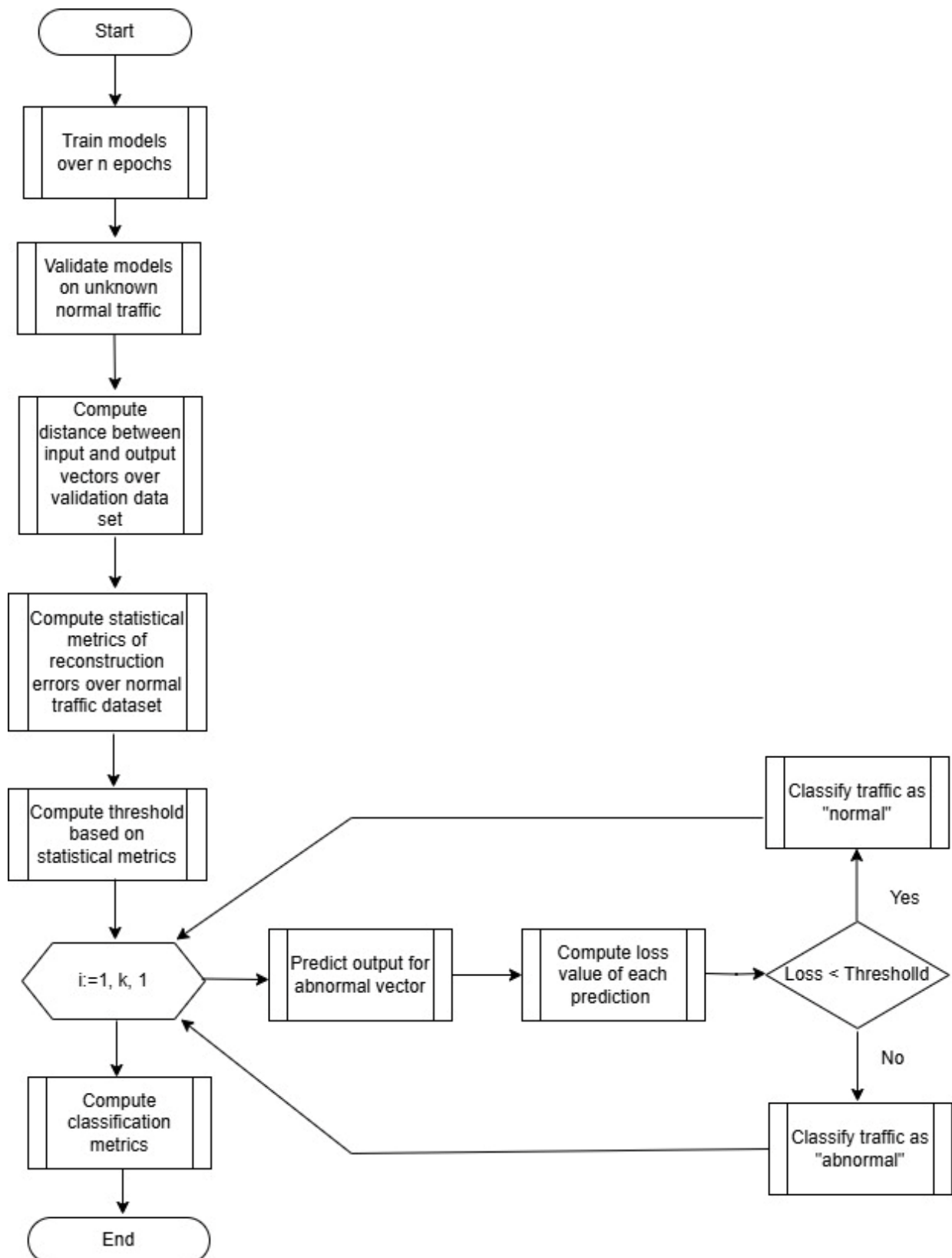


Fig. 2. Reconstruction error based algorithm for traffic classification

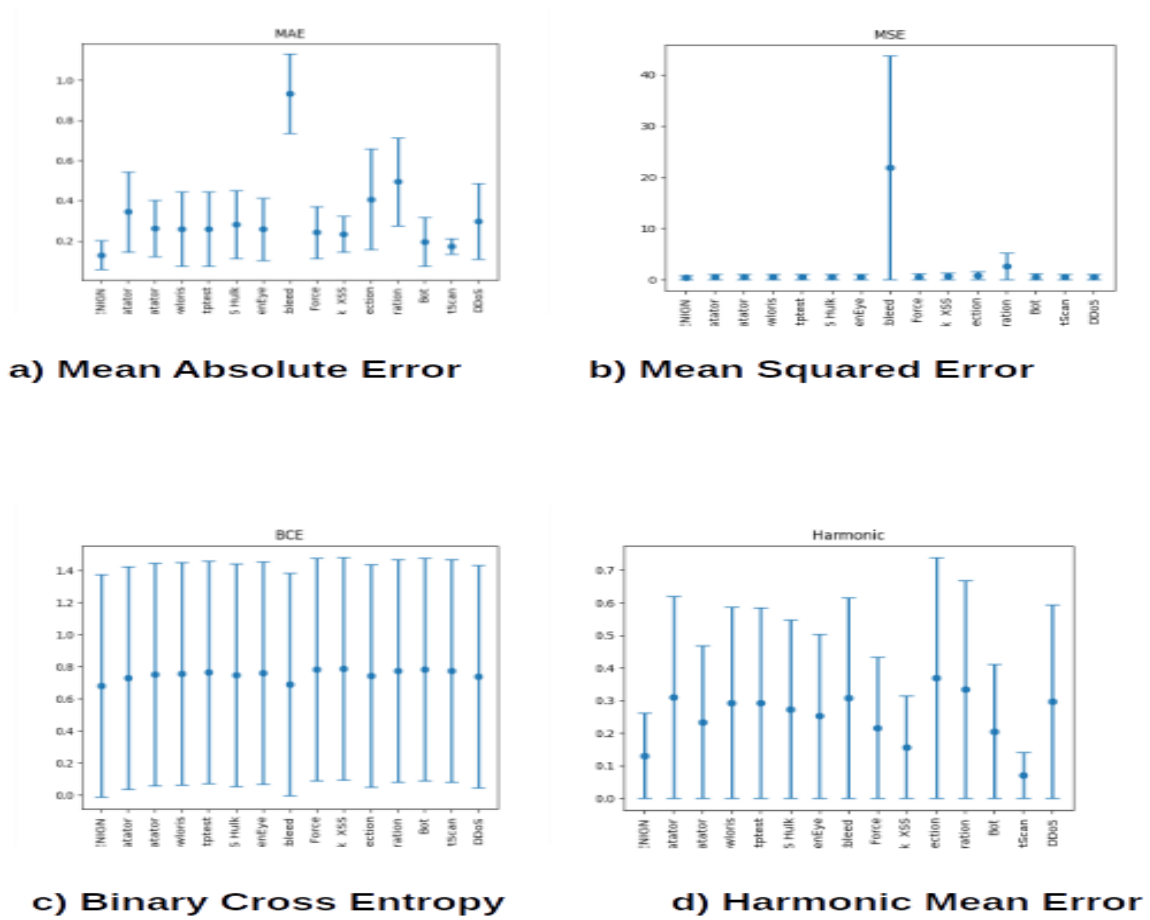


Fig. 3. Confidence intervals of errors for different traffic classes

From fig. 3 it is noticeable that the mean square error is the most suitable error measure for our algorithm as it shows the least degree of overlap between the distributions of normal traffic reconstruction errors and malicious traffic.

A more detailed picture of the mean absolute error for each class of malicious traffic can be seen in figs. 4–6. Each graph represents the difference between the average vector of the corresponding malicious traffic and its reconstruction generated by the auto encoder.

Also based on these graphs it is easy to have an idea of the malicious traffic attributes that make the greatest contribution to its detection.

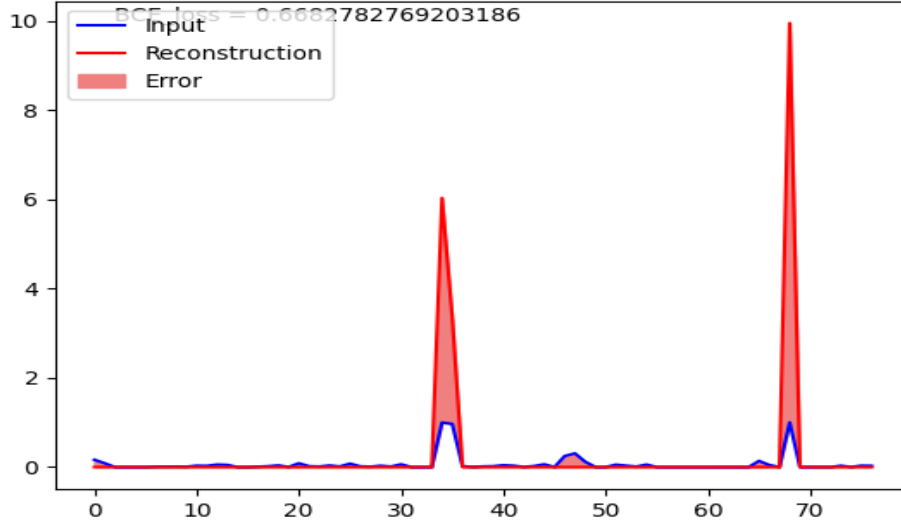


Fig. 4. Average reconstruction error for benign traffic

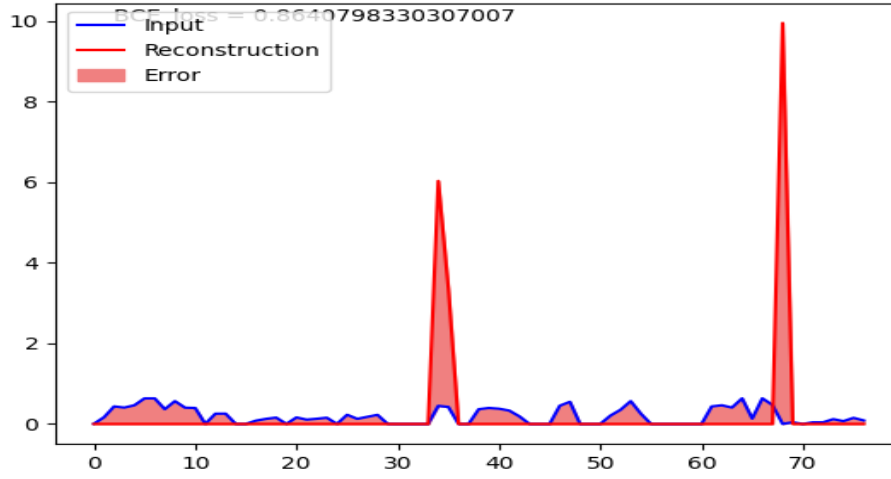


Fig. 5. Average reconstruction error for DDoS traffic

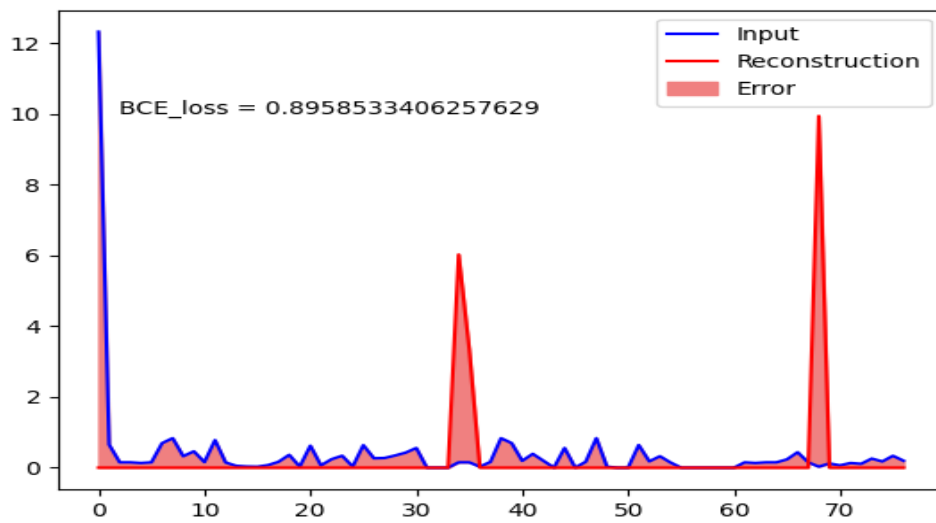


Fig. 6. Average reconstruction error for Infiltration traffic

In machine learning, particularly in the context of training neural networks, an epoch refers to one complete pass through the entire training dataset. During training, the model's parameters (e.g., weights in a neural network) are updated iteratively to minimize the loss function. The number of epochs is a hyperparameter that determines how many times the model will see the entire dataset during training.

To evaluate the quality of the proposed algorithm based on the auto encoder, the following metrics were selected: Accuracy, Recall, F1 score, ROC curve and the area under the curve value.

1. Accuracy measures the proportion of correctly classified instances (both true positives and true negatives) out of the total number of instances.
2. Recall measures the proportion of actual positives that are correctly identified by the model.
3. The F1 score is the harmonic mean of precision and recall, providing a balance between the two.
4. The ROC (Receiver Operating Characteristic) curve is a graphical representation of a model's performance across different classification thresholds.
5. AUC (area under the ROC curve) measures the entire area under the ROC curve, providing a single value to summarize the model's performance across all classification thresholds.

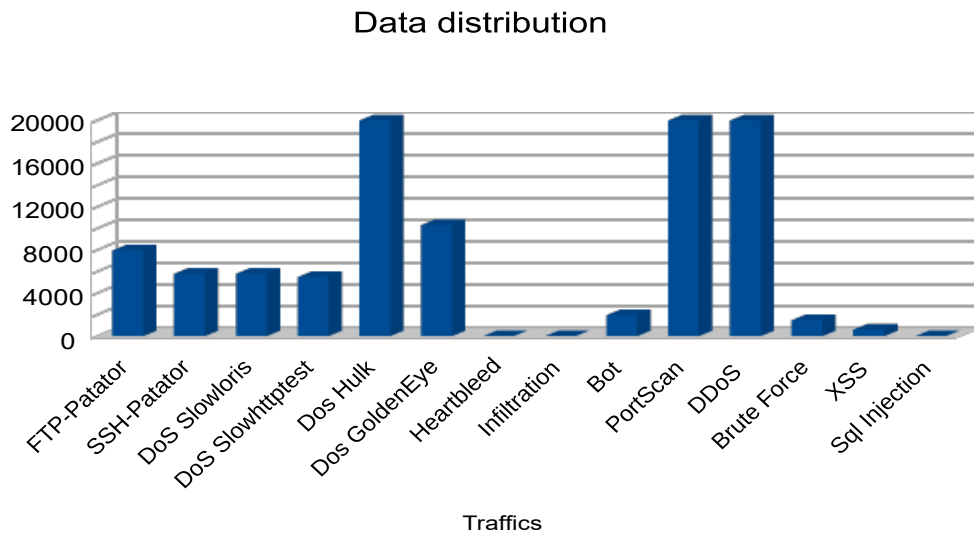
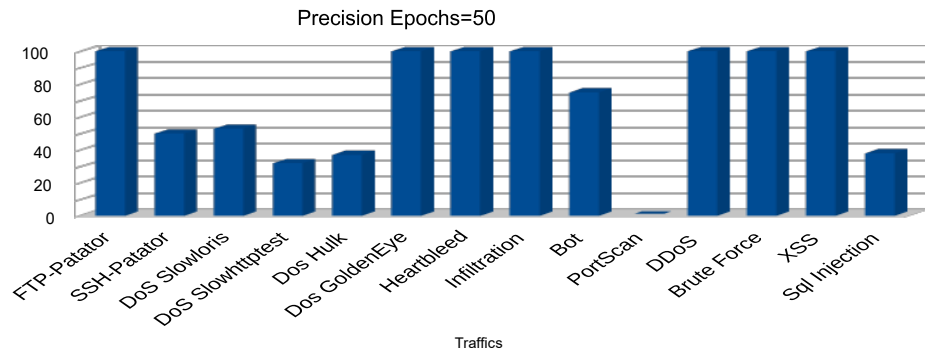


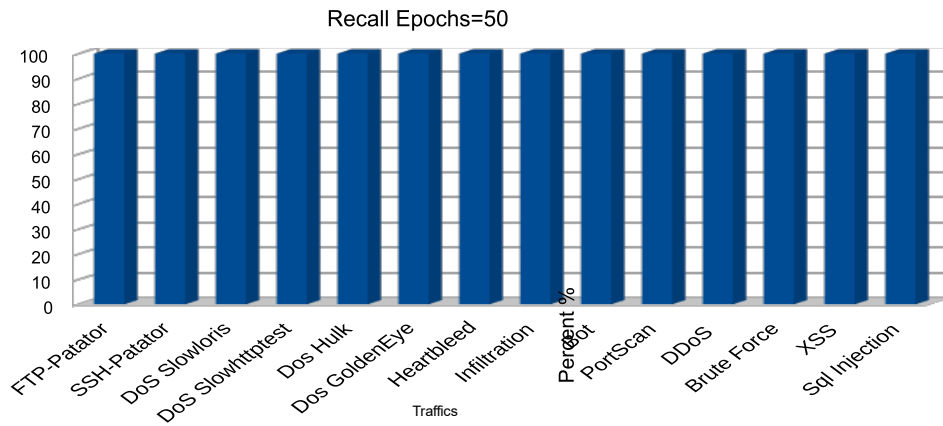
Fig. 7. *Distribution of dataset*

The above metrics were measured after training epochs equal to 50, 150 and 200.

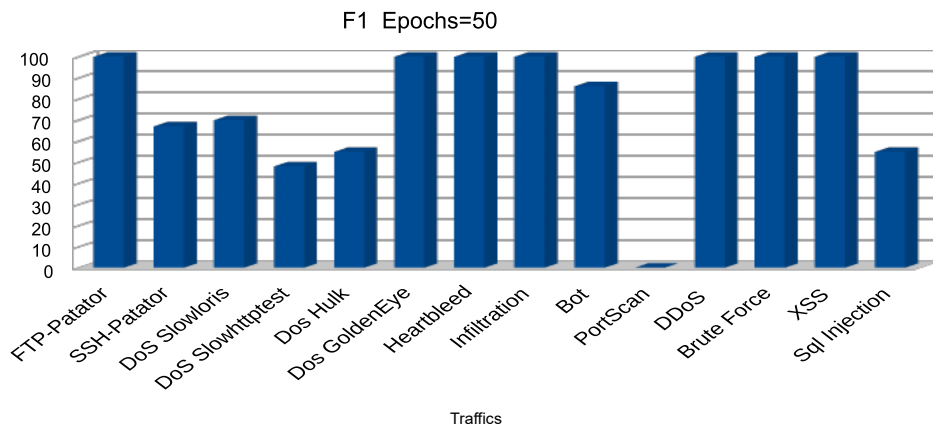
In figure 7 we have the distribution of malicious traffic based on which the tests were conducted. From this histogram we can observe a strong imbalance between different classes of malicious traffic, which reflects the frequency of these attacks in the real world. Thus, we observe that the "Denial of Service (DoS) " attacks are the most common, while more complex types of attacks from a technical point of view such as "Heartbleed" and "Infiltration" are very few. In order to maintain an objective picture of the world, it was decided not to artificially increase the number of rare types of attacks.



a) Precision metric for 50 epochs

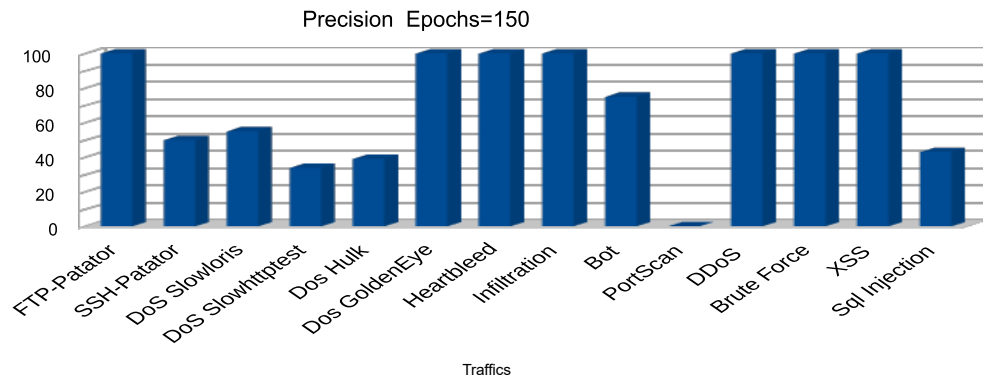


b) Recall metric for 50 epochs

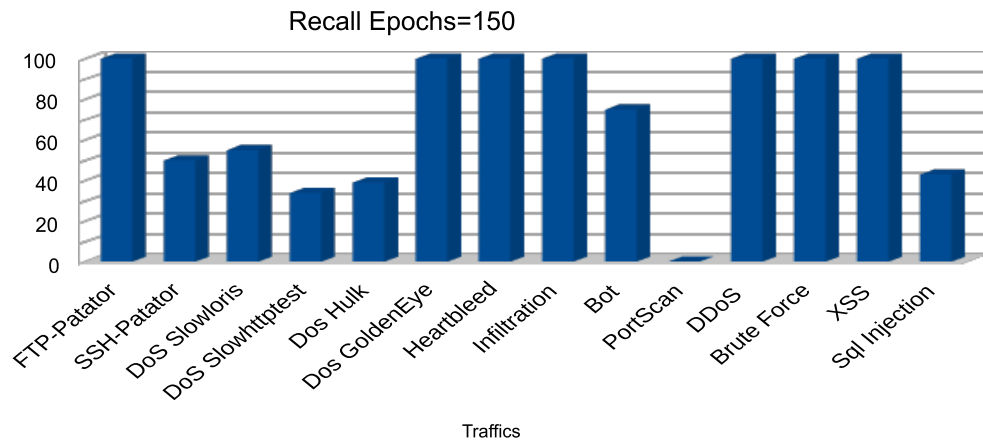


c) F1 metric for 50 epochs

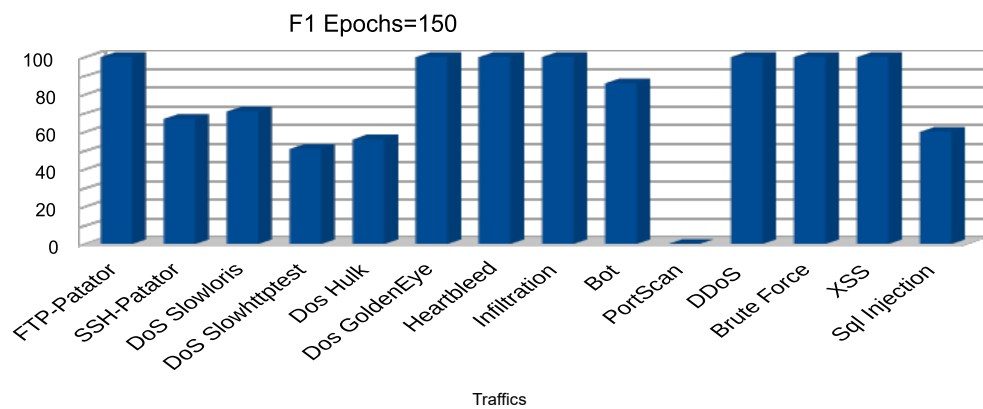
Fig. 8. Autoencoder metrics for 50 training epochs



a) Precision metric for 150 epochs

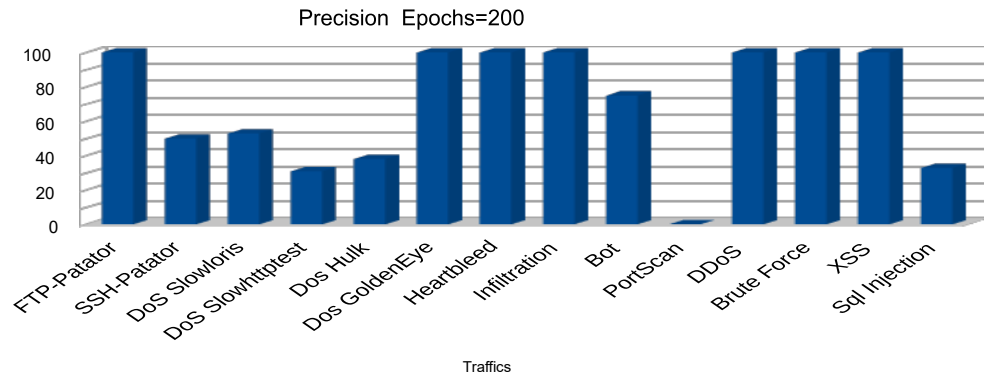


b) Recall metric for 150 epochs

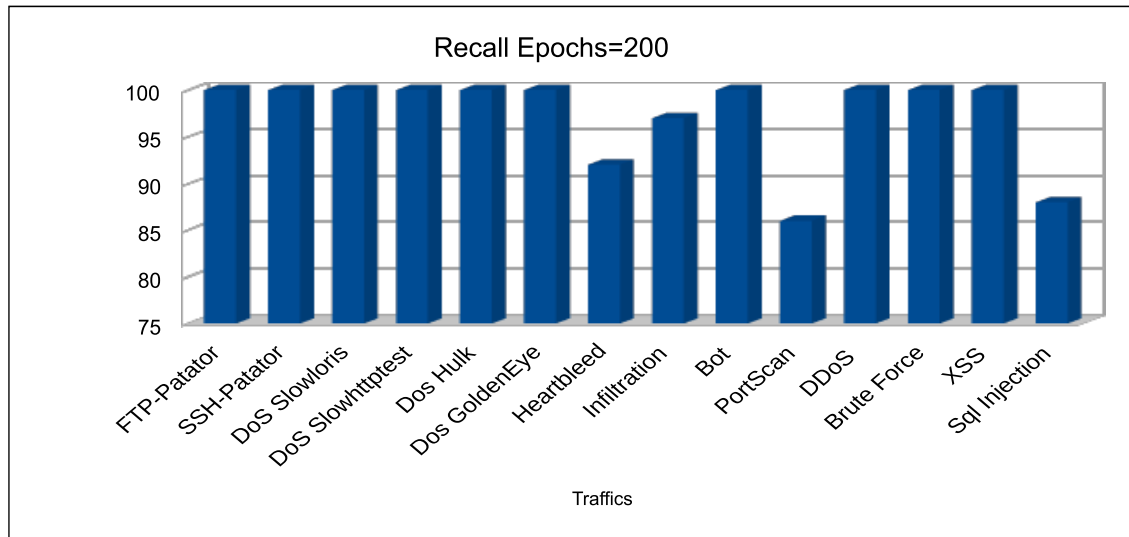


c) F1 metric for 150 epochs

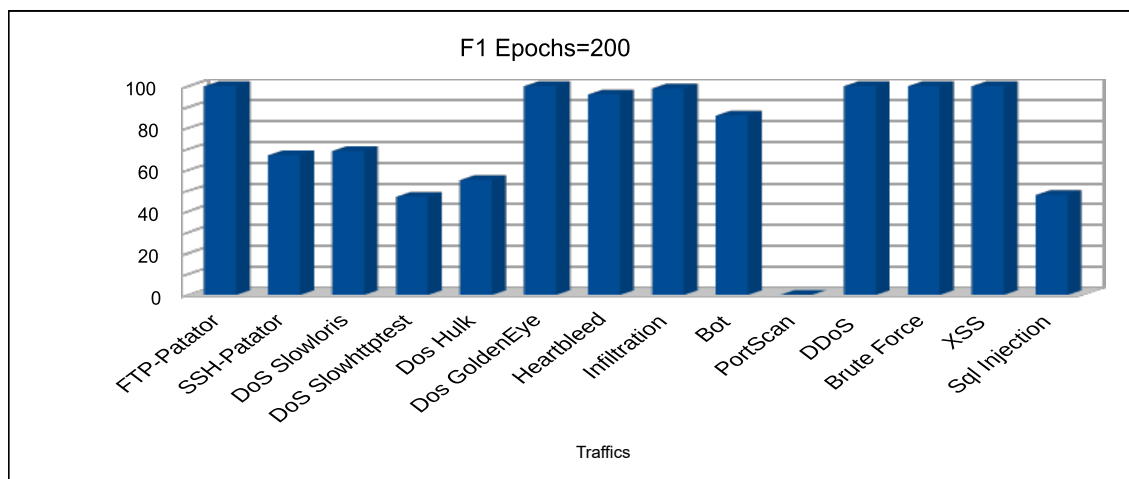
Fig. 9. Autoencoder metrics for 150 training epochs



a) Precision metric for 200 epochs



b) Recall metric for 200 epochs



c) F1 metric for 200 epochs

Fig. 10. Autoencoder metrics fo 200 training epochs

The obtained results displayed in figures 8 b) – c), 9 a) – c), 10 a) – c) show that increasing the number of autoencoder training epochs from 50 to 200 did not have a significant effect on the quality of predictions. It is also interesting to note that the proposed algorithm allows detecting rare types of attacks with 100 % accuracy, 97 % response, and 99 % F1 for the ‘Infiltration’ traffic and 100% accuracy, 92 % response, and 98 % F1 for the ‘Heartbleed’ traffic. Among the ‘Denial of Service’ attacks, the best predictions were obtained for the distributed ‘DDoS’ attack with 100 % accuracy, 99% response, and 99% F1. The model also showed high performance in detecting botnet attacks with 75 % accuracy, 99 % response, and 85 % F1. Among web attacks, the model showed 100 percent performance for all types of attacks except for "Sql injection", for which 30% accuracy, 87 % recall, and 45 % F1 were observed. The lowest results were obtained for "PortScan" traffic with zero accuracy and F1 percentages. This result is explained by the fact that the distribution of reconstruction error for the "Portscan" class is completely covered by the distribution of reconstruction error for normal traffic. The area under ROC curve is the most significant indicator of the model’s performance in the case of a strong imbalance between classes, as in this case. The obtained values for this metric are depicted in table 2 fully correlate with the previous results and show which types of attacks or traffic the model copes with detecting better.

Table 2. Area under ROC curve.

Type of Attack	Area under ROC curve (AuC)
Botnet	0.8
DDoS	1.00
DoS Golden Eye	1.00
DoS Hulk	0.69
DoS Slowhttptest	0.66
DosSlowloris	0.76
FTP Patator	1.00
Heartbleed	1.00
Infiltration	1.00
Portscan	0.50
SSH Patator	0.75
Brute force	1.00
SQL Injection	0.67
XSS	1.00

Variational Auto Encoder Results. As can be seen from the results of testing a simple autoencoder, the proposed model performed well with the primary detection of all types of critical attacks except for the "PortScan" traffic, that is an essential part of network maintenance but can be used by attackers to find open doors. In this section, we present the results of using the variational model for the primary detection of attacks and track the influence of various distribution functions of the parameterization coefficient ϵ on performance metrics. Particular attention will be paid to the t-distribution and the study of the influence of the degree of freedom of this distribution and the number of training epochs on the quality of forecasting "PortScan" traffic. The t-distribution is well known in statistics and serves as a generalization of the normal distribution when the variance with the formula is not known in advance

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi \cdot n} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

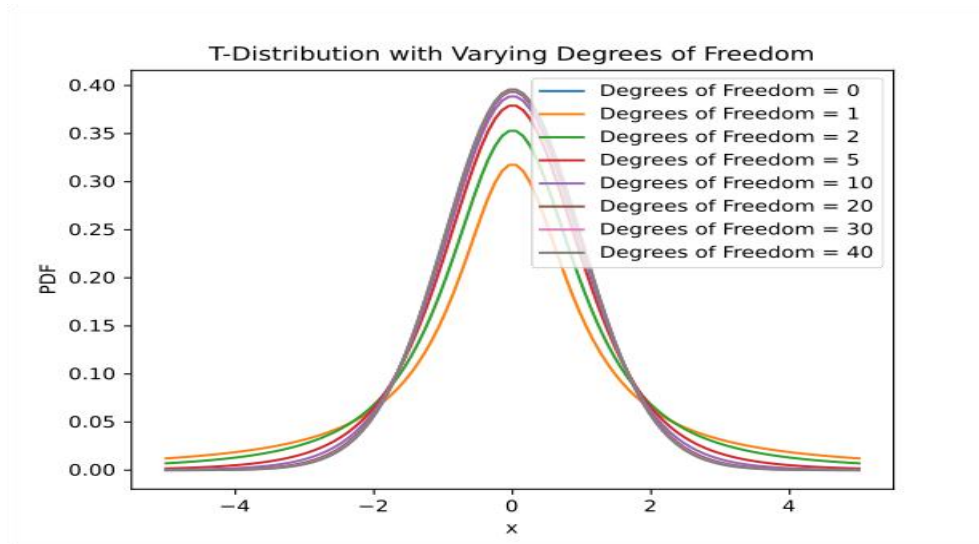


Fig. 11. Probability density of Student's distribution.

The test scenario for the variational autoencoder was conducted for the training epoch values of 50, 100, and 150 and the degrees of freedom values of 10, 20, 30, 40, and 50. The obtained experimental results show a significant improvement in the model's performance in detecting 'PortScan' traffic. In Fig. 12a-d), the area under the ROC curve ranges from 80 % to 80.5 %, which is 30% higher than the result obtained from the simple autoencoder.

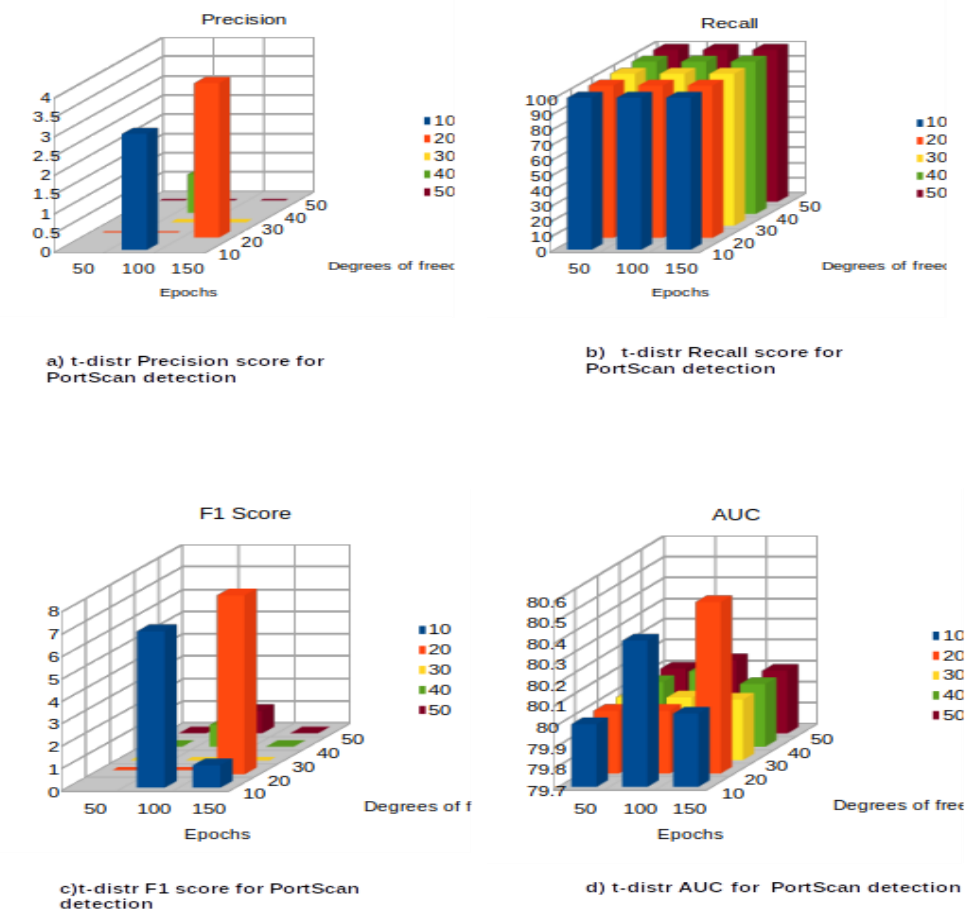


Fig. 12. Performance metrics of the variational model for detecting "PortScan" traffic

Conclusion

Primary detection of increasingly sophisticated new types of attacks is a topical issue in the field of computer security. The ability of the architecture of autoencoder families based on deep neural networks to reconstruct large-scale datasets with minimal deviations prompted us to consider these models as objects of our research. Thus, in the course of this work, an algorithm for detecting zero-day attacks based on an auto encoder and a variational auto encoder consisting of 17 layers was proposed. The studies showed: a direct correlation between the indicators of various metrics (accuracy, response, F1-measure, ROC curve and area under the curve) and the degree of overlap of the distribution densities of the reconstruction error of normal traffic and detected malicious traffic; the ability of the auto encoder to cope with the primary detection of all types of attacks, with the exception of "PortScan" traffic due to the first point above. The achieved indicators for all metrics in their totality exceed the results of similar studies; the use of parametric distribution functions (Student's t-distribution in this particular study) for the parameterization coefficient ϵ of the variational model allows us to more clearly separate the distribution densities of the reconstruction error of normal traffic and "PortScan" traffic.

Comparisons of the obtained results with the results of a similar work [2], in which autoencoders based on deep neural networks were used, are given in Table 2.

Table 3. Comparison of the obtained results

	few-shot learning model	Proposed auto encoder
FTP	Accuracy: 95% Precision: - Recall: 99%	Accuracy: 99.99% Precision: 100% Recall: 99.99%
SSH	Accuracy: 99% Precision: - Recall: 100%	Accuracy: 91.82% Precision: 50.28% Recall: 99.97%
PortScan	Accuracy: 99% Precision: Recall: 99%	Accuracy: 60% Precision:0 Recall: 85%
DoS	Accuracy: 98% Precision: Recall: 99%	Accuracy: 99.99% Precision: 99.98% Recall: 99.99%
DDoS	Accuracy: 94% Precision: - Recall: 96%	Accuracy: 99.99% Precision: 100% Recall: 99.99%
Bot	Accuracy: 90.6% Precision: - Recall: 93.8%	Accuracy: 98.46% Precision: 74.85% Recall: 99.93%

Analysis of these results shows that the proposed model is inferior only in detecting "PortScan" traffic, and in other cases it surpasses similar models.

References

1. Qiu, W., Ma, Y., Chen, X., Yu, H. and Chen, L. (2022), "Hybrid intrusion detection system based on Dempster–Shafer evidence theory", *Elsevier BV*, vol. 117, pp. 4–6. doi:10.1016/j.cose.2022.102709. EDN: VYGLJS.

2. Yang, J., Li, H., Shao, S., Zou, F. and Wu, Y. (2022), "FS-IDS: A Framework for Intrusion Detection Based on Few-Shot Learning", *Elsevier BV*, vol. 122, pp. 3–8. doi:10.1016/j.cose.2022.102899. EDN: FLHKJN.
3. Zhang, J., Chen, R., Zhang, Y., Han, W., Gu, Z., Yang, S. and Fu, Y. (2024), "MF2POSE: Multitask feature fusion pseudo-siamese network for intrusion detection using category-distance promotion loss", *Elsevier BV*, vol. 283, pp. 3–5. doi:10.1016/j.knosys.2023.111110. EDN: ODGZZD.
4. Gupta, N., Jindal, V. and Bedi, P. (2021), "LIO-IDS: Handling Class Imbalance Using LSTM and Improved One-Vs-One Technique in Intrusion Detection System", *Elsevier BV*, vol. 192, pp. 6–7. doi:10.1016/j.comnet.2021.108076. EDN: TECCVW.
5. Olszewski, D., Iwanowski, M. and Graniszewski, W. (2024), "Dimensionality Reduction for Detection of Anomalies in the IoT Traffic Data", *Elsevier BV*, vol. 151, pp. 137–151. doi:10.1016/j.future.2023.09.033. EDN: VPAVLN.
6. Gao, M., Wu, L., Li, Q. and Chen, W. (2023), "Anomaly Traffic Detection in IoT Security Using Graph Neural Networks", *Elsevier BV*, vol. 76, pp. 3–7. doi:10.1016/j.jisa.2023.103532. EDN: FNAUKF.
7. Ma, Q., Sun, C., Cui, B. and Jin, X. (2021), "A Novel Model for Anomaly Detection in Network Traffic Based on Kernel Support Vector Machine", *Elsevier BV*, vol. 104, pp. 3–7. doi:10.1016/j.cose.2021.102215. EDN: BJBXZA.
8. Wang, X., Wang, Z., Wang, E. and Sun, Z. (2024), "Spatial-Temporal Knowledge Distillation for Lightweight Network Traffic Anomaly Detection", *Elsevier BV*, vol. 137, pp. 3–7. doi:10.1016/j.cose.2023.103636. EDN: XKZBIV.
9. Wu, Y., Hu, Y., Wang, J., Feng, M., Dong, A. and Yang, Y. (2024), "An Active Learning Framework Using Deep Q-Network for Zero-Day Attack Detection", *Elsevier BV*, vol. 139, pp. 2–6. doi:10.1016/j.cose.2024.103713. EDN: ZPMTLG.
10. Shen, S., Cai, C., Li, Z., Shen, Y., Wu, G. and Yu, S. (2024), "Deep Q-Network-Based Heuristic Intrusion Detection Against Edge-Based SIoT Zero-Day Attacks", *Elsevier BV*, vol. 150, pp. 3–7. doi:10.1016/j.asoc.2023.111080.
11. Akshaya, S. and Padmavathi, G. (2024), "Enhancing Zero-Day Attack Prediction: A Hybrid Game Theory Approach with Neural Networks", *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, pp. 643–663.
12. Soltani, M., Ousat, B., Siavoshani, M. J. and Jahangir, A. H. (2023), "An Adaptable Deep Learning-Based Intrusion Detection System to Zero-Day Attacks", *Elsevier BV*, vol. 76, pp. 3–6. doi:10.1016/j.jisa.2023.103516. EDN: SWBFSN.
13. Blaise, A., Bouet, M., Conan, V. and Secci, S. (2020), "Detection of Zero-Day Attacks: An Unsupervised Port-Based Approach", *Elsevier BV*, vol. 180, pp. 4–6. doi:10.1016/j.comnet.2020.107391. EDN: OKBLHN.
14. Chen, W., Wang, Z., Chang, L., Wang, K., Zhong, Y., Han, D., Duan, C., Yin, X., Yang, J. and Shi, X. (2024), "Network Anomaly Detection via Similarity-Aware Ensemble Learning with ADSim", *Elsevier BV*, vol. 247, pp. 4–5. doi:10.1016/j.comnet.2024.110423. EDN: HLGASK.

Information about the author:

A. M. A. E. Djeguede – post graduate student of the Mathematical Modeling and artificial intelligence Department, RUDN University (6, Miklukho-Maklaya St., Moscow, Russia, 117198).

Научная статья

УДК 004.932.2

DOI: 10.17072/1993-0550-2025-2-65-87

<https://elibrary.ru/sijpkl>

Модели сегментации сердечных артерий по коронарографическим снимкам

Владислав Антонович Бочкарев¹, Александр Андреевич Усынин², Александр Денисович Осипов³, Михаил Русланович Аухадиев⁴, Роман Владиславович Шаров⁵, Марина Александровна Барулина⁶

^{1,2,3,4,5,6} Пермский государственный национальный исследовательский университет,
г. Пермь, Россия

¹pikova98@gmail.com

²alexandrusynin2306@gmail.com

³alexosipov03@yandex.com

⁴tyver9@gmail.com

⁵roma752275@mail.ru

⁶mab@psu.ru

Аннотация. Сердечно-сосудистые заболевания продолжают лидировать среди причин смертности. Снизить смертность и повысить точность диагностики можно с помощью разработки решений на основе искусственного интеллекта. При этом важнейшей из задач является сегментация сердечных артерий. Точная сегментация коронарных сосудов на ангиографических снимках критически важна для выявления стеноза и планирования интервенционных процедур. Однако автоматизированные методы сегментации сталкиваются с рядом проблем, связанных с затруднением распознавания сердечных артерий: неравномерное распределение контрастного вещества, артефакты движения, наложение теней от анатомических структур. В данной работе представлена модель сегментации сосудов на основе модифицированной архитектуры U-Net, включающей residual-блоки и механизмы внимания (SCSE) с предварительным обучением энкодера на искусственном датасете для выделения сосудистых признаков. Особое внимание уделено устойчивости модели к шумам – ключевой проблеме ангиографических данных, с которой не справляются многие существующие методы. Модель обучалась на смешанном наборе данных из 1285 двумерных коронарографических снимков, аннотированных экспертами. Предложенный подход достиг значения IoU 0.54 и F1-score 0.79 на тестовой выборке, демонстрируя устойчивость к шуму и артефактам. Дополнительная постобработка с адаптивной фильтрацией улучшила качество бинарных масок, устраняя ложные срабатывания от катетеров и металлических объектов. Внешняя оценка на 50 изображениях из независимого клинического набора показала IoU 0.50 и F1-score 0.75, что превосходит базовые методы сегментации, такие как классическая U-Net (IoU 0.42). Результаты подчеркивают эффективность предложенной модели для сегментации сосудов в условиях реальных ангиограмм, а также демонстрируют потенциал подхода как основы для последующей 3D-



Эта работа © 2025 Бочкарев В. А., Усынин А. А., Осипов А. Д., Аухадиев М. Р., Шаров Р.В., Барулина М. А. распространяется по лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>.

реконструкции сосудистой сети, что может улучшить диагностику и планирование лечения стеноза коронарных артерий.

Ключевые слова: сердечно-сосудистые заболевания; коронарография; сегментация коронарных сосудов; предобработка изображений; U-Net; механизмы внимания; глубокое обучение

Для цитирования: Бочкарев В. А., Усынин А. А., Осипов А. Д., Аухадиев М. Р., Шаров Р. В., Барулина М. А. Модели сегментации сердечных артерий по коронарографическим снимкам // Вестник Пермского университета. Математика. Механика. Информатика. 2025. Вып. 2(69). С. 65–88. DOI: 10.17072/1993-0550-2025-2-65-87. <https://elibrary.ru/sijpkl>

Статья поступила в редакцию 16.04.2025; одобрена после рецензирования 25.05.2025; принята к публикации 09.07.2025.

Research article

Models of Cardiac Artery Segmentation From Coronary Angiographic Images

Vladislav A. Bochkarev¹, Alexander A. Usynin², Alexander D. Osipov³, Mikhail R. Aukhadiev⁴, Roman V. Sharov⁵, Marina A. Barulina⁶

^{1,2,3,4,5,6} Perm State University, Perm, Russia

¹pikova98@gmail.com

²alexandrusynin2306@gmail.com

³alexosipov03@yandex.com

⁴tyver9@gmail.com

⁵roma752275@mail.ru

⁶mab@psu.ru

Abstract. Cardiovascular diseases remain the leading cause of death. Mortality can be reduced and diagnostic accuracy can be improved by developing artificial intelligence-based solutions. The most important task here is segmentation of cardiac arteries. Accurate segmentation of coronary vessels on angiographic images is critical for detecting stenosis and planning interventional procedures. However, automated segmentation methods face a number of problems associated with difficulty in recognizing cardiac arteries: uneven distribution of contrast agent, motion artifacts, and superposition of shadows from anatomical structures. This paper presents a vessel segmentation model based on a modified U-Net architecture, including residual blocks and attention mechanisms (SCSE) with pre-training of the encoder on an artificial dataset for extracting vascular features. Particular attention is paid to the model's resistance to noise, a key problem of angiographic data that many existing methods cannot cope with. The model was trained on a mixed dataset of 1285 2D coronary angiogram images annotated by experts. The proposed approach achieved an IoU of 0.54 and an F1-score of 0.79 on the test set, demonstrating robustness to noise and artifacts. Additional post-processing with adaptive filtering improved the quality of binary masks, eliminating false positives from catheters and metal objects. External evaluation on 50 images from an independent clinical dataset showed an IoU of 0.50 and an F1-score of 0.75, which outperforms baseline segmentation methods such as the classical U-Net (IoU 0.42). The results highlight the effectiveness of the proposed model for vessel segmentation in real-world angiograms and demonstrate the potential of the approach as a basis for subsequent 3D reconstruction of the vascular network, which may improve diagnosis and treatment planning for coronary artery stenosis.

Keywords: cardiovascular diseases; coronary angiography; coronary vessel segmentation; image preprocessing; U-Net; attention mechanisms; deep learning

For citation: Bochkarev, V. A., Usynin, A. A., Osipov, A. D., Auhadiev, M. R., Sharov, R. V. and Barulina, M. A. (2025), "Models of Cardiac Artery Segmentation Based on Coronary Angiographic Images", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), p. 65–88. DOI: 10.17072/1993-0550-2025-2-65-87. <https://elibrary.ru/sijpkl>

The article was submitted 16.04.2025; approved after reviewing 25.05.2025; accepted for publication 09.07.2025.

Введение

Сердечно-сосудистые заболевания (ССЗ) являются ведущей причиной смертности и инвалидности во всем мире [1]. По данным Всемирной организации здравоохранения, только в 2021 году ССЗ унесли жизни более 20 миллионов человек, что составляет около трети всех зарегистрированных смертей на планете [2]. Среди множества патологических процессов, лежащих в основе ССЗ, особое внимание привлекает стеноз коронарных артерий – патологическое сужение сосудов, которое нарушает нормальное кровоснабжение миокарда и может стать причиной инфаркта миокарда, сердечной недостаточности и других критических осложнений [3]. Успешное лечение таких состояний во многом зависит от точной и своевременной диагностики, а также от грамотного планирования интервенционных процедур, таких как стентирование или коронарное шунтирование, что подчеркивает важность современных подходов к визуализации и анализу сосудистой сети сердца [4]. В этом контексте детальная визуализация коронарных сосудов становится незаменимым инструментом, предоставляя врачам информацию о морфологии, анатомических особенностях и функциональном состоянии сосудов, что критически важно для принятия клинических решений [5].

Исследования в области автоматической сегментации коронарных сосудов ведутся как российскими, так и зарубежными учеными. Например, ряд работ подчеркивает важность точного выделения сосудистых структур на 2D-ангиограммах как основы для построения 3D-моделей, что требует применения сложных алгоритмов обработки изображений [6]. Такие подходы приобретают особую актуальность в ситуациях, когда традиционные методы синхронизации данных, такие как использование электрокардиограммы (ЭКГ) для определения фазы сердечного цикла, недоступны в реальной клинической практике [7]. Автоматизация процессов сегментации и анализа не только сокращает время, необходимое для подготовки диагностических данных, но и повышает воспроизводимость результатов, что особенно ценно для специалистов с ограниченным опытом, таких как молодые кардиохирурги [8]. Кроме того, использование ИИ позволяет минимизировать человеческий фактор, снижая вероятность ошибок, связанных с субъективной интерпретацией изображений [9].

Для диагностики и лечения ССЗ специалисты традиционно полагаются на стандартную коронарографию – метод, который позволяет получать двумерные (2D) изображения сосудов в реальном времени. Этот подход уже давно зарекомендовал себя в клинической практике благодаря своей доступности и информативности [5]. В последние годы искусственный интеллект, и, в частности, глубокое обучение, показал высокую эффективность в анализе медицинских изображений, включая задачи сегментации сосудистых структур [8]. Одной из наиболее успешных архитектур для таких задач является U-Net – сверточная нейронная сеть, разработанная для сегментации биомедицинских изображений и демонстрирующая высокую точность в выделении сложных структур на зашумленных данных [10].

Для повышения эффективности U-Net часто модифицируется: например, добавляются residual-блоки для улучшения сходимости обучения [11], а также

механизмы внимания, такие как SCSE (Spatial and Channel Squeeze-and-Excitation), которые позволяют модели фокусироваться на наиболее значимых областях изображения [12]. Эти улучшения особенно полезны при работе с ангиографическими данными, где низкий контраст и шумы являются распространенными проблемами.

Целью данной работы является разработка и валидация метода автоматической сегментации коронарных сосудов на двумерных ангиографических снимках с использованием глубокого обучения. Предложенный подход направлен на преодоление ограничений традиционного анализа за счет применения модифицированной модели глубокого обучения, обученной на смешанном датасете, включающем реальные и синтетические данные, что обеспечивает устойчивость к вариациям качества изображений.

В качестве базовой архитектуры для построения автоматического сегментатора была выбрана архитектура U-Net, как наиболее гибкая из существующих базовых архитектур, и способная эффективно сочетать локальные, глобальные признаки через skip-связи [10], что особенно важно для сегментации объектов с тонкими границами, такими как тонкие сосуды на ангиограммах с низким контрастом.

2. Методы исследования

2.1. Наборы данных

2.1.1. Исходный набор данных (Raw)

Данные для исследования включали 36000 анонимизированных необработанных DICOM файлов, каждый DICOM содержал 30–60 изображений коронарной ангиографии, предоставленных Федеральным центром сердечно-сосудистой хирургии им. С.Г. Суханова (ФЦССХ). Эти изображения представляли собой рентгеновские ангиограммы, выполненные с использованием стандартного оборудования и контрастного вещества, введенного через катетер (рис. 1).

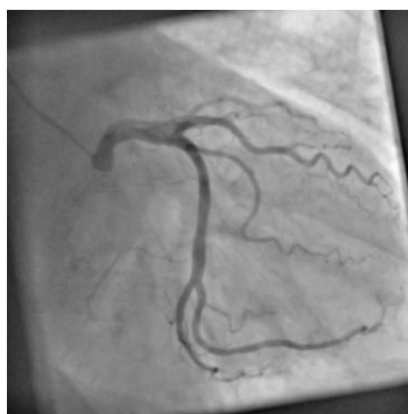


Рис. 1. Снимок коронарографии

Из-за низкого контраста не все изображения подходили для анализа, так как на таких изображениях плохая видимость сердечно-сосудистых артерий. Поэтому были взяты по три кадра из середины каждого DICOM файла, когда контраст в большинстве случаев заполнил сердечные артерии. Таким образом, был сформирован датасет (Raw-датасет) из 108000 изображений.

2.1.2. Набор данных для классификатора (Good/Bad)

Для последующей фильтрации изображений был сформирован датасет "Good/Bad", включающий 500 изображений с низким качеством и 200 изображений с высоким качеством, отобранных вручную (рис. 2).

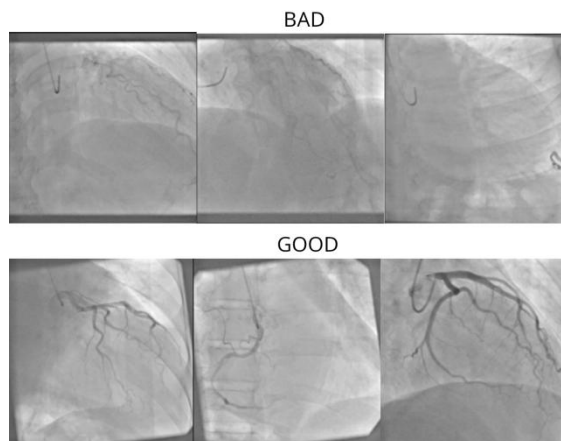


Рис. 2. Пример датасета (Good/Bad)

Критериями отнесения изображений к категории "плохих" служили: некорректная экспозиция (например, чрезмерная затемненность или пересвет), наличие выраженных артефактов (в том числе сильное размытие, обусловленное движением сердца), недостаточная визуализация контраста, а также низкая различимость сосудистой структуры. В свою очередь, к категории "хороших" относились изображения с отчетливо визуализированными артериями и выраженным контрастированием.

Данный датасет был использован для обучения модели классификации изображений на два класса на основе архитектуры ResNet-50 [11]. После завершения обучения модель была применена к полному набору изображений (108 000 снимков), по результатам чего было отобрано 14 000 снимков высокого качества, составивших итоговый RawGood-датасет.

2.1.3. Размеченный набор данных (Segmentation)

1285 изображений коронарографии из RawGood-датасета были размечены с помощью экспертов – сотрудников ФЦССХ. Разметкой являлось выделение маски, соответствующей сердечным артериям. Датасет сформирован из необработанных снимков путем разметки в программе "Label-Studio" [13], тем самым был получен Segmentation-датасет (рис. 3).



Рис. 3. Пример датасета (Segmentation)

2.1.4. Набор данных с шумами

Ключевым недостатком Segmentation-датасета является его "стерильность". В полученном наборе данных крайне мало примеров разметки зашумленных снимков, что ограничивает вариативность данных. Для решения этой проблемы был выбран метод искусственного добавления шумов, для которого был создан дополнительный датасет. При решении задачи сегментации были выделены следующие ключевые проблемы реальных данных ангиографических снимков (рис. 4).

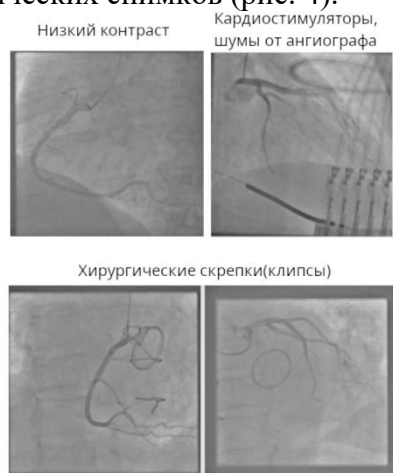


Рис. 4. Проблемные для сегментации снимки коронарографии

Первая проблема – низкий контраст. Контрастное вещество распределяется неравномерно, что приводит к слабой видимости тонких сосудов, особенно на периферии [6]. Это усложняет задачу сегментации, так как модель должна отличать сосуды от фона с минимальными визуальными подсказками. Вторая проблема – наличие таких артефактов, как:

- клапаны и скобы. Металлические объекты, такие как хирургические скобы или клапаны, создают яркие пятна, которые могут быть ошибочно приняты за сосуды;
- кардиостимуляторы. Провода и электроды кардиостимуляторов пересекают область сердца, добавляя линейные структуры, схожие с сосудами по форме;
- шумы от ангиографа. Случайные артефакты, вызванные движением пациента или оборудования, ухудшают качество изображения;
- длинный катетер. Катетер, используемый для введения контрастного вещества, часто виден на снимках как длинная изогнутая линия, которую модель может ошибочно идентифицировать как сосуд, из-за схожей интенсивности и формы [6].

Таким образом, для последующего добавления шумов был создан Noise-датасет. Он состоит из снимков, имитирующих естественные шумы, появляющиеся на типичных ангиограммах. Шумы были созданы путем копирования артефактов и искажений с реальных снимков Raw-датасета и верифицированы врачами. Был создан 31 образец уникальных шумов. Датасет состоит из 5 типов шумов (рис. 5):

1. Сердечные клапаны – (механические или биологические) протезы, представляющие собой анатомические или имплантируемые структуры, которые создают яркие тени или пятна на ангиографических снимках из-за высокой плотности материала, что может затруднять выделение сосудов.

2. Имплантируемые кардиостимуляторы – медицинские устройства с металлическими электродами и проводами, которые пересекают область сердца, формируя линейные артефакты, схожие по форме с сосудами, что приводит к проблемам при сегментации.
3. Ангиографические катетеры – гибкие трубки, используемые для введения контрастного вещества, которые отображаются на снимках как длинные изогнутые линии с интенсивностью, близкой к сосудам, что делает их частой причиной ложных срабатываний в задачах сегментации.
4. Хирургические скрепки (клипсы) – металлические фиксаторы, оставшиеся от предыдущих операций, которые проявляются на ангиограммах как темные скрюченные кривые, ошибочно принимаемые моделью за элементы сосудистой сети.
5. Шумы от ангиографа – элементы ангиографической машины, которые иногда попадают в кадр.

Цель создания данного датасета в том, чтобы в процессе обучения модели иногда добавлять шум к исходному изображению, так, чтобы модель научилась игнорировать похожие структуры (рис. 6). Было принято решение не создавать отдельный датасет с исходными снимками и шумами, а перенести процесс создания зашумленных примеров прямо в процесс предобработки данных при обучении модели.

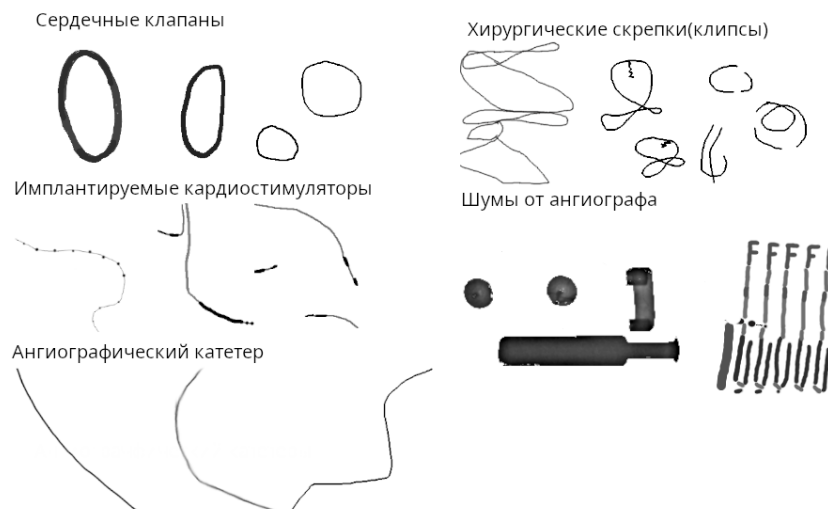


Рис. 5. Срез Noise-датасета с естественными шумами разбитыми по группам

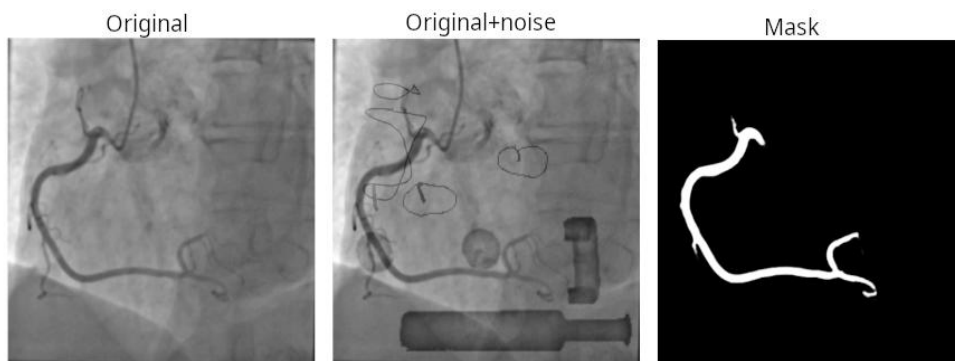


Рис. 6. Создание зашумленного примера, используя данные из (Segmentation) и (Noise) датасетов

2.1.5. Искусственные данные для предобучения (Filter)

Для повышения эффективности обучения модели сегментации и улучшения ее способности извлекать признаки сосудистых структур был разработан искусственный датасет, предназначенный для предварительного обучения (pre-training) модели сегментации на прокси-задаче [14].

Искусственный набор данных (Filter-датасет) был создан на основе (RawGood) данных. Каждое изображение подвергалось многоэтапной обработке для формирования одноканального изображения, акцентирующего внимание на сосудистых структурах. Размер каждого изображения был приведен к стандартному разрешению 512x512 пикселей, что соответствует входным требованиям модели U-Net [10]. Этапы обработки:

1. Фильтрация Mean Shift. Исходные снимки обрабатывались с использованием алгоритма Mean Shift [15], который представляет собой метод кластеризации в пространстве интенсивностей и координат. Mean Shift эффективно сглаживает шумы и однородные области фона, сохраняя при этом резкие границы объектов, такие как сосуды. Параметры фильтра (ширина окна $h_s = 16$ для пространственного измерения и $h_r = 32$ для диапазона интенсивности) были подобраны эмпирически для оптимального подавления небольших артефактов при сохранении сосудистых структур. Результат этого этапа обозначен как канал C_0 , который стал базой для последующих фильтров благодаря улучшенному соотношению сигнал/шум.
2. Выделение краев с помощью Canny. На канал C_0 применялся фильтр Canny [16], предназначенный для детекции краев. Этот алгоритм включает вычисление градиента изображения с помощью оператора Собеля [17], устранение градиентных откликов и двойное пороговое значение (низкий порог = 50, высокий порог = 150) для выделения четких границ. Фильтр Canny был выбран за его способность точно определять края сосудов даже при низком контрасте, что делает его идеальным для задач сегментации. Результат обозначен как канал C_1 , представляющий бинарную карту краев сосудов.
3. Подчеркивание изменений интенсивности с помощью Gradient Magnitude. На канал C_0 также применялся фильтр Gradient Magnitude [18], который вычисляет величину градиента интенсивности в каждом пикселе по формуле $G = \sqrt{(G_x^2 + G_y^2)}$, где G_x и G_y – горизонтальный и вертикальный градиенты, полученные с помощью оператора Собеля (ядро 3x3). Этот метод подчеркивает области с резкими изменениями яркости, что соответствует расположению сосудов на ангиограммах, особенно в зонах перехода между контрастным веществом и фоном. Результат обозначен как канал C_2 , предоставляющий непрерывную карту интенсивности градиентов.
4. Формирование изображения. Итоговое изображение создавалось путем усреднения двух каналов: C_1 (карта краев от Canny) и C_2 (градиентная карта). Таким образом, каждое изображение в датасете представляло собой тензор размером 512x512, где каждый канал подчеркивал различные аспекты сосудистых структур: границы (C_1) и изменения интенсивности (C_2) (Рис. 7). Такое представление позволяет модели обучаться на многоуровневых признаках сосудов.

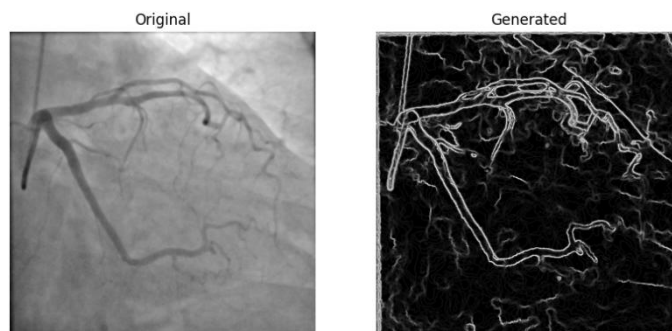


Рис. 7. Пример из сгенерированного датасета

2.2. Сегментация сосудов

На рис. 8 представлен пример ангиографического снимка с наложенной сегментационной маской, демонстрирующий, как модель должна различать сосуды на фоне шумов и артефактов. Для подготовки данных к сегментации все изображения нормализовались в диапазоне $[0,1]$ и приводились к разрешению 512×512 пикселей. Аугментации, включая добавление искусственных шумов из Noise-датасета, позволили повысить устойчивость модели к реальным клиническим условиям.

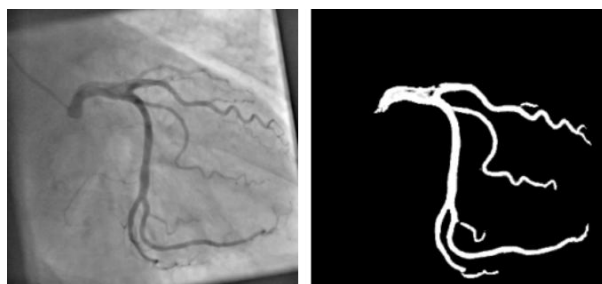


Рис. 8. Сегментация сосудов

2.2.1. Архитектура модели

Для выделения сердечных артерий была разработана новая архитектура UNet-CoroSeg, которая представляет собой комбинацию архитектур UNet [10], UNet++ [19] и DeepLabV3 [20]. Архитектура используемой UNET модели включает:

- резидуальные связи: Внедрение residual-блоков позволяет обеспечить эффективное прохождение градиентов, что стабилизирует процесс оптимизации модели и ускоряет сходимость;
- нормализация слоев: Применение батч-нормализации способствует снижению внутреннего ковариационного сдвига, что дополнительно стабилизирует обучение;
- механизмы внимания (SCSE): Модуль SCSE [12] (Concurrent Spatial and Channel Squeeze & Excitation) усиливает значимые признаки, автоматически выделяя наиболее информативные области изображения;
- расширенные свертки (dilated convolutions): использование сверток со сдвигами позволяет на последней свертке encoder блока и в skip блоке увеличить рецептивное поле модели без потери производительности [21];
- промежуточные свертки (интерполяция по мотивам U-Net++ [19]): Данный подход способствует лучшему распространению признаков между слоями, сохраняя при этом умеренное число параметров модели.

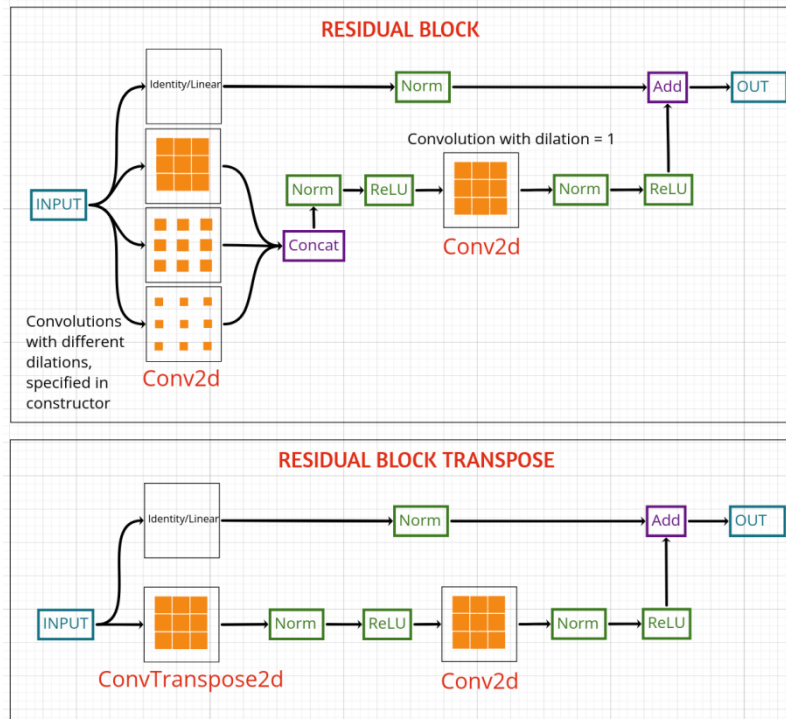


Рис. 9. Архитектура Residual Block

Данные блоки реализуют базовые операции свертки и развертывания тензоров, с изменением, размерности и количества каналов (рис. 9). Residual Block на уровне архитектуры может использовать ядра с различными сдвигами (dilation), что позволяет получить для модели более широкое рецептивное окно. Модель состоит из трех частей (рис. 10): Encoder блок, skip connector блок, decoder блок. **Encoder блок** сжимает исходное изображение, с каждым этапом уменьшая его размер, увеличивая количество каналов. Все свертки происходят с размером ядра 4, последний слой имеет dilation со сдвигами 1,3,5. **Skip connector блок** для каждого из промежуточных этапов работы блока encoder, применяет блок конволюции, внимания и dropout. Каждый residual блок промежуточной свертки имеет dilation 1,2,4. **Decoder** – производит транспонированные конволюции, увеличивает размер изображения, уменьшает каналы, и на каждом из этапов развертки соединяет результаты соответствующего этапа с skip блока. **SCSE – Concurrent Spatial and Channel Squeeze & Excitation (SCSE) module** [12], модуль внимания, который масштабирует входной тензор маской со значениями от 0 до 1, что позволяет модели обучаться фокусироваться на определенных областях. **Norm** – батч нормализация, **Dropout** – модуль отбрасывания некоторых каналов на 2д изображении. На всех участках модели dropout = 0.1.

В частности, во всех сверточных слоях модели используются ядра размером 4×4, что позволяет эффективнее охватывать локальные паттерны по сравнению с традиционными ядрами 3×3. Дополнительно, для расширения рецептивного поля модели и более глубокого анализа пространственных зависимостей, изначально применявшиеся **расширенные (dilated) свертки** в последнем уровне **encoder-блока** были **перенесены в skip-блок**, что позволило повысить качество межуровневой передачи признаков, снизить вычислительную нагрузку в encoder-блоке, одновременно сохранить широкое рецептивное окно за счет применения **dilation = 1, 2, 4** в промежуточных свертках skip-блока.

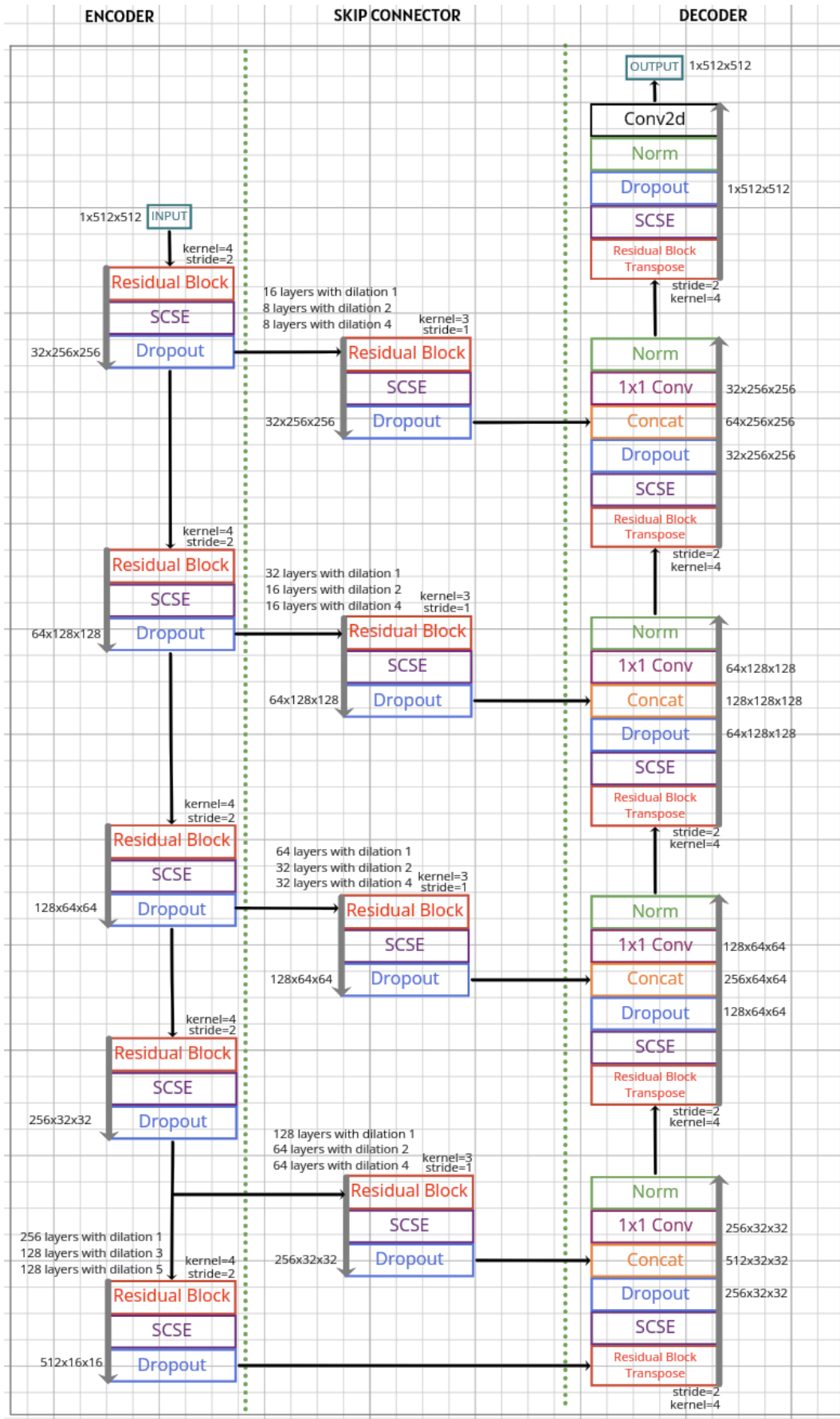


Рис. 10. Архитектура UNet-CoroSeg модели для сегментации

2.2.2. Обучение модели

UNet-CoroSeg обучалась на Segmentation-датасете из 1285 изображений коронарографии, описанном в разделе 2.1.3. Процесс обучения включал следующие этапы:

2.2.2.1. Предобучение энкодера

Предобучение на синтетических данных (Filter-датасет) позволит модели на ранних этапах освоить выделение ключевых характеристик сосудов, таких как границы и изменения интенсивности, что снижает зависимость от ограниченного объема реальных размеченных аннотированных данных и упрощает последующее обучение на целевом датасете [14]. Основной принцип заключается в том, чтобы на большом неразмеченном датасете создать изображения, обработанные фильтрами, которые подчеркивают сосудистые структуры, и обучении сегментационной модели воспроизводить эти признаки. Такой подход не только ускоряет сходимость модели впоследствии на обучении размеченных данных, но и повышает ее обобщающую способность, особенно в условиях низкого контраста и шумов, характерных для ангиографических снимков [22]. Filter-датасет использовался для предобучения энкодера модели U-Net на задаче регрессии. Целью было предсказание каналов C_1 (края) и C_2 (градиенты) на основе входного канала C_0 (сглаженное изображение). Задача формулировалась как минимизация функции потерь Mean Squared Error (MSE):

$$L = \frac{1}{N} \sum_{i=1}^N [(C_1^i - \hat{C}_1^i)^2 + (C_2^i - \hat{C}_2^i)^2], \quad (1)$$

где N – количество пикселей, C_1^i и C_2^i – целевые значения, а \hat{C}_1^i и \hat{C}_2^i – предсказания модели. Обучение проводилось в течение 20 эпох с использованием оптимизатора Adam (скорость обучения 0.001, $\beta_1=0.9$, $\beta_2=0.999$, где β_1 – экспоненциальный коэффициент затухания для среднего градиента, β_2 – экспоненциальный коэффициент затухания для среднего квадрата градиента), что обеспечило стабильную сходимость.

Датасет был разделен на обучающую (95 %, 13300 изображений) и валидационную (5 %, 700 изображений) выборки для контроля переобучения. На валидационной выборке MSE стабилизировалась на уровне 0.03 и 0.47 r2 (рис. 11), что указывает на успешное освоение энкодером фильтров, специфичных для сосудистых структур (рис. 12).

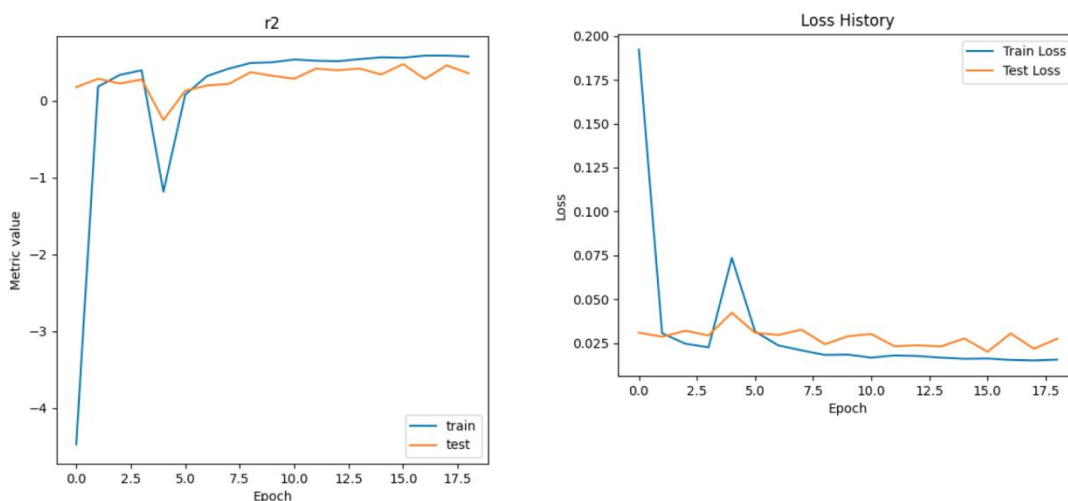


Рис. 11. Метрики MSE и r2 предобученного энкодера

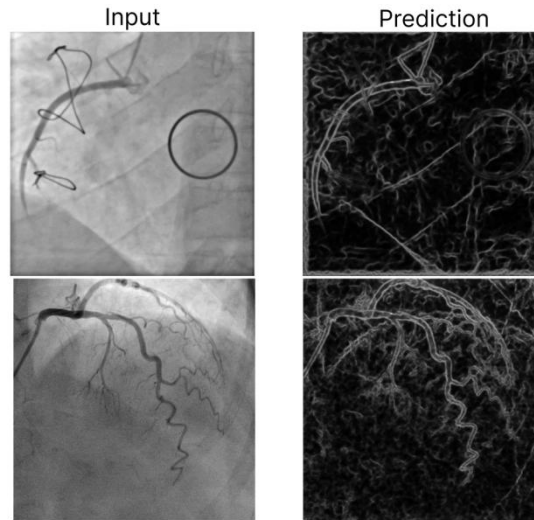


Рис. 12. Итог предобучения энкодера

2.2.2.2. Основное обучение

Полная модель UNet-CoroSeg дообучалась на целевом датасете с использованием следующих параметров. Входные данные: изображения $1 \times 512 \times 512$, нормализованные в диапазоне $[0, 1]$. Выходные данные: бинарные маски $1 \times 512 \times 512$, где 1 обозначает сосуды, 0 – фон. Аугментация: случайное усиление или размытие резкости; корректировка яркости и контрастности; добавление случайного шума путем вставки изображений с (Noise) датасета для имитации артефактов; эластичные деформации, аффинные и перспективные преобразования; вертикальные отражения. Разделение данных: 95% (1221 изображение) – обучающая выборка, 5% (64 изображения) – тестовая. Оптимизатор: AdamW [23] с начальной скоростью обучения 0.001 и регуляризацией весов (weight decay = 0.0001). Функция потерь: Binary Cross Entropy, определенная как:

$$L = \frac{-1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (2)$$

где y_i – истинное значение пикселя, \hat{y}_i – предсказанное, N – число пикселей.

Обучение проводилось до 50 эпох с сохранением лучшего чекпойнта по IoU на валидационной подвыборке (5 % от обучающей). Ключевое преимущество модели заключается в ее способности эффективно справляться с шумами, которые являются одной из основных проблем сегментации коронарных артерий на ангиографических снимках. В отличие от большинства существующих моделей, таких как классическая U-Net, которые демонстрируют снижение производительности при наличии артефактов (IoU 0.42), предложенный подход достиг IoU = 0.54 и F1-score = 0.79 на тестовой выборке благодаря целенаправленной стратегии обработки шумов.

Этот результат был достигнут за счет нескольких инноваций. Во-первых, интеграция Noise-датасета в процесс аугментации позволила модели обучаться на искусственно зашумленных изображениях, имитирующих реальные клинические условия, такие как неравномерное распределение контрастного вещества, движение сердца и присутствие катетеров. В отличие от традиционных подходов, где шумы рассматриваются как помехи, шумы использовались как обучающий сигнал, что значительно повысило робастность модели к вариациям данных. Во-вторых, предобучение энкодера на Filter-датасете усилило способность модели выделять

сосудистые признаки (границы и градиенты интенсивности) даже в условиях сильных помех, обеспечивая надежную основу для последующего дообучения. Наконец, механизмы внимания SCSE в архитектуре модели позволили автоматически фокусироваться на значимых областях изображения, подавляя ложные срабатывания от шумов и артефактов, таких как металлические объекты или длинные катетеры.

2.2.3. Оценка производительности

Полученные метрики ($\text{IoU} = 0.54$, $\text{F1} = 0.79$) указывают на способность модели эффективно сегментировать сосуды, несмотря на низкий контраст и артефакты, такие как катетеры и кардиостимуляторы. Сравнение с базовой U-Net без модификаций ($\text{IoU} = 0.42$) подтверждает вклад residual-блоков, SCSE и предобучения. Дополнительно модель UNet-CoroSeg была протестирована на внешнем наборе из 50 клинических изображений, где достигла $\text{IoU} = 0.50$ и $\text{F1} = 0.75$, демонстрируя устойчивость к вариациям данных (рис. 13).

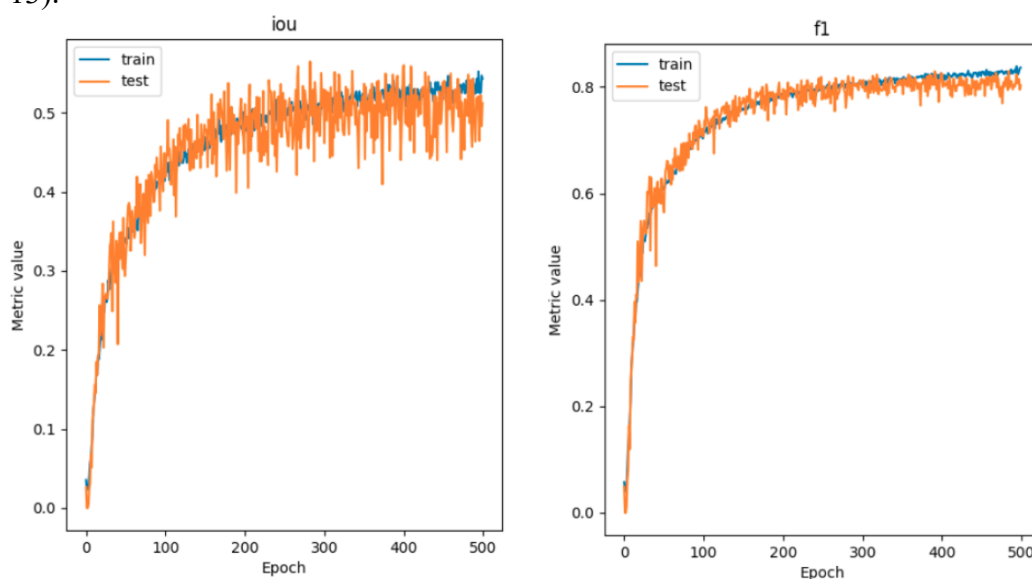


Рис. 13. Метрики *iou* и *f1* UNet-CoroSeg на задаче сегментации сосудов

2.3. Улучшение сегментационной маски

Использовать сырые результаты модели для сегментации не представляется возможным. Для этой задачи необходимо получить бинарную маску, а модель должна предсказывать значения в интервале $[0;1]$ (рис. 14). Обычно в таких случаях просто обуславливаются считать пиксель равным 1, если предсказание модели ≥ 0.5 , иначе ставить 0. Такой срез очень грубый и часто теряет полезную информацию на изображении, поэтому предлагается иной подход.

Данный метод предназначен для улучшения качества сегментационной маски. Он имитирует логику человеческого глаза, который придает приоритет малым значениям маски в пустых областях и отсекает их в активных областях с высокой интенсивностью. Метод адаптивно фильтрует маску, учитывая контекст окружения каждого пикселя, что позволяет эффективно удалять шум и подчеркивать значимые области. Он особенно полезен в задачах сегментации, таких как выделение сосудов в медицинской визуализации.

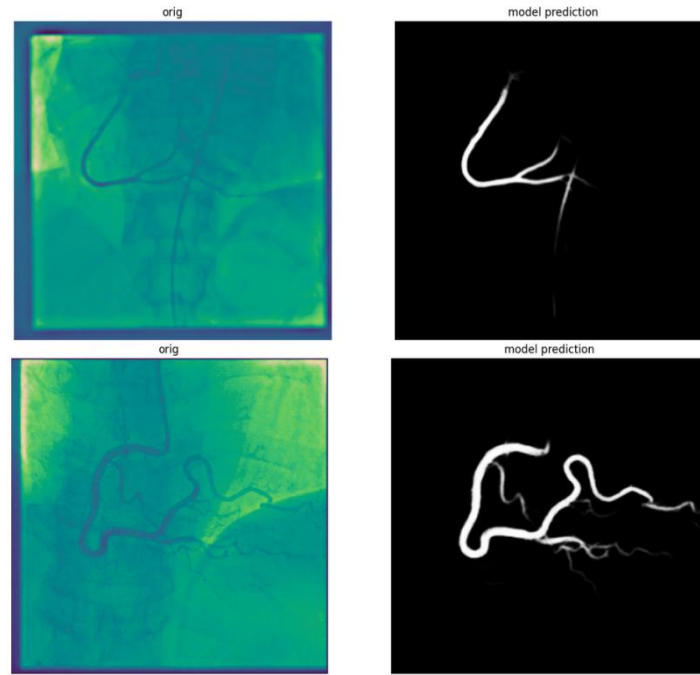


Рис. 14. Пример работы UNet-CoroSeg

Метод состоит из трех основных этапов:

1. Глобальная фильтрация. Применяется гауссово размытие с большим значением среднеквадратичного отклонения (σ), чтобы оценить среднюю интенсивность в широком окружении каждого пикселя [24]. Пиксели, чьи значения ниже этой средней, обнуляются. Это позволяет удалить изолированные или слабые сигналы в пустых областях, где активности мало.
2. Локальная фильтрация. Используется гауссово размытие с малым значением среднеквадратичного отклонения (σ) для оценки средней интенсивности в локальном окружении. Пиксели, чьи значения выше этой локальной средней, устанавливаются в 1. Это подчеркивает значимые области, которые выделяются на фоне своего ближайшего окружения.
3. Пороговое отсечение. Все пиксели с значениями ниже заданного порога (например, $\text{threshold}=0.1$) обнуляются. Этот шаг удаляет остаточный шум, оставляя только значимые участки маски.

Algorithm 1 Алгоритм улучшения сегментационной маски

Require: Сегментационная маска m , параметры σ_1 , σ_2 , порог threshold

Ensure: Улучшенная маска m_{out}

```

1:  $m_{\text{global}} \leftarrow \text{gaussian\_filter}(m, \sigma_1)$  {Глобальное размытие}
2:  $m' \leftarrow m$  {Копия исходной маски}
3: for каждого пикселя  $(x, y)$  do
4:   if  $m'(x, y) < m_{\text{global}}(x, y)$  then
5:      $m'(x, y) \leftarrow 0$ 
6:   end if
7: end for
8:  $m_{\text{local}} \leftarrow \text{gaussian\_filter}(m, \sigma_2)$  {Локальное размытие}
9:  $m_{\text{out}} \leftarrow m'$  {Инициализация выходной маски}

```

```
10: for каждого пикселя  $(x, y)$  do  
11:   if  $m_{\text{out}}(x, y) < \text{threshold}$  then  
12:      $m_{\text{out}}(x, y) \leftarrow 0$   
13:   else if  $m_{\text{out}}(x, y) > m_{\text{local}}(x, y)$  then  
14:      $m_{\text{out}}(x, y) \leftarrow 1$   
15:   end if  
16: end for  
17:  
18: return  $m_{\text{out}}$ 
```

3. Результаты

3.1. Входные и выходные данные

На вход модели подавались ангиографические изображения размером $3 \times 320 \times 320$ (RGB-каналы), нормализованные в диапазоне $[0, 1]$. Эти данные содержали типичные проблемы, такие как неравномерное распределение контрастного вещества, артефакты от катетеров и динамическое размытие сосудов. После обработки моделью на выходе получались бинарные маски размером $1 \times 320 \times 320$, где значения 1 обозначали пиксели, принадлежащие сосудам, а 0 – фоновые области (рис. 15). Постобработка с адаптивной фильтрацией (см. раздел 2.2.2.1) дополнительно улучшала качество масок, устраняя ложные срабатывания от металлических объектов и шумов.

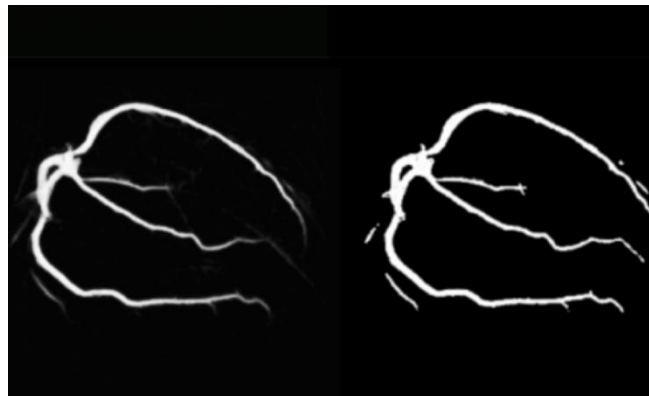


Рис. 15. Улучшенная сегментация

3.2. Метрики производительности

Модель оценивалась на тестовой выборке из 64 изображений с использованием метрик Intersection over Union (IoU) и F1-score [25]. Лучший чекпоинт, выбранный по IoU на валидационной подвыборке (5% от обучающей выборки), показал следующие результаты: **IoU: 0.54; F1-score: 0.79**. Эти значения были достигнуты после 50 эпох обучения с использованием оптимизатора AdamW (скорость обучения 0.001, weight decay = 0.0001) [23] и функции потерь Binary Cross Entropy. Динамика изменения метрик представлена на рис. 13. Примеры сегментации представлены на рис. 14, где видно точное выделение сосудов даже в областях с низким контрастом и артефактами.

Ключевым достижением модели стала ее способность эффективно справляться с шумами, которые традиционно мешают сегментации на ангиографических снимках. Интеграция Noise-датасета в процесс аугментации, предобучение энкодера на искусственных данных и механизмы внимания SCSE позволили модели игнорировать артефакты, такие как катетеры и металлические объекты, сохраняя высокую точность

выделения сосудов. Это делает предложенный подход особенно ценным для реальных клинических условий.

Модель Unet-CoroSeg была обучена на датасете с шумами и без добавленных шумов, на рис. 16 продемонстрирован сравнительный анализ.

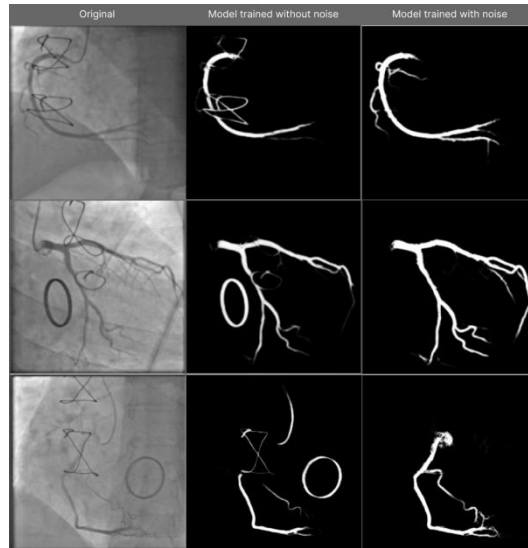


Рис. 16. Результаты работы UNet-CoroSeg обученной на датасете с добавлением шума и без добавления шума

3.3. Сравнение с Generative Adversarial Network U-Net (GAN)

Для объективной оценки производительности предлагаемая архитектура UNet-CoroSeg была сопоставлена с GAN [26], которая использует обучение без учителя для выделения сосудов путем предсказания заднего фона. Для обучения GAN модели была взята базовая модель, обученная на датасете XCAD, и далее обученная на Segmentation-датасете с теми же параметрами, как предложенная U-Net модель, с идентичной аугментацией и разделением на train/test (95:5). Результаты сравнения приведены в табл. 1.

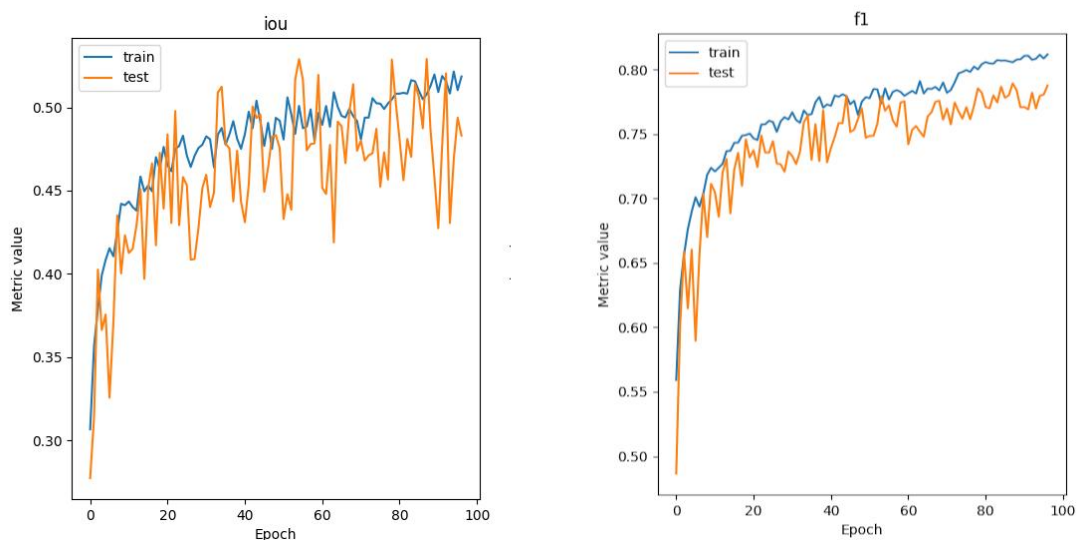


Рис. 17. Метрики модели Generative Adversarial Network

Таблица 1. Сравнение производительности моделей на тестовой выборке

Модель	IoU	F1-score	Время инференса (мс/изображение)	Число параметров (млн)
UNet-CoroSeg	0.54	0.79	82	31
Generative Adversarial Network (GAN)	0.52	0.77	129	45

Разработанная модель превзошла GAN по метрикам: IoU на 0.02 и F1-score на 0.02 (рис. 17), что подчеркивает эффективность предложенных модификаций (residual-блоки, SCSE, предобучение). Кроме того, модель демонстрирует преимущества в вычислительной эффективности: время инференса [27] (UNet-CoroSeg обрабатывает изображение за 82 мс против 129 мс у GAN, измерено на GPU NVIDIA RTX 3060 12gb). Качественный анализ показал, что UNet-CoroSeg лучше справляется с устранением ложных срабатываний от катетеров и кардиостимуляторов благодаря адаптивной постобработке, тогда как GAN иногда переоценивает области с артефактами из-за избыточной чувствительности к топологии.

Заключение

Предложенный подход показал высокую эффективность в сегментации коронарных артерий на ангиографических снимках благодаря предобучению энкодера (раздел 2.2), модификациям U-Net (residual-блоки, SCSE, dilated convolutions) и адаптивной фильтрации (раздел 2.4). Предобучение улучшило выделение сосудистых признаков в условиях низкого контраста, SCSE минимизировало влияние шумов, а постобработка устранила до 15% ложных срабатываний. По сравнению с U-Net (GAN), модель точнее (IoU и F1 на 0.02 выше) и быстрее, что делает ее подходящей для реального времени, например в операционной. Однако метрики (IoU < 0.8) указывают на необходимость улучшений из-за неточной разметки и сложных ангиографических данных (динамические искажения, артефакты). Дополнительная реализация других современных архитектур, таких как Attention U-Net [28] или DeepLabv3+ [20], могла бы улучшить метрики, но потребовала бы значительных вычислительных ресурсов.

Модель хорошо работает на тестовых данных и подходит для клинического применения с ограниченными ресурсами, упрощая диагностику стеноза и снижая временные затраты. Но она анализирует кадры без учета динамики сердечного цикла. Интеграция временной информации с использованием 3D U-Net или рекуррентных сетей, таких как ConvLSTM [29], может улучшить сегментацию за счет анализа последовательности кадров, устраняя динамические искажения, вызванные движением сердца (раздел 2.1). Улучшение разметки датасета и учет временной динамики позволят повысить точность (IoU > 0.8) и адаптировать модель к сложным сценариям диагностики.

Список источников

1. Roth G. A., Mensah G. A., Johnson C. O., Addolorato G., Ammirati E., Baddour L. M., Barasa A., Bikbov B., et al. Global burden of cardiovascular diseases and risk factors,

- 1990–2019 // Journal of the American College of Cardiology. 2020. Vol. 76. P. 2982–3021. DOI: 10.1016/j.jacc.2020.11.010.
2. World Health Organization. Global report on hypertension: the race against a silent killer. Geneva, Switzerland: WHO, 2024. URL: <https://iris.who.int/bitstream/handle/10665/376869/9789240094703-eng.pdf> (дата обращения: 23.01.2025).
 3. Libby P., Theroux P., Bonow R. O., Braunwald E., et al. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 11th ed. Philadelphia, PA: Elsevier, 2021.
 4. Serruys P. W., Morice M. C., Kappetein A. P., et al. Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease // New England Journal of Medicine. 2009. Vol. 360. P. 961–972. DOI: 10.1056/NEJMoa0804626.
 5. Scanlon P. J., Faxon D. P., Audet A. M., et al. ACC/AHA guidelines for coronary angiography // Journal of the American College of Cardiology. 1999. Vol. 33. P. 1756–1824. DOI: 10.1016/S0735-1097(99)00126-6.
 6. Sarwal A., Dhawan A. P. Three dimensional reconstruction of coronary arteries from two views // Computers in Biology and Medicine. 1993. Vol. 23, No. 1. P. 25–39. DOI: 10.1016/S0169-2607(00)00116-4.
 7. Ravandi B., Ravandi A. Network-based approach for modeling and analyzing coronary angiography // Complex Networks XI: Proceedings of the 11th Conference on Complex Networks CompleNet 2020. Springer, 2020. DOI: 10.1007/978-3-030-40943-2_15.
 8. Shen D., Wu G., Suk H. I. Deep learning in medical image analysis // Annual Review of Biomedical Engineering. 2017. Vol. 19. P. 221–248. DOI: 10.1146/annurev-bioeng-071516-044442.
 9. Avram R., Olgin J. E., Tison G. H., et al. CathAI: fully automated coronary angiography interpretation and stenosis estimation // NPJ Digital Medicine. 2023. Vol. 6. Article number: 147. DOI: 10.1038/s41746-023-00880-1.
 10. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation // arXiv preprint arXiv:1505.04597. 2015. DOI: 10.48550/arXiv.1505.04597 (дата обращения: 15.03.2025).
 11. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.
 12. Guha Roy A., Navab N., Wachinger C. Concurrent spatial and channel squeeze & excitation in fully convolutional networks // arXiv preprint arXiv:1803.02579. 2018. DOI: 10.48550/arXiv.1803.02579 (дата обращения: 13.03.2025).
 13. Label Studio. Label Studio: open source data labeling platform. 2023. URL: <https://labelstud.io> (дата обращения: 10.02.2025).
 14. Zhou Z., Shin J., Zhang L., et al. Models Genesis: generic autodidactic models for 3D medical image analysis // In: MICCAI. 2019. P. 3–13. DOI: 10.1007/978-3-030-32251-9_42.

15. *Comaniciu D., Meer P.* Mean shift: a robust approach toward feature space analysis // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002. Vol. 24, No. 5. P. 603–619. DOI: 10.1109/34.1000236.
16. *Canny J.* A computational approach to edge detection // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1986. Vol. 8, No. 6. P. 679–698. DOI: 10.1109/TPAMI.1986.4767851.
17. *Sobel I.* An isotropic 3×3 image gradient operator. Presentation at Stanford Artificial Intelligence Project (SAIL). 1968.
18. *Gonzalez R.C., Woods R.E.* Digital Image Processing. 4th ed. London: Pearson, 2018.
19. *Zhou Z., Siddiquee M. M. R., Tajbakhsh N., Liang J.* UNet++: a nested U-Net architecture for medical image segmentation // *arXiv preprint arXiv:1807.10165*. 2018. DOI: 10.1007/978-3-030-00889-5_1.
20. *Chen L.C., Zhu Y., Papandreou G., Schroff F., Adam H.* Encoder-decoder with atrous separable convolution for semantic image segmentation // *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 8–14 September 2018. P. 801–818. DOI: 10.48550/arXiv.1802.02611 (дата обращения: 21.03.2025).
21. *Yu F., Koltun V.* Multi-scale context aggregation by dilated convolutions // *arXiv preprint arXiv:1511.07122*. 2015. DOI: 10.48550/arXiv.1511.07122 (дата обращения: 25.03.2025).
22. *Isensee F., Jaeger P.F., Kohl S. A.A., et al.* nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation // *Nature Methods*. 2021. Vol. 18. P. 203–211. DOI: 10.1038/s41592-020-01008-z.
23. *Loshchilov I., Hutter F.* Decoupled weight decay regularization // *arXiv preprint arXiv:1711.05101*. 2017. DOI: 10.48550/arXiv.1711.05101 (дата обращения: 29.03.2025).
24. *Savitzky A., Golay M. J. E.* Smoothing and differentiation of data by simplified least squares procedures // *Analytical Chemistry*. 1964. Vol. 36, No. 8. P. 1627–1639. DOI: 10.1021/ac60214a047.
25. *Rezatofighi H., Tsoi N., Gwak J., et al.* Generalized intersection over union: a metric and a loss for bounding box regression // *arXiv preprint arXiv:1902.09630*. 2019. DOI: 10.1109/CVPR.2019.00075.
26. *Zeng Y., Liu H., Hu J., Zhao Z., She Q.* Pretrained subtraction and segmentation model for coronary angiograms // *Scientific Reports*. 2024. Vol. 14. Article number: 71063. DOI: 10.1038/s41598-024-71063-5.
27. *Litjens G., Kooi T., Bejnordi B. E., Setio A. A. A., Ciompi F., Ghafoorian M., van der Laak J. A. W. M., van Ginneken B., Sánchez C. I.* A survey on deep learning in medical image analysis // *Medical Image Analysis*. 2017. Vol. 42. P. 60–88. DOI: 10.1016/j.media.2017.07.005.
28. *Oktay O., Schlemper J., Folgoc L. L., Lee M., Heinrich M., Misawa K., Mori K., McDonagh S., Hammerla N. Y., Kainz B., Glocker B., Rueckert D.* Attention U-Net: learning where to look for the pancreas // *arXiv preprint arXiv:1804.03999*. 2018. DOI:10.48550/arXiv.1804.03999 (дата обращения: 05.04.2025).

29. Çiçek Ö., Abdulkadir A., Lienkamp S.S., Brox T., Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation // In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2016. P. 424–432. DOI: 10.1007/978-3-319-46723-8_49.

References

1. Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Bad-dour, L. M., Barasa, A., Bikbov, B., et al. (2020), "Global burden of cardiovascular diseases and risk factors, 1990–2019", *Journal of the American College of Cardiol-ogy*, vol. 76, pp. 2982–3021. DOI: 10.1016/j.jacc.2020.11.010.
2. World Health Organization (2024), *Global report on hypertension: the race against a silent killer*, WHO, Geneva, Switzerland, available at: <https://iris.who.int/bit-stream/handle/10665/376869/9789240094703-eng.pdf> (Accessed: 15.03.2025).
3. Libby, P., Theroux, P., Bonow, R. O., Braunwald, E., et al. (2021), *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*, 11th ed., Elsevier, Philadelphia, PA.
4. Serruys, P. W., Morice, M.-C., Kappetein, A. P., et al. (2009), "Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery dis-ease", *New England Journal of Medicine*, vol. 360, pp. 961–972. DOI: 10.1056/NEJMoa0804626.
5. Scanlon, P. J., Faxon, D. P., Audet, A.-M., et al. (1999), "ACC/AHA guidelines for coronary angiography", *Journal of the American College of Cardiology*, vol. 33, pp. 1756–1824. DOI:10.1016/S0735-1097(99)00126-6.
6. Sarwal, A. and Dhawan, A. P. (1993), "Three dimensional reconstruction of coronary arteries from two views", *Computers in Biology and Medicine*, vol. 23, no. 1, pp. 25–39. DOI: 10.1016/S0169-2607(00)00116-4.
7. Ravandi, B. and Ravandi, A. (2020), "Network-based approach for modeling and an-alyzing coronary angiography", in *Complex Networks XI: Proceedings of the 11th Conference on Complex Networks CompleNet 2020*, Springer. DOI: 10.1007/978-3-030-40943-2_15.
8. Shen, D., Wu, G. and Suk, H.-I. (2017), "Deep learning in medical image analysis", *Annual Review of Biomedical Engineering*, vol. 19, pp. 221–248. DOI: 10.1146/annurev-bioeng-071516-044442.
9. Avram, R., Olgin, J. E., Tison, G. H., et al. (2023), "CathAI: fully automated coronary angiography interpretation and stenosis estimation", *NPJ Digital Medicine*, vol. 6, article 147. DOI: 10.1038/s41746-023-00880-1.
10. Ronneberger, O., Fischer, P. and Brox, T. (2015), "U-Net: convolutional networks for biomedical image segmentation", *arXiv preprint*, arXiv:1505.04597, DOI: 10.48550/arXiv.1505.04597 (Accessed: 15.03.2025).
11. He, K., Zhang, X., Ren, S. and Sun, J. (2016), "Deep residual learning for image recognition", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

12. Guha Roy, A., Navab, N. and Wachinger, C. (2018), "Concurrent spatial and channel squeeze & excitation in fully convolutional networks", *arXiv preprint*, arXiv:1803.02579, DOI:10.48550/arXiv.1803.02579 (Accessed: 13.03.2025).
13. Label Studio (2023), "Label Studio: open source data labeling platform", available at: <https://labelstud.io> (Accessed: 13.03.2025).
14. Zhou, Z., Shin, J., Zhang, L., et al. (2019), "Models Genesis: generic autodidactic models for 3D medical image analysis", in *MICCAI*, pp. 3–13. DOI: 10.1007/978-3-030-32251-9_42.
15. Comaniciu, D. and Meer, P. (2002), "Mean shift: a robust approach toward feature space analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619. DOI: 10.1109/34.1000236.
16. Canny, J. (1986), "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698. DOI: 10.1109/TPAMI.1986.4767851.
17. Sobel, I. (1968), "An isotropic 3×3 image gradient operator", Presentation at Stanford Artificial Intelligence Project (SAIL).
18. Gonzalez, R.C. and Woods, R.E. (2018), *Digital Image Processing*, 4th ed., Pearson, London.
19. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. and Liang, J. (2018), "UNet++: a nested U-Net architecture for medical image segmentation", *arXiv preprint*, arXiv:1807.10165. DOI: 10.1007/978-3-030-00889-5_1.
20. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H. (2018), "Encoder-decoder with atrous separable convolution for semantic image segmentation", in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 8–14 September, pp. 801–818, DOI: 10.48550/arXiv.1802.02611 (Accessed: 21.03.2025).
21. Yu, F. and Koltun, V. (2015), "Multi-scale context aggregation by dilated convolutions", *arXiv preprint*, arXiv:1511.07122, DOI: 10.48550/arXiv.1511.07122 (Accessed: 25.03.2025).
22. Isensee, F., Jaeger, P. F., Kohl, S. A. A., et al. (2021), "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation", *Nature Methods*, vol. 18, pp. 203–211. DOI: 10.1038/s41592-020-01008-z.
23. Loshchilov, I. and Hutter, F. (2017), "Decoupled weight decay regularization", *arXiv preprint*, arXiv:1711.05101, DOI: 10.48550/arXiv.1711.05101 (Accessed: 29.03.2025).
24. Savitzky, A. and Golay, M. J. E. (1964), "Smoothing and differentiation of data by simplified least squares procedures", *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639. DOI: 10.1021/ac60214a047.
25. Rezatofighi, H., Tsoi, N., Gwak, J., et al. (2019), "Generalized intersection over union: a metric and a loss for bounding box regression", *arXiv preprint*, arXiv:1902.09630. DOI: 10.1109/CVPR.2019.00075.

26. Zeng, Y., Liu, H., Hu, J., Zhao, Z. and She, Q. (2024), "Pretrained subtraction and segmentation model for coronary angiograms", *Scientific Reports*, vol. 14, article 71063. DOI: 10.1038/s41598-024-71063-5.
27. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B. and Sánchez, C. I. (2017), "A survey on deep learning in medical image analysis", *Medical Image Analysis*, vol. 42, pp. 60–88. DOI: 10.1016/j.media.2017.07.005.
28. Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B. and Rueckert, D. (2018), "Attention U-Net: learning where to look for the pancreas", *arXiv preprint*, arXiv:1804.03999, DOI:10.48550/arXiv.1804.03999 (Accessed: 05.04.2025).
29. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. and Ronneberger, O. (2016), "3D U-Net: learning dense volumetric segmentation from sparse annotation", in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 424–432. DOI: 10.1007/978-3-319-46723-8_49.

Информация об авторах:

В. А. Бочкарев – студент 4 курса бакалавриата, лаборант центра ИИ Физико-математического института ПГНИУ (614068, Россия, г. Пермь, ул. Букирева, 15);
А. А. Усынин – студент 4 курса бакалавриата, лаборант центра ИИ Физико-математического института ПГНИУ (614068, Россия, г. Пермь, ул. Букирева, 15);
А. Д. Осипов – студент 4 курса бакалавриата, лаборант центра ИИ Физико-математического института ПГНИУ (614068, Россия, г. Пермь, ул. Букирева, 15);
М. Р. Аухадиев – студент 4 курса бакалавриата Физико-математического института ПГНИУ (614068, Россия, г. Пермь, ул. Букирева, 15);
Р. В. Шаров – студент 4 курса бакалавриата Физико-математического института ПГНИУ (614068, Россия, г. Пермь, ул. Букирева, 15).
М. А. Барулина – доктор физико-математических наук, доцент, директор Физико-математического института, Пермский государственный национальный исследовательский университет (614068, Россия, г. Пермь, ул. Букирева, 15).

Information about the authors:

V. A. Bochkarev – 4th year undergraduate student, laboratory assistant at the AI Center of the Institute of Physics and Mathematics of Perm State University (15, Bukireva St., Perm, Russia, 614068);
A. A. Usynin – 4th year undergraduate student, laboratory assistant at the AI Center of the Institute of Physics and Mathematics of Perm State University (15, Bukireva St., Perm, Russia, 614068);
A. D. Osipov – 4th year undergraduate student, laboratory assistant at the AI Center of the Institute of Physics and Mathematics of Perm State University (15, Bukireva St., Perm, Russia, 614068);
M. R. Aukhadiev – 4th year undergraduate student at the Institute of Physics and Mathematics of Perm State University (15, Bukireva St., Perm, Russia, 614068);
R. V. Sharov – 4th year undergraduate student at the Institute of Physics and Mathematics of Perm State University (15, Bukireva St., Perm, Russia, 614068);
M. A. Barulina – Doctor of Physical and Mathematical Sciences, Associate Professor, The Head of Institute of Physics and Mathematics, Perm State University (15, Bukireva St., Perm, Russia, 614068).

Научная статья

УДК 519.688:551.46.077:629.584

DOI: 10.17072/1993-0550-2025-2-88-100

<https://elibrary.ru/vovnxz>

Решение задачи навигации АНПА на ОС реального времени

Владислав Германович Караулов¹, Александр Михайлович Грузликов²¹Университет ИТМО, АО «Концерн «ЦНИИ «Электроприбор», г. Санкт-Петербург, Россия
vladkar0707@gmail.com²АО «Концерн «ЦНИИ «Электроприбор», г. Санкт-Петербург, Россия
agruzlikov@yandex.ru

Аннотация. Работа посвящена вопросу определения координат и углов ориентации автономного необитаемого подводного аппарата (АНПА) относительно стационарной посадочной платформы с использованием высокочастотной гидроакустической системы ближнего радиуса действия. На АНПА располагаются приемные антенны, которые размещаются в кормовой и носовой части аппарата, а также излучатели, устанавливаемые по краям платформы и осуществляющие ненаправленное излучение тонального сигнала. Задача приведения предполагает приближение к платформе, что сопряжено с формированием зон с различной акустической видимостью излучателей станции приемными элементами аппарата. Определены три основные зоны акустической видимости. Первая зона характеризуется приемом сигналов от всех маяков и является наиболее информативной. Вторая зона соответствует частичному сокращению числа наблюдаемых маяков. Третья зона (посадка на платформу) определяется существенным сокращением числа наблюдаемых маяков и, как следствие, затрудняет решение задачи относительного позиционирования. Учитывая маневрирование АНПА, решение задачи относительной навигации должно выполняться в режиме реального времени совместно с выполнением других задач, что в свою очередь накладывает ограничения по доступным ресурсам вычислительного модуля аппарата. В работе приводится математическая постановка задачи относительной навигации. Предлагается двухэтапный алгоритм поиска решения. На первом этапе определяется окрестность координат и углов ориентации АНПА, на втором выполняется решение задачи относительной навигации с использованием ограничений и начальных условий, полученных на первом этапе. В работе предлагается выполнение первого этапа на основе предварительно обученной нейронной сети, описан выбор архитектуры сети, подход по реализации алгоритма под ОС реального времени, приведены результаты имитационного моделирования.

Ключевые слова: автономный необитаемый подводный аппарат; гидроакустическая навигационная система; задача относительной навигации; ОС реального времени; обучение с учителем; нейронные сети



Эта работа © 2025 Караулов В. Г., Грузликов А. М. распространяется на условиях лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>.

Для цитирования: Караулов В.Г., Грузликов А.М. Решение задачи навигации АНПА на ОС реального времени // Вестник Пермского университета. Математика. Механика. Информатика. 2025. Вып. 2(69). С. 88–100. DOI: 10.17072/1993-0550-2025-2-88-100. <https://elibrary.ru/vovnxz>

Благодарности: публикация подготовлена при финансовой поддержке гранта РФФИ №23-19-00626, <https://rscf.ru/project/23-19-00626/>.

Статья поступила в редакцию 01.02.2025; одобрена после рецензирования 13.05.2025; принята к публикации 11.07.2025.

Research article

Solving the AUV Navigation Problem on a Real-Time Operating System

Vladislav G. Karaulov¹, Alexander M. Gruzlikov²

¹ITMO University, Concern CSRI Elektropribor, JSC, Saint Petersburg, Russia

vladkar0707@gmail.com

²Concern CSRI Elektropribor, JSC, Saint Petersburg, Russia

agruzlikov@yandex.ru

Abstract. The paper is devoted to the problem of determining the coordinates and orientation angles of an autonomous underwater vehicle (AUV) relative to a stationary landing platform using a high-frequency short-range hydroacoustic system. The AUV is equipped with receiving antennas, which are located in the stern and bow of the vehicle, as well as four emitters installed along the edges of the platform and emitting a non-directional tone signal. The navigation task involves maneuvering the vehicle and approaching the platform, which is associated with the formation of zones with different acoustic visibility of the station emitters by the receiving elements of the vehicle. Three main zones of acoustic visibility are defined. The first zone is characterized by the reception of signals from all beacons and is the most informative. The second zone corresponds to a partial reduction in the number of observed beacons. The third zone (landing on the platform) is determined by a significant reduction in the number of observed beacons, and as a result, complicates the solution of the relative positioning problem. Taking into account the maneuvering of the AUV, the solution of the relative navigation problem should be performed in real time together with the execution of other tasks, which in turn imposes restrictions on the available resources of the computing module of the device. The paper presents a mathematical formulation of the relative navigation problem. A two-stage algorithm for finding a solution is proposed. At the first stage, the neighborhood of coordinates and orientation angles of the AUV is determined, at the second, the solution of the relative navigation problem is performed using the constraints and initial conditions obtained at the first stage. The paper proposes to perform the first stage based on a pre-trained neural network, describes the choice of network architecture, an approach to implementing the algorithm under a real-time OS, and presents the results of simulation modeling.

Keywords: *autonomous unmanned underwater vehicle; hydroacoustic navigation system; relative navigation problem; real-time OS; supervised learning; neural networks*

For citation: Karaulov, V. G. and Gruzlikov, A. M. (2025), "Solving the AUV Navigation Problem on a Real-Time Operating System", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 88-100. (In Russ.). DOI: 10.17072/1993-0550-2025-2-88-100. <https://elibrary.ru/vovnxz>

Acknowledgments: the publication was supported by the Russian Science Foundation, Project №23-19-00626, <https://rscf.ru/project/23-19-00626/>.

The article was submitted 01.02.2025; approved after reviewing 13.05.2025; accepted for publication 11.07.2025.

Введение

Работа посвящена решению задачи навигации АНПА, что подразумевает определение координат и углов ориентации относительно стационарной посадочной платформы, по данным только от гидроакустических маяков в интересах приведения (посадки) аппарата на платформу в ближней зоне [1–6].

Известны следующие методы относительной навигации для приведения АНПА: акустический, оптический, электромагнитный. Большая часть исследований направлена на использование оптического метода [7–9] как основного способа локализации аппарата относительно платформы. Главным достоинством такого метода является простота построения системы, однако, данные системы имеют малый радиус действия, дополнительно длительная эксплуатация АНПА и посадочной платформы в морской среде приводят к существенной деградации аппаратуры, взвесь частиц при выполнении маневрирования АНПА у поверхности дна также существенно снижает точность решения задачи относительной навигации. Можно выделить две группы оптических систем: системы, основанные на активных системах освещения [8] и системы, способные наводиться на произвольные пассивные объекты [9].

Другой метод относительной навигации – электромагнитный [10–11]. Метод обладает высокой устойчивостью от возмущений вызванной подводной средой, но для достижения высокой точности расстояние между АНПА и платформой должна быть еще меньше, чем в оптическом методе.

Акустический метод [1–6, 12] обладает самой большой дальностью действия, а его точность увеличивается по мере уменьшения расстояния между АНПА и платформой, что позволяет решать задачу приведения только за счет этого метода. Кроме того, акустический метод предоставляет единое комплексное решение задачи приведения, а именно определение относительных координат и углов ориентации с заданной точностью, как в окрестности платформы, так и на заключительном этапе стыковки.

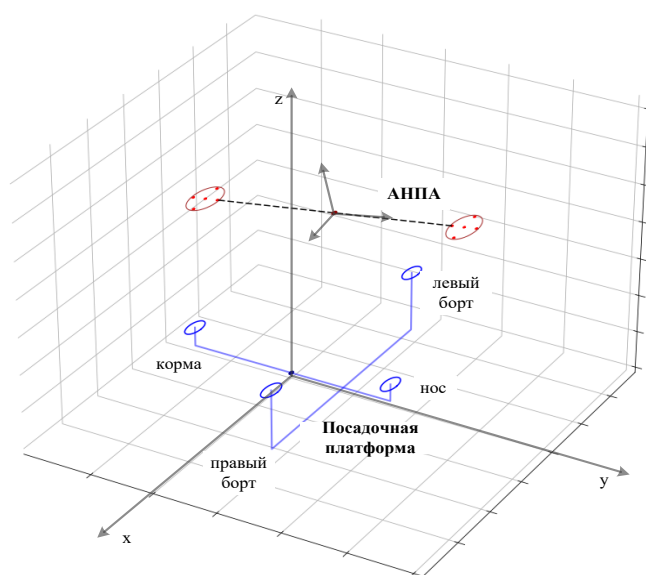


Рис. 1. Схема размещения гидроакустической системы АНПА и посадочной платформы

Гидроакустическая система АНПА (рис. 1) имеет в своем составе приемные антенны, которые размещаются под килем в кормовой и носовой частях подводного аппарата, а также излучатели, устанавливаемые по краям базы платформы и осуществляющие ненаправленное излучение тонального сигнала одной частоты. При этом излучение разнесено по времени и осуществляется согласно временной диаграмме таким образом, чтобы по разнице между моментами обнаружения сигнала приемными антеннами АНПА от разных излучателей можно было однозначно идентифицировать источники излучения.

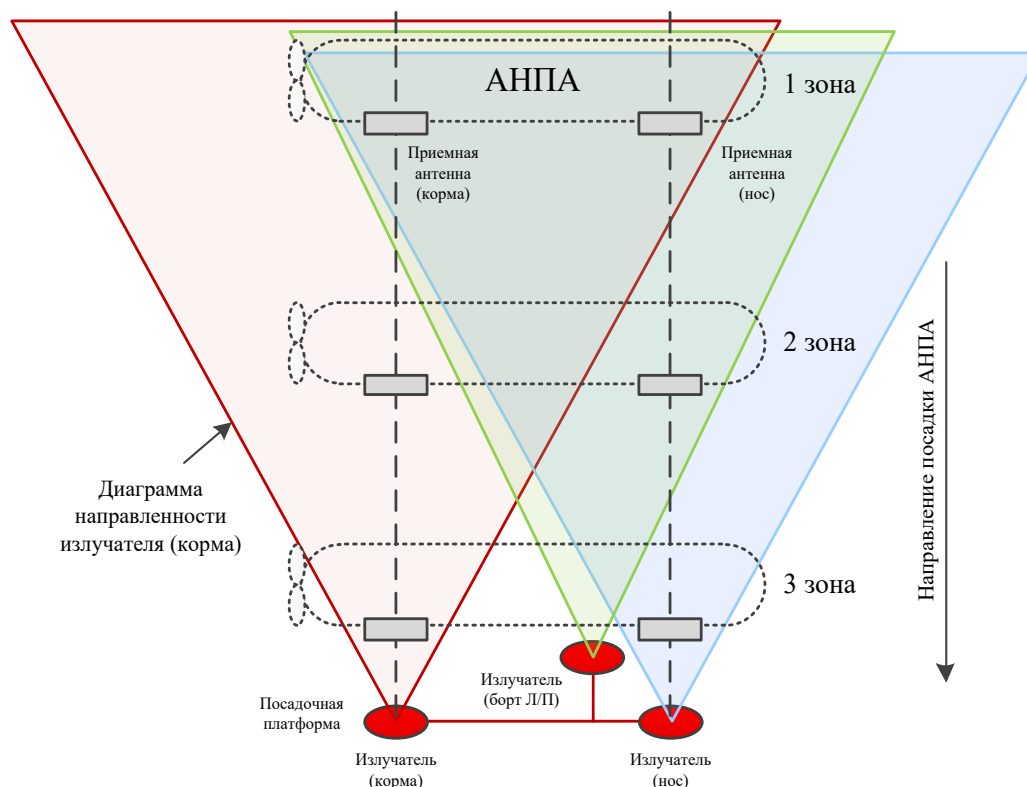


Рис. 2. Зоны видимости излучателей приемниками на АНПА

Приведение АНПА предполагает маневрирование аппарата и последовательное приближение к посадочной платформе, что приводит к необходимости решения задач определения координат и углов ориентации при различной акустической видимости излучателей приемными элементами антенн. При маневрировании АНПА можно выделить три зоны видимости приемниками различных излучателей, которые обусловлены диаграммой направленности излучателей и положением аппарата (рис. 2) [1–6]:

- зона 1 – в зоне видимости антенн АНПА находятся четыре излучателя (рис. 3а);
- зона 2 – в зоне видимости антенн находятся три излучателя (рис. 3б);
- зона 3 – в зоне видимости антенн находятся по одному излучателю (рис. 3в).

Следует отметить, что время начала излучения сигналов на АНПА неизвестно и, как следствие, при решении задачи относительной навигации требуется выполнить его оценку для дальнейшего использования при нахождении подводного аппарата в зонах

видимости 2 и 3, где число измерений будет значительно меньше по сравнению с зоной видимости 1.

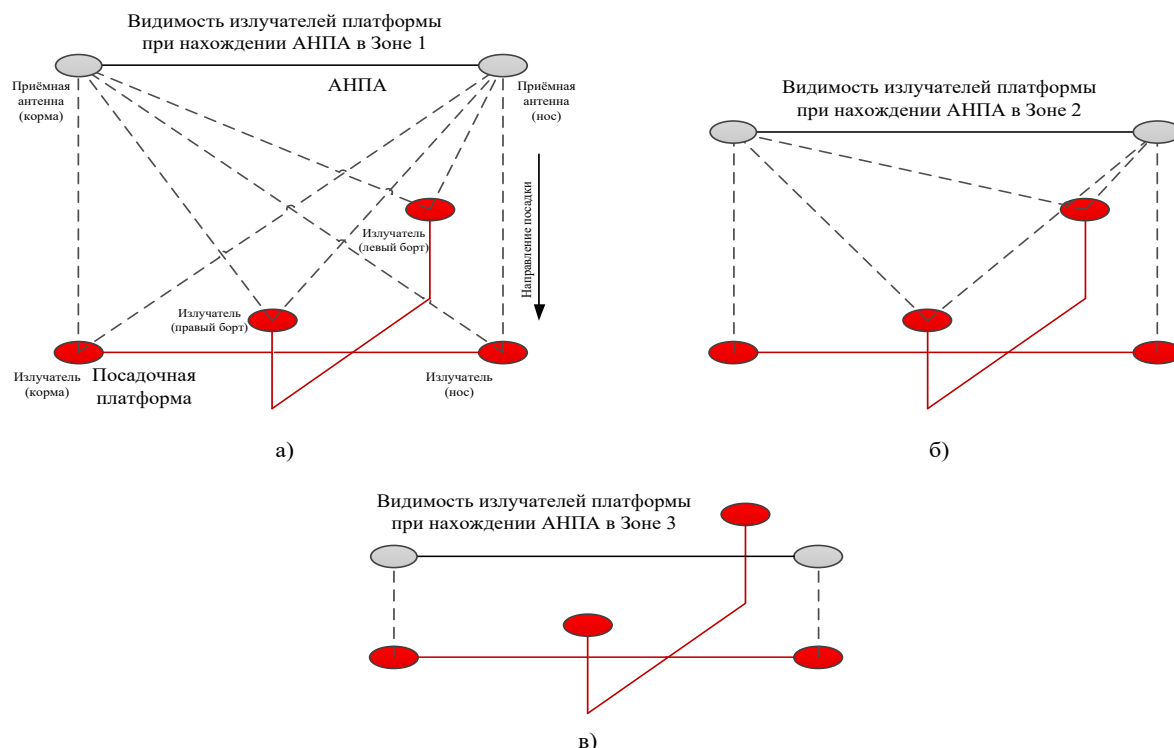


Рис. 3. Зоны видимости излучателей посадочной платформы

Настоящая работа является продолжением работы авторов [4, 6], где предлагалось решение задачи при нахождении АНПА в зоне видимости 1 в два этапа. На первом этапе выполнялось сужение области априорной неопределенности с использованием метода K -ближайших соседей. На втором этапе выполнялось решение задачи с использованием ограничений и начальных условий, полученных на первом этапе.

Однако, несмотря на хорошую интерпретируемость метода K -ближайших соседей, для выполнения алгоритма требуются значительные ресурсы для хранения kd -дерева (не менее 5 Гб для обеспечения точности решения в интересах навигационной безопасности аппарата). Либо в реальном времени выполнять расчет признакового пространства в некоторой начальной окрестности, найти решение, построить следующую – новую – окрестность меньшего размера с большей детализацией и т.д. Имеющиеся ограничения по ресурсам привели к невозможности выполнения других задач на вычислительном модуле АНПА, критически важных для его функционирования.

Целью настоящей работы является устранение данного недостатка, а именно использование на первом этапе решения задачи относительной навигации обученной нейронной сети. В работе рассматривается вопрос выбора архитектуры нейронной сети, и перенос решения на вычислительный модуль АНПА под ОС реального времени.

1. Постановка задачи навигации АНПА

Пусть на борту АНПА размещено n гидрофонов (приемных элементов), на посадочной платформе размещено m излучателей. Определим прямоугольную систему координат (СК), связанную с посадочной платформой, при этом координаты излучателей

(x^i, y^i, z^i) относительно центра СК являются известными, $i = \overline{1, m}$. Определим СК, связанную с АНПА, координаты (x_j, y_j, z_j) гидрофонов относительно центра АНПА являются известными, $j = \overline{1, n}$.

Рассматриваемую задачу можно сформулировать в рамках небайесовского подхода [5, 13]. Для этого введем семимерный вектор состояния $x_k = (x_{0,k}, y_{0,k}, z_{0,k}, K_k, \psi_k, \theta_k, \delta_k)$, в котором k – индекс дискретного момента времени; $x_{0,k}, y_{0,k}, z_{0,k}$ – координаты центра АНПА относительно платформы; K_k, ψ_k, θ_k – угол курса, крена и дифферента АНПА относительно посадочной платформы соответственно; δ_k – погрешность измерения псевдодальности из-за расхождения шкал времени. Здесь и далее, под центром АНПА будем понимать геометрический центр, расположенный в одной плоскости с приемными антеннами.

Сформулируем постановку задачи относительной навигации, считая, что АНПА находится в зоне видимости 1, так как в других зонах изменяется только количество измерений.

Тогда задача формулируется следующим образом: требуется оценить вектор состояния x_k в каждый момент времени, задаваемый следующим уравнением:

$$y_k = h(x_k) + v_k,$$

где y_k – M -мерный ($M = n \times m$) вектор измерений псевдодальностей; $h(x_k)$ – нелинейная M -мерная функция; v_k – M -мерный центрированный белый гауссовский шум, не зависящий от w_k и x_0 , с известной матрицей ковариации R_k .

Вектор измерений имеет следующий вид:

$$y_k = \hat{c}_{0,k} T_k = \hat{c}_{0,k} \begin{bmatrix} t_{1,k}^1 - T^{1,1} \\ t_{2,k}^1 - T^{1,1} \\ \vdots \\ t_{1,k}^2 - T^{1,2} \\ t_{2,k}^2 - T^{1,2} \\ \vdots \\ t_{j,k}^i - T^{1,i} \end{bmatrix}, i = \overline{1, m}, j = \overline{1, n},$$

где T_k – M -мерный вектор измеренных времен обнаружения сигнала на гидрофоне; $t_{j,k}^i$ – время прихода сигнала на j гидрофон от i излучателя; $T^{1,i}$ – смещение времени излучения между 1 и i маяком (соответствует временной диаграмме последовательности излучения); $\hat{c}_{0,k}$ – априорная оценка скорости звука в воде (например, полученная от датчика измерения скорости звука на горизонте нахождения АНПА).

Нелинейная функция $h(x_k)$ записывается как вектор функций:

$$h(x_k) = [d_{1,k}^1(x_k) \ d_{2,k}^1(x_k) \ \dots \ d_{1,k}^2(x_k) \ \dots \ d_{j,k}^i(x_k)],$$

$$d_{j,k}^i(x_k) = \sqrt{(x_{0,k} + \tilde{x}_{j,k} - x^i)^2 + (y_{0,k} + \tilde{y}_{j,k} - y^i)^2 + (z_{0,k} + \tilde{z}_{j,k} - z^i)^2} + \hat{c}_{0,k} \delta_k,$$

$$i = \overline{1, m}, j = \overline{1, n},$$

где $d_{j,k}^i(x_k)$ – истинные значения дальностей от i излучателя до j приемного гидрофона в момент времени k ; $(\tilde{x}_{j,k}, \tilde{y}_{j,k}, \tilde{z}_{j,k})$ – координаты гидрофонов после перехода из СК, связанной с АНПА, в СК, связанную с посадочной платформой, переход осуществляется через матрицу поворота $A(K_k, \psi_k, \theta_k)$ [15]:

$$\begin{bmatrix} \tilde{x}_{j,k} \\ \tilde{y}_{j,k} \\ \tilde{z}_{j,k} \end{bmatrix} = A(K_k, \psi_k, \theta_k) \begin{bmatrix} x_j \\ y_j \\ z_j \end{bmatrix}, j = \overline{1, n}.$$

Для того чтобы найти оценку вектора состояния x_k^* в рамках описанной постановки задачи, требуется минимизировать следующий функционал:

$$\begin{aligned} x_k^* &= \operatorname{argmin}_{x_k} (\|y_k - h(x_k)\|_{R_k}) = \\ &= \operatorname{argmin}_{x_k} \left[\sum_{i,j} \{ \hat{c}_{0,k}(t_{j,k}^i - T^{1,i}) - d_{j,k}^i(x_k) \}^T R_k^{-1} \{ \hat{c}_{0,k}(t_{j,k}^i - T^{1,i}) - d_{j,k}^i(x_k) \} \right], \quad (1) \\ i &= \overline{1, m}, j = \overline{1, n}. \end{aligned}$$

Кроме описанной постановки задачи, рассматриваемая задача была сформулирована и в рамках байесовского подхода как задача рекуррентной фильтрации [5–7].

2. Алгоритм решения задачи относительной навигации

Стоит отметить, что задача относительной навигации при нахождении АНПА в зоне видимости 1 является многоэкстремальной, и применение традиционных численных методов на всей области определения x_k является несостоятельным. Как было сказано выше, рассмотрим двухэтапный алгоритм поиска решения. На первом этапе, для уменьшения области априорной неопределенности, найдем окрестность координат и углов ориентации АНПА, на втором этапе найдем оценку вектора состояния с учетом результатов на первом этапе. Разберем данное решение более подробно в контексте выполнения первого этапа – сужение области неопределенности, с использованием предварительно обученной нейронной сети.

Этап 1. Поиск окрестности координат и углов ориентации. Воспользуемся подходом на основе метода обучения с учителем, где обучающая выборка будет сформирована с использованием решения геометрической задачи на заданной координатной сетке и таблицы углов ориентации (в соответствии с дистанцией распространения сигнала и с учетом характеристики направленности излучателей). В качестве исходной информации для решения определим две группы:

- 1) $(t_a^i - t_b^i)_{i \in \overline{1, m}}$ – разница времен приема сигнала между парами гидрофонов (a, b) , расположенных на разных антеннах, при излучении на фиксированном излучателе;
- 2) $(t_j^a - t_j^b - T^{a,b})_{j \in \overline{1, n}}$ – разница времен приема сигнала между парами излучателей (a, b) на фиксированном гидрофоне и смещением по времени излучения сигнала излучателями (известное значение $T^{a,b}$, заданное согласно временной диаграмме излучения).

Выбор искомого метода решения определим на этапе моделирования.

Этап 2. Оценивание вектора состояния на втором этапе осуществляется за счет минимизации функционала (1) с помощью алгоритма Левенберга–Марквардта [3, 16–17]. За начальное приближение компонент вектора состояния берутся результаты с первого этапа.

Стоит отметить, что для решения задачи определения координат и углов ориентации при нахождении АНПА в 1 зоне достаточно выполнение одного – первого этапа, например можно сформировать обучающую выборку с «мелким» шагом, который обес-

печит требования по точности решения задачи навигации в зоне видимости 1. Однако выполнение второго этапа алгоритма позволяет оценить δ_k , т.е. выполнить синхронизацию шкалы времени излучения на платформе со шкалой времени на АНПА и, соответственно, решать задачу относительной навигации для зоны видимости 3, оперируя только оценками псевдодальностей излучатель – приемный гидрофон АНПА.

Алгоритм решения задачи относительной навигации для зон видимости 2 и 3 был определен ранее авторами в работах [2–6].

3. Результаты моделирования

Рассмотрим решение задачи относительной навигации на следующем практическом примере. Пусть на АНПА размещены две приемные антенны, которые располагаются в кормовой и носовой части аппарата, каждая антенна содержит по 5 приемных гидрофонов (n равно 10). Излучатели размещены по краям платформы (m равно 4).

Для выбора архитектуры сети рассмотрим 6 полносвязных нейронных сетей (англ. Fully Connected Network, FCN) различным количеством нейронов в скрытых слоях. Обучающая выборка состоит из 1368000 точек при нахождении АНПА в зоне видимости 1 и с учетом различных вариантов углов ориентации аппарата (допустимые границы крена и дифферента).

На вход нейронной сети подается 160 признаков (25×4 – из состава первой группы, и 6×10 – из состава второй группы исходной информации), выходом нейронной сети является 6 параметров (координаты и углы ориентации). В качестве функции активации используется ReLU (Rectified Linear Unit). Алгоритм оптимизации был выбран AdamW [18].

По результатам имитационного моделирования были получены оценки среднеквадратических погрешностей (СКП) координат АНПА при использовании полносвязных архитектур нейронных сетей.

*Результаты СКП оценки координат и углов ориентации
для 6 вариантов моделей сетей*

Варианты архитектур		СКП x_0 , м	СКП y_0 , м	СКП z_0 , м	СКП K , °	СКП ψ , °	СКП θ , °
FCN 1:	Без BN	1.36	0.71	1.83	3.84	0.02	0.02
	С BN	1.10	1.35	1.13	1.55	0.001	0.0002
FCN 2:	Без BN	1.33	1.54	1.48	3.94	0.03	0.03
	С BN	1.37	0.94	1.65	3.19	0.001	0.0002
FCN 3:	Без BN	1.44	1.96	1.15	1.02	0.0004	0.0002
	С BN	1.44	1.51	2.92	3.49	0.001	0.001
FCN 4:	Без BN	1.81	0.81	2.00	3.28	0.01	0.002
	С BN	0.99	0.98	1.84	2.60	0.001	0.01
FCN 5:	Без BN	1.04	1.12	1.04	2.36	0.001	0.04
	С BN	0.99	1.11	1.44	2.05	0.001	0.03
FCN 6:	Без BN	2.54	3.16	1.29	3.78	0.02	1.04
	С BN	2.33	2.89	1.87	3.51	0.001	0.81

В таблице представлены результаты для шести вариантов нейронной сети с применением пакетной нормализации (англ. Batch Normalization, BN) и без нее. Параметры архитектур FCN следующие:

- 1) 5 слоев по 160 нейронов в каждом;
- 2) 10 слоев по 160 нейронов в каждом;
- 3) 15 слоев по 160 нейронов в каждом;
- 4) 5 слоев с количеством нейронов 160-160-500-250-50;
- 5) 10 слоев с количеством нейронов 160-500-1000-3000-6000-4000-1000-500-250-160;
- 6) 5 слоев с количеством нейронов 10000-5000-1000-160.

По результатам моделирования получено, что применение BN в ряде архитектур позволяет улучшить обучение модели. В результате лучшую точность показала 1 модель с применением BN.

Однако наличие 160 входных признаков оказывается избыточным. Чтобы показать этот факт, применим математический аппарат дерева решений для ранжирования признаков по их значимости (весам). На рисунке 4 представлена гистограмма распределения признаков по их значимости. Очевидно, что полученный результат имеет хорошую физическую интерпретируемость – наиболее значимые признаки получаются на гидрофонах, имеющих наибольшую разницу по дальности от точки излучения.

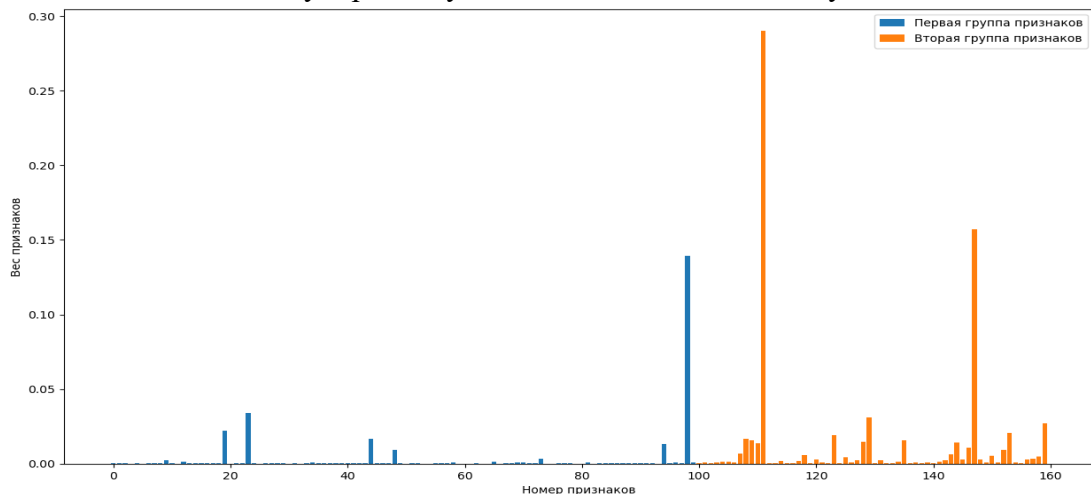


Рис. 4. Распределение весов признаков

Таким образом, отобрав 20 наиболее важных признаков, обучим лучшую модель из таблицы, изменив только входной слой.

В результате получим следующие оценки СКП: СКП x_0 равна 0.59 м; СКП y_0 равна 0.52 м; СКП z_0 равна 0.89 м; СКП K равна 2.05°; СКП ψ равна 0.008°; СКП θ равна 0.0004°. Для переноса решения под ОС реального времени (ОС QNX) воспользуемся следующим подходом:

- 1) проведем обучение нейронной сети с использованием современных библиотек на базе языка программирования python;
- 2) выгрузим полученную архитектуру и весовые коэффициенты сети в файл;
- 3) под ОС QNX реализуем восстановление архитектуры из файла с использованием конструкций базовых элементов сети на языке программирования C/C++ реализующих суммирование и функции активации на основе ReLU.

Заметим, что для выполнения первого этапа решения задачи относительной навигации на основе полносвязной нейронной сети требуется не более 45 Мб.

Моделирование решения задачи относительной навигации с использованием выбранной архитектуры нейронной сети состояло в имитации посадки АНПА с заданным креном и дифферентом на посадочную стационарную платформу. Решение задачи в зоне 3 (видимость приемных гидрофонов только одного излучателя) проводилось с переходом аппарата из зоны 1. Полученные оценки в зоне 3 при имитации вертикальной посадки АНПА с почти нулевыми значениями крена и дифферента составили: СКП x_0 равна 0.17 м; СКП y_0 равна 0.15 м; СКП z_0 равна 0.10 м; СКП K равна 1.0°; СКП ψ равна 0.006°; СКП θ равна 0.0004°.

Полученные результаты соответствуют требованиям по обеспечению безаварийной посадки АНПА на платформу.

Заключение

В работе приведена постановка задачи навигации АНПА относительно стационарной посадочной платформы. Показана возможность применения полносвязной нейронной сети для уменьшения области априорной неопределенности знания относительных координат и углов ориентации АНПА на вычислительном модуле с ОС реального времени.

Предлагаемый алгоритм позволил сократить объем используемой памяти более чем в 100 раз, по сравнению с предлагаемым ранее авторами решением и обеспечить решение задачи относительной навигации с заданной точностью.

Список источников

1. Кошаев Д. А. Относительное позиционирование и определение ориентации автономного необитаемого подводного аппарата по данным от гидроакустических маяков // Гирскопия и навигация. 2022. Т. 30, № 4 (119). С. 122–141. DOI: 10.17285/0869-7035.00107 EDN: HQJONJ
2. Грузликов А. М. Навигация АНПА в ближнем поле в интересах решения задачи приведения к причальному устройству // Сб. матер. XXIX Санкт-Петербург. междунар. конф. по интегрированным навигационным системам. 2022. С. 138–140. EDN: RKTNAM
3. Грузликов А. М., Караулов В. Г., Мухин Д. А., Шалаев Н. А. Результаты апробации алгоритма позиционирования и определения ориентации подводного аппарата по данным от гидроакустических маяков // Известия ЮФУ. Технические науки. 2023. № 1. С. 265–275. DOI: 10.18522/2311-3103-2023-1-265-274 EDN: KJGDVW
4. Караулов В. Г., Степанов О. А., Грузликов А. М., Литвиненко Ю. А. Решение задачи позиционирования АНПА относительно стационарной посадочной платформы с применением оптимизации на фактор-графе // Матер. XXXI Санкт-Петербург. междунар. конф. по интегрированным навигационным системам. 2024. EDN: IRRXPI
5. Караулов В. Г., Грузликов А. М. Использование неравенства Рао-Крамера при решении задачи относительной навигации подводного аппарата // Матер. XIV Всерос. совещания по проблемам управления. 2024. С. 1746–1752. EDN: GMBQNI

6. Караулов В. Г., Грузликов А. М., Золотаревич В. П. Рекуррентный алгоритм позиционирования и определения ориентации АНПА // Сб. матер. XXX Санкт-Петербург. междунар. конф. по интегрированным навигационным системам. 2023. С. 54–57. EDN: OLZNFX
7. Pinheiro P. M., Neto A. A., Grando R. B. et al. Trajectory Planning for Hybrid Unmanned Aerial Underwater Vehicles with Smooth Media Transition // Journal Intel Ro-bot Syst. 104, 46. 2022. DOI: 10.1007/s10846-021-01567-z EDN: FBDRDS
8. Liu S., Xu H., Lin Y., Gao L. Visual Navigation for Recovering an AUV by Another AUV in Shallow Water // Sensors. 2019, 19, 1889. DOI: 10.3390/s19081889
9. Subsea Docking Station [Электронный ресурс]. 2019. URL: <https://www.bluelogic.no/news-and-media/subsea-docking-station-sds->.
10. Peng S, Liu J, Wu J, Li C, Liu B, Cai W, Yu H. A Low-cost Electromagnetic Docking Guidance System for Micro Autonomous Underwater Vehicles // Sensors. 2019, 19(3):682. DOI: 10.3390/s19030682
11. Feezor M. D., Sorrell F. Y., Blankinship P. R., and Bellingham J. G. Autonomous Underwater Vehicle Homing/Docking via Electromagnetic Guidance // Journal of oceanic engineering. October 2001. Vol. 26, № 4. P. 515–521. DOI: 10.1109/48.972086
12. Кебкэл К. Г., Машошин А. И. Гидроакустические методы позиционирования автономных необитаемых подводных аппаратов // Гироскопия и навигация. 2016. Т. 24, №3 (94). С. 115–130. DOI: 10.17285/0869-7035.2016.24.3.115-130 EDN: WTPSNR
13. Степанов О. А. Применение теории нелинейной фильтрации в задачах обработки навигационной информации. СПб: ГНЦ РФ ЦНИИ "Электроприбор", 1998. 370 с. ISBN: 5-900780-15-5 EDN: RUVATH
14. Степанов О. А., Моторин А. В., Золотаревич В. П. [и др.] Рекуррентные и нерекуррентные алгоритмы в задачах обработки навигационной информации. Отличия и взаимосвязь с алгоритмами оптимизации на графах // XXXI Санкт-Петербург. междунар. конф. по интегрированным навигационным системам: Сб. докл. Санкт-Петербург, 27–29 мая 2024 г. СПб.: "Концерн "Центральный научно-исследовательский институт "Электроприбор", 2024. С. 336–344. EDN: IBOXVJ
15. Емельянец Г. И., Степанов А. П. Интегрированные инерциально-спутниковые системы ориентации и навигации. СПб.: ГНЦ РФ АО "Концерн "ЦНИИ Электроприбор", 2016. ISBN: 978-5-91995-029-5 EDN: XSSBEF
16. Levenberg K. A method for the solution of certain non-linear problems in least squares // Quarterly of Appl. Math. 1944. Vol. 2. P. 164–168. DOI: 10.1090/qam/10666
17. Marquardt D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters // Journal of The Society for Industrial and Applied Mathematics. 1963. Vol.11. P. 431–441.
18. Loshchilov I., Hutter F. Decoupled weight decay regularization // International Conference on Learning Representations. arXiv:1711.05101. 2019.

References

1. Koshaev, D. A. (2022), "AUV Relative Position and Attitude Determination Using Acoustic Beacons", *Gyroscopy and Navigation*, no 13, pp. 262–275. doi: 10.1134/S2075108722040083.

2. Gruzlikov, A. M. (2022), "Short and Ultra-Short Baseline Navigation of the AUV for Bringing it to the Bottom Docking Device", *Proceedings of the 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*, pp. 1–3.
3. Gruzlikov, A. M., Karaulov, V. G., Mukhin, D. A., and Shalaev, N. A. (2023), "Results of Testing the Positioning Algorithm and Determining the Orientation of the Underwater Vehicle Based on Data from Hydroacoustic Beacons", *Izvestiya SFedU. Engineering Sciences. Section 4. Communication, Navigation, and Guidance*, pp. 265–275. doi: 10.18522/2311-3103-2023-1-265-274. (In Russian).
4. Karaulov, V. G., Stepanov, O. A., Gruzlikov, A. M., and Litvinenko, Y. A. (2024), "Recursive Algorithm for AUV Positioning and Determining Its Orientation", *Proceedings of the 31st Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*, pp. 1–4.
5. Karaulov, V. G. and Gruzlikov, A. M. (2024), "Using the Cramer-Rao Inequality in Solving the Problem of Relative Navigation of an Underwater Vehicle", *Materials of the XIV All-Russian Meeting on Control Problems*, pp. 1746–1752. (In Russian).
6. Karaulov, V. G., Gruzlikov, A. M., and Zolotarevich, V. P. (2023), "Recursive Algorithm for AUV Positioning and Determining Its Orientation", *Proceedings of the 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*, pp. 1–4.
7. Pinheiro, P. M., Neto, A. A., Grando, R. B., et al. (2022), "Trajectory Planning for Hybrid Unmanned Aerial Underwater Vehicles with Smooth Media Transition", *Journal Intelligent Robotic Systems*, 104, p. 46.
8. Liu, S., Xu, H., Lin, Y., & Gao, L. (2019), "Visual Navigation for Recovering an AUV by Another AUV in Shallow Water", *Sensors*, no 19, p. 1889.
9. Blue Logic AS. (2019), *Subsea Docking Station SDS*. Available at: <https://www.bluelogic.no/news-and-media/subsea-docking-station-sds-> (Accessed: November 2023).
10. Peng, S., Liu, J., Wu, J., et al. (2019), "A Low-Cost Electromagnetic Docking Guidance System for Micro Autonomous Underwater Vehicles", *Sensors*, no 19(3), p. 682.
11. Feezor, M. D., Sorrell, F. Y., Blankinship, P. R., and Bellingham, J. G. (2001), "Autonomous Underwater Vehicle Homing/Docking via Electromagnetic Guidance", *Journal of Oceanographic Engineering*, no 26(4), pp. 515–521.
12. Kebkal, K. G., and Mashoshin, A. I. (2017), "AUV Acoustic Positioning Methods", *Gyroscopy and Navigation*, vol. 8(1), pp. 80–89. doi: 10.17285/0869-7035.2016.24.3.115-130.
13. Stepanov, O. A. (1998), *Primenenie teorii nelinejnoj fil'tracii v zadachah obrabotki navigacionnoj informacii* [Application of Nonlinear Filtering Theory in Processing Navigation Information (In Russian)], St. Petersburg: CSRI Elektropribor, 370 p.
14. Stepanov, O. A., Zolotarevich, V.P., Litvinenko, Y. A., Isaev, A. M. and Motorin, A. V. (2024), "Recursive and Nonrecursive Algorithms Applied to Navigation Data Processing: Differences and Interrelation with Factor Graph Optimization Algorithms", *Proceedings of the 31th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*, pp. 336–344.

15. Emelianstev, G. I. and Stepanov, A. P. (2016), *Integrirovannye inercial'no-sputnikovye sistemy orientacii i navigacii* [Integrated Inertial-Satellite Orientation and Navigation Systems (In Russian)], St. Petersburg: CSRI Elektropribor.
13. Levenberg, K. (1944), "A Method for the Solution of Certain Non-Linear Problems in Least Squares", *Quarterly of Applied Mathematics*, no 2, pp. 164–168.
14. Marquardt, D.W. (1963), "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", *SIAM Journal on Numerical Analysis*, no 11, pp. 431–441.
15. Loshchilov, I., and Hutter, F. (2019), "Decoupled Weight Decay Regularization", *Proceedings of the International Conference on Learning Representations*, arXiv preprint arXiv:1711.05101.

Информация об авторах:

В. Г. Караулов – инженер-программист, Государственный научный центр Российской Федерации АО Концерн «ЦНИИ "Электроприбор"» (197046, Россия, г. Санкт-Петербург, ул. Малая посадская, 30);

А. М. Грузликов – к. т. н., начальник отдела, Государственный научный центр Российской Федерации АО Концерн «ЦНИИ "Электроприбор"» (197046, Россия, г. Санкт-Петербург, ул. Малая посадская, 30).

Information about the authors:

V. G. Karaulov – Software Engineer at State Research Center of the Russian Federation – Concern CSRI Elektropribor, JSC (30, Malaya Posadskaya St., Saint Petersburg, Russia, 197046);

A. M. Gruzlikov – PhD, Head of the Department at State Research Center of the Russian Federation – Concern CSRI Elektropribor, JSC (30, Malaya Posadskaya St., Saint Petersburg, Russia, 197046).

Обзорная статья

УДК 81.322

DOI: 10.17072/1993-0550-2025-2-101-122

<https://elibrary.ru/ykakqn>

Анализ подходов к автоматизации разметки паралингвистических характеристик в русскоязычных речевых данных

Евгений Николаевич Радченко¹, Екатерина Владимировна Исаева²

¹Национальный исследовательский технологический университет "МИСИС", г. Москва, Россия
turnipseason@gmail.com

²Пермский государственный национальный исследовательский университет, г. Пермь, Россия
ekaterinaisae@psu.ru

Аннотация. Разработка систем синтеза речи с возможностью управления речевыми характеристиками посредством естественного языка имеет практический интерес, поскольку предоставляет интуитивно понятный способ влияния на результат генерации. Вместе с тем, для русскоязычных данных наблюдается недостаток как подобных систем, так и размеченных наборов данных, необходимых для их создания. Ручная разметка больших наборов данных является ресурсоемким процессом, требующим не только экспертных знаний предметной области, но и согласованности разметчиков между собой. В связи с этим, актуальным является исследование подходов к автоматизации аннотации паралингвистических характеристик русскоязычной речи, позволяющих унифицировать существующую разметку и ускорить ее масштабирование. В данной статье рассмотрены основные подходы к разметке таких паралингвистических характеристик, как паузы, ударения, а также высота и тембр голоса. Особое внимание уделено обзору доступных программных реализаций описанных методов. Ключевым выводом по итогам анализа стало наличие достаточного количества программных средств, пригодных для аннотации "базовых" характеристик в русскоязычной речи. Паузы и фундаментальная частота могут выделяться с помощью методов, не использующих лингвистическую информацию, в то время как для разметки ударений существуют методы, основанные на нейронных сетях и учитывающие контекст высказывания для снятия омографии, достигающие значения метрики Assurance в 98%. В то же время автоматическая разметка более сложных характеристик, таких как тембр и выражаемые эмоции, остается малоизученной. Данные результаты указывают на необходимость дополнительных исследований в области методов автоматической аннотации паралингвистических характеристик в русскоязычных речевых данных.

Ключевые слова: автоматическая аннотация; разметка аудио; разметка текста; паралингвистические характеристики; генерация речи



Эта работа © 2025 Радченко Е. Н., Исаева Е. В. распространяется по лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <http://creativecommons.org/licenses/by/4.0/>.

Для цитирования: Радченко Е. Н., Исаева Е. В. Анализ подходов к автоматизации разметки паралингвистических характеристик в русскоязычных речевых данных // Вестник Пермского университета. Математика. Механика. Информатика. 2025. № 2(69). С. 101–122. DOI: 10.17072/1993-0550-2025-2-101-122. <https://elibrary.ru/ykakqn>

Статья поступила в редакцию 28.04.2025; одобрена после рецензирования 15.06.2025; принята к публикации 11.07.2025.

Review article

Analysis of Approaches to Paralinguistic Feature Annotation Automation in Russian Speech

Evgenii. N. Radchenko¹, Ekaterina. V. Isaeva²

¹National University of Science and Technology "MISIS", Moscow, Russia
urnipseason@gmail.com

²Perm State University, Perm, Russia
ekaterinaisae@psu.ru

Abstract. The development of speech synthesis systems with the ability to control speech characteristics using natural language is of practical interest, since it provides an intuitive way to influence the results of the generation. At the same time, for Russian-language data there exists a shortage of both such systems and labeled datasets required to create them. Manual labeling of large datasets is a resource-intensive process that requires not only expert knowledge, but also inter-annotator labeling consistency. In this regard, the task of automating the annotation of paralinguistic characteristics of Russian-language speech becomes relevant, allowing to unify the labeling already existing in available datasets as well as accelerate its scaling to unlabeled ones.

This article considers the main approaches to the annotation of such paralinguistic characteristics as pauses, stresses, as well as the pitch and timbre of the voice. In particular, attention is paid to reviewing available software implementations of the methods described.

The key conclusion from the analysis was the existence of a sufficient number of methods suitable for annotating "basic" characteristics in Russian-language speech. Pauses and fundamental frequency can be extracted using methods that do not use linguistic information, while for stress annotation there are methods based on neural networks and, thus, taking into account the context of the utterance to resolve stress placement in homographs, achieving an Accuracy metric score as high as 98%. At the same time, automatic annotation of more complex characteristics, such as timbre and expressed emotions, remains poorly studied. These results indicate the need for additional research in the field of methods for automatic annotation of paralinguistic features in Russian-language speech corpora.

Keywords: *automatic annotation; audio annotation; text annotation; paralinguistic characteristics; speech generation*

For citation: Radchenko, E. N. and Isaeva, E. V. (2025), "Analysis of Approaches to Paralinguistic Feature Annotation Automation in Russian Speech", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 101–122. (In Russ.). DOI: 10.17072/1993-0550-2025-2-101-122. <https://elibrary.ru/> EDN

The article was submitted 28.04.2024; approved after reviewing 15.06.2024; accepted for publication 11.07.2025.

1. Введение

Естественная речь характеризуется спектром так называемых паралингвистических характеристик: длительностью пауз, ударениями, тембром и выражаемыми эмоциями говорящего, а также другими особенностями. В англоязычной литературе совокупность этих характеристик также называется *речевым стилем*. Современные алгоритмы генерации речи, основанные на глубоких нейросетях, такие как ParlerTTS и LibriTTSP [1, 2, 3], позволяют задавать желаемые характеристики синтезированного голоса, используя короткие текстовые описания.

Для использования такого способа задания стиля необходимы размеченные аудиотекстовые корпуса, включающие транскрипции и аннотацию соответствующих признаков. В то время как такие корпуса активно разрабатываются для англоязычных данных, существующие русскоязычные корпуса, такие как Dusha, ПРУД, РИНКО [4, 5, 6], обычно разработаны под конкретные узконаправленные задачи, и требуют дополнительной разметки для полноценного использования вышеобозначенных методов.

Таким образом, исследование подходов к автоматизации стилевой разметки представляется перспективным направлением, способным существенно упростить создание обучающих данных. Данная статья направлена на анализ существующих подходов к автоматической разметке паралингвистических характеристик речи и оценку их применимости к русскоязычным данным. Поскольку понятие паралингвистических характеристик охватывает широкий спектр различных качеств голоса, материал статьи поделен на подразделы, каждый из которых анализирует методы автоматизации разметки применительно к конкретным характеристикам. Для удобства чтения названия метрик оценки качества работы моделей машинного обучения в статье приводятся на английском.

2. Методы решения задачи автоматизации разметки паралингвистических характеристик

2.1. Автоматическая разметка пауз

В данном разделе будут рассмотрены методы моделирования невокализованных пауз (синтаксических и гезитационных пауз [7], если они не являлись заполненными).

Наиболее прямолинейным подходом к решению данной задачи является выделение синтаксических пауз, исходя из пунктуации. Стоит отметить, что в существующих наборах данных транскрибация зачастую получена путем автоматического распознавания речи. Так, например, в расшифровках аудио в датасете Dusha, пунктуация отсутствует. Для восстановления разметки могут быть использованы существующие программные решения, такие как "Восстановление пунктуации для русского языка" [8] и "Punctuation and casing restoration for the Russian Language (BERT-based)" [9] на основе нейронных сетей.

Наилучшее качество по метрике F1 (среднего гармонического между метриками Precision и Recall) такие модели показывают для восстановления знаков точки (0.93 и 0.7) и запятой (0.92 и 0.77). При восстановлении знаков вопроса результаты приведенных моделей по F1 различались на 0.2%, составив 0.418 и 0.42, соответственно. Примечательно, что для модели [9] это обуславливалось "средними" значениями метрик Precision (0.55) и Recall (0.34), в то время как у модели [8] – высоким значением метрики Precision (0.76) при сравнительно низком значении метрики Recall (0.29).

Наличие пунктуации может само по себе помочь расстановке пауз моделью в момент синтеза, если модель обучается синтезировать паузы, исходя из скрытых представ-

лений, не используя непосредственную аннотацию [10]. Однако некоторые исследования, проведенные для китайского языка, показывают, что разметка "просодических границ", определяющихся на уровне слогов, слов и даже отдельных предложений, также может улучшить качество синтеза. Так, например, в работе "Automatic Prosody Annotation with Pre-Trained Text-Speech Model" [11] предложен иерархический подход к разметке просодических границ для китайского языка с использованием механизма кросс-внимания между аудио- и текстовой расшифровкой. Это позволяет модели выучить соответствие между промежутками звучащей речи в аудио и тексте, чтобы затем автоматически размечать просодические границы (здесь – поделенные на классы пауз на уровне отдельных иероглифов, слов и так далее). Авторы отмечают, что поскольку разметка проводилась без участия человека, система демонстрирует повышенную консистентность в сравнении с ручной разметкой, что, в свою очередь, приводит к улучшенному качеству синтезированной речи.

Для получения прямой разметки пауз по длительности могут применяться также алгоритмы детекции речевой активности (Voice Activity Detection, VAD). Модель транскрипции WhisperX [12] поддерживает аннотацию на уровне слов и может, таким образом, быть использована без модификаций даже для корпусов, изначально не содержащих транскрипцию аудиоданных. В условиях ограниченных вычислительных ресурсов на полную транскрипцию возможно также использование гибридного подхода с такими алгоритмами, как WebRTC [17], использующим вероятностный подход на основе смесей распределений Гаусса и Лапласа, или Silero VAD [18], основанный на сверточных нейронных сетях (Convolutional Neural Network, CNN). Согласно официальной документации, на восьми наборах данных зашумленной речи Silero VAD показывает медианное значение метрики ROC-AUC в 0.95. Медианное значение ROC-AUC у WebRTC на тех же данных – 0.76.

Примерный алгоритм в случае использования методов VAD будет состоять из следующих этапов:

1. Детекция речевой активности выбранным алгоритмом VAD;
2. Сопоставление временных промежутков, выделенных в шаге 1, с текстовой расшифровкой из изначального датасета;
3. Аннотация текстовой расшифровки полученными метками длительности пауз.

Второй этап может быть реализован за счет использования эвристики, например, путем проставления временных меток по тексту пропорционально длительности аудио. При таком подходе пауза, находящаяся в середине аудио, будет аннотирована как находящаяся примерно в середине текста, если текст сегментирован по словам. Данный подход является наиболее простым в реализации, однако его качество сильно зависит от равномерности темпа речи и подходит преимущественно для случаев, когда речь не содержит значительных промежутков ускорения или замедления.

Другим возможным решением может быть повторное использование модели транскрипции, но не на всем аудио, а только на временных промежутках, выделенных моделью VAD. Для сопоставления полученной транскрипции с уже имеющейся в исходном наборе данных, можно использовать временное окно, охватывающее соответствующий сегмент аудио, а также дополнительные, например, 100 миллисекунд до и после него. Сравнение текстов можно осуществлять с использованием метрик посимвольного сходства (например, сходство Левенштейна в реализации библиотеки RapidFuzz [13]). Если значение сходства для некоторого участка исходной транскрипции превышает заданный порог, то метка паузы может быть проставлена в данном отрывке текста.

2.2. Автоматическая разметка ударений

2.2.1. Подход с использованием правил (словарей)

Ударения являются одной из базовых характеристик речи и потому также представляют интерес с точки зрения автоматизации их разметки. "Наивный" подход к автоматизации разметки ударений (акцентуации) включает в себя разметку по заранее определенным правилам. Разметка в таком случае осуществляется согласно подключаемому к алгоритму словарю, содержащему информацию об ударениях в тех или иных словоформах. Данный подход является простым в реализации и не требует больших вычислительных ресурсов.

Программную реализацию данного подхода можно найти, например, у А. Полякова [14]. В документации указано, что программа способна размечать два вида ударений: первичные (в словах типа "ёлка", "база") и второстепенные (в словах типа "авиабаза"). Если одна и та же словоформа может иметь несколько ударений, то данная программа проставляет оба. Существует также возможность подключения пользовательских словарей.

Одним из минусов данного подхода является сложность определения ударений в омографах (зАмок, замОк), а также при использовании в поэтических текстах, где авторское ударение может отличаться от общепринятого. Проблема акцентуации омографов может быть частично решена с помощью расширения алгоритма и использования, например, конечных автоматов [15]. В сочетании с данными о частотности тех или иных словоформ, такой подход достиг значения метрики Ассигасу в 96.15% на небольшом, вручную размеченном авторами статьи, корпусе, содержащем 7689 токенов. Данный корпус, однако, был собран из материалов, ориентированных на изучающих русский язык (диалоги, отрывки из классических произведений, а также вручную подобранные предложения). Расширение такой разметки на другие домены и масштабирование на использование большего количества данных представляется трудозатратным.

Рассмотрим подходы, позволяющие учитывать контекст и использующиеся таким образом для решения проблем с акцентуацией омографов.

2.2.2. Нейросетевой подход

Для учета контекста могут использоваться нейронные сети. В частности, такие архитектуры как рекуррентные нейронные сети (Recurrent Neural Network, RNN), их разновидности, такие как сети с долгой краткосрочной памятью (Long Short-Term Memory, LSTM), а также более продвинутые архитектуры типа "Трансформер", предназначенные для обработки последовательных данных. Представленные в 2017 году в статье "Attention is all you need" [16], сети трансформерной архитектуры обрабатывают входящие последовательности с помощью так называемого "механизма внимания". Благодаря ему стала возможной обработка не только отдельных слов и кратковременных контекстов, но также учет расширенного, по сравнению с RNN и LSTM, контекстного окна и ускорение обучения моделей за счет использования параллелизации вычислений.

Примером разметки ударений нейросетевым подходом является разработка И. Гусева [17]. В программном комплексе имеется возможность использовать разновидность модели трансформерной архитектуры deberta-2 или LSTM. Данные для обучения собираются из открытых источников (Викисловарь, "Грамматический словарь русского языка" А. А. Зализняка [18], а также вручную размеченный набор данных), а затем подаются на вход для обучения модели. В официальной документации сказано, что таким образом удалось достичь значения 89.73% по метрике Assurasy, без указания, однако, на какой из моделей данное значение было получено.

Для русскоязычных текстов сугубо нейросетевой подход исследовался в рамках разметки поэзии, не проверенной профессиональными редакторами и выложенной в открытом доступе на ресурсе *stihi.ru* [19]. Авторы приводят примеры некоторых удачных определений ударения, однако на момент написания статьи их код являлся недоступным для использования и модификации.

Существуют и программные средства с открытым исходным кодом, направленные на акцентуацию в области поэтических текстов. Например, программный пакет "RussianPoetryScansionTool" [20, 21] позволяет расставлять ударения в текстах, а также оценивать их стихотворный размер и рифму. Пользователю предоставляется возможность использовать на выбор одну из четырех архитектур моделей: многослойный перцептрон с ReLU-активациями, LSTM, CNN или модель трансформерной архитектуры с собственными весами. В официальной документации не приведены метрики качества работы моделей, однако есть приведенный пример использования библиотеки для аннотации стихотворения. Примечательно, что библиотека ставит ударения в словах, содержащих букву "ё", однако не всегда ставит ударение в односложных словах ("сто́ль", "что́б"). Также отмечено, что основные и второстепенные (при наличии) ударения обозначаются разными символами.

Наконец, благодаря использованию нейросетевой архитектуры двунаправленных LSTM, авторам статьи "Automated Word Stress Detection in Russian" [22, 23] удалось добиться микроусредненного (по классам слов, содержащих от двух до девяти слогов) значения Ассигасу в 0.979 на наборе данных в 1154067 уникальных тестовых примеров. Такое значение было получено при использовании модели, учитывающей контекст в формате окончания предыдущего слова. Модель, не использовавшая данную информацию, показала себя незначительно хуже – для нее микроусредненная метрика Ассигасу равнялась 0.977.

Важным фактом является то, что значение метрики Ассигасу обеих моделей было значительно ниже при тестировании на пятидесяти омографах. Для модели, учитывающей контекст, она составила 0.819, а для модели, не учитывающей контекст – 0.77. Эксперименты авторов также показали, что использование данных из размеченных корпусов является предпочтительным использованию данных из словарей, поскольку в первом случае слова находятся в контексте и позволяют модели учитывать частотность возможных ударений.

Стоит отметить, что, поскольку нейросетевые подходы основываются на частотных закономерностях языка, они также подвержены изменениям в обучающей выборке и могут плохо показывать себя на примерах, слабо репрезентированных в обучающих данных, таких как авторские изменения ударения в поэтическом домене, если модель была обучена сугубо на прозаических текстах.

2.2.3. Комбинированный подход

Комбинированный подход к разметке ударений в поэтических текстах в русском языке представлен в статье "Комбинированный Словарно-Нейросетевой Акцентуатор Для Разметки Русского Поэтического Текста" [24]. Как было отмечено ранее, разметка таких текстов представляет особую сложность, поскольку наряду с омографами может содержать также и авторские ударения, обусловленные ритмикой конкретного стихотворения.

Взяв за основу разработку [22], не используя сторонние библиотеки для POS-теггинга (такие, как *ru morphology* или *SpaCy*), авторы учитывали морфологический контекст с помощью использования флексии предшествующего слова. Обучающая выборка собиралась из "Грамматического словаря русского языка" А. А. Зализняка и устного подкорпуса Национального корпуса русского языка (НКРЯ) [25]. Авторы отмечают, что

ошибки, совершаемые словарным и нейросетевым акцентуаторами, отличаются по своей сути. Словарный акцентуатор ошибается в определении ударений в словах, отсутствующих в словаре, а также в случаях с неоднозначным ударением. Двухнаправленная LSTM-модель ошибалась, например, при определении ударения в словах с подразумеваемой, но не обозначенной буквой "ё", что может указывать на необходимость предварительной "ёфикации" (в наиболее простой реализации: за счет словарей) текстов перед использованием такого подхода.

Авторы рассмотрели несколько способов совмещения подходов – изначальная разметка словарным акцентуатором с последующей разметкой нейросетевым акцентуатором, изначальная разметка нейросетевым акцентуатором с последующей разметкой словарным акцентуатором, а также случайный выбор разметки тем или иным акцентуатором для каждого слова.

В итоговом пайплайне реализована одновременная разметка обоими акцентуаторами, причём результаты разметки нейросетевого акцентуатора учитываются только в случае неоднозначной разметки, отсутствия разметки словарным акцентуатором (при условии отсутствия в слове буквы "ё"), либо наличия и разметки, и буквы "ё" (кроме слов с дефисом). Результаты разработки реализованы в виде библиотеки `ru-accent-poet` [26] на языке Python, доступной для скачивания.

За счет использования словаря при разметке ударений для однозначных слов и использовании нейросетей для разметки ударений в омографах, удалось достичь результатов, превосходящих использование только одного из методов. Качество работы сравнивалось на вручную размеченных авторами стихах, а также на 100 строках из поэтического подкорпуса НКРЯ. На данной выборке комбинированный подход показал качество около 0.98 по метрике Ассигасу, по сравнению с 0.93 у словарного и 0.94 у отдельно нейросетевого подходов, соответственно.

2.3. Автоматическая разметка высоты и тембра

Рассмотрим такие характеристики как высота и тембр голоса. Под высотой будем понимать слуховое ощущение частоты звука, а под тембром – признак, позволяющий слушателю различать звуки одинаковой высоты и громкости, но различного генезиса [27]. Высота звука тесно связана с его "фундаментальной частотой" (частота основного тона, F_0), в то время как тембр – с обертонами, то есть всей частью звукового спектра, не относящейся к фундаментальной частоте [28].

Исследования показывают, что высота голоса играет ключевую роль в восприятии социальных качеств, таких как надежность, авторитетность, лидерские качества говорящего [29, 30, 31, 32]. Таким образом, моделирование и автоматическая разметка высоты и тембра являются особенно актуальными при, например, создании голосовых ассистентов, где формирование доверительного отношения пользователей является одним из наиболее важных факторов в разработке.

2.3.1. Выделение фундаментальной частоты и тональных контуров

Алгоритмы выделения фундаментальной частоты, как правило, опираются на анализ аудиосигнала во временной либо частотной области, также существуют и гибридные подходы [33]. Поскольку данная статья рассматривает разметку в целях генерации аудио, содержащего характеристики только одного говорящего, то рассматриваться будут только базовые алгоритмы, применяющиеся при определении F_0 в вышеуказанном сценарии.

При работе с речью в частотной области часто применяются мел-кепстральные коэффициенты (Mel-frequency Cepstrum Coefficient, MFCC) – представление, полученное

через обратное преобразование Фурье от логарифма спектра мощности сигнала. Использование мел-шкалы позволяет учесть нелинейную связь между воспринимаемой и физической частотой, возникающую в результате особенностей человеческой физиологии.

Большинство алгоритмов, использующих преимущественно временное представление входного сигнала для определения F_0 , основаны на принципе автокорреляции. Входной сигнал разделяется на части (фреймы), от которых высчитывается автокорреляционная функция, отображающая сходство сигнала с самим собой. В наиболее простом подходе первый максимум данной функции и будет фундаментальной частотой. На базе этого подхода были разработаны несколько алгоритмов, таких как AMDF [34], YIN [35], а также его вероятностная модификация pYIN [36] и другие.

В алгоритме YIN используется кумулятивное нормализованное среднее, что позволяет ему быть более устойчивым к колебаниям в амплитуде входного сигнала, делая период F_0 более выраженным по отношению к остальным. Программная реализация YIN и pYIN доступна в библиотеке librosa [37].

Подходы, использующие исключительно анализ в частотной области, пользуются меньшей популярностью и их реализации не настолько распространены, как реализации гибридных подходов или подходов, основанных на анализе амплитудно-временных характеристик. Среди известных гибридных подходов можно отметить HARVEST [37] и YAAPT [38].

HARVEST извлекает основные частоты-кандидаты F_0 с помощью набора фильтров с разными частотами, анализируя спектральные компоненты, после чего уточняет их с использованием "мгновенной" частоты. Затем несколько кандидатов F_0 оцениваются в каждом фрейме. Для формирования финального F_0 -контура применяется алгоритм соединения соседних кадров, опирающийся на предположение о плавности изменения высоты тона, что делает его более устойчивым к локальным шумам (проблема, проявляющаяся при покадровой обработке сигналов). Реализация алгоритма доступна в библиотеке ruworld [39], также предоставляющей реализацию алгоритма DIO [40].

Ядром алгоритма YAAPT является метод нормализованной кросс-корреляции, использующийся вместо обычной автокорреляционной функции. На этапе предобработки над входным сигналом производится нелинейное преобразование, позволяющее восстановить слабые компоненты F_0 . Для выбора наиболее правдоподобной F_0 используется динамическое программирование, что делает алгоритм устойчивым к искажениям и эффективным при работе как с записями высокого качества, так и с, например, телефонной речью. Реализация алгоритма YAAPT доступна для Python в библиотеке AMFM_decompy [41].

Существуют также статьи, описывающие успешное применение методов традиционного машинного обучения (алгоритмов К-Среднего, модели смесей Гауссовских распределений, метода опорных векторов [42]) и CNN [43] для решения задачи выделения F_0 .

Помимо отдельной частоты F_0 можно выделять также тональный контур, представляющий собой изменение тона на протяжении отдельного отрезка звучащей речи. Одним из наиболее распространенных программных решений для работы с анализом речи, дающим возможность выделения тональных контуров, является программный пакет Praat [44], доступный для использования с языком Python с помощью библиотеки Parselmouth [45]. Алгоритм, применяющийся в Parselmouth для выделения F_0 по умолчанию, является автокорреляционным [46], однако пользователю предоставлена возможность выбрать и другие методы.

Для разметки текстовых данных, поступающих в модель на момент синтеза, можно использовать моделирование тональных контуров на основе частотных характеристик частей речи, как это было сделано для тамильского языка в статье "Utilizing POS-Driven

Pitch Contour Analysis for Enhanced Tamil Text-to-Speech Synthesis" [47], однако возможность успешного использования такого подхода применительно к русскому языку требует дополнительных исследований.

Рассмотренные выше методы в основном были направлены на разметку базовых характеристик звучащей речи – ударения, пауз и тона. Хотя их модуляция и способствует повышению естественности синтезированной речи, этого недостаточно для моделирования речевого многообразия. В следующем разделе будут рассмотрены способы автоматической разметки более сложных аспектов стиля, таких как акцент и эмоция говорящих.

2.3.2. Разметка характеристик тембра в текстовом формате

Наиболее простым с точки зрения естественного языка способом задать желаемый голос является непосредственное задание характеристик словесным образом. Ряд исследований сфокусировался на создании таких словесных описаний. Например, авторами статьи "Dream Voice: Text Guided Voice Conversion" [48] была разработана система из десяти ключевых слов, разделенных на две категории в зависимости от уровня субъективности. Первая категория составляла базовые характеристики, такие как пол и возраст говорящего, вторая же соответствовала более абстрактным понятиям, таким как сила или теплота голоса. Разметка проводилась вручную экспертами, а в итоговом датасете было 900 говорящих.

Авторы статьи [1] развили идею стилистической разметки, предложив подход к ее автоматизации. Дополнительно к этому ими была предложена обширная система тегов, охватывающих как присущие отдельным говорящим характеристики ("Intrinsic tags"), такие как акцент и пол, так и ситуативные характеристики, такие как выражаемая в речи эмоция ("Situational tags"). Стоит отметить, что для разметки ситуативных характеристик не подходит использование исключительно методов, основанных на анализе тональности текста, так как интонация высказывания может отличаться от семантики предложения.

Авторы также выделяли уровень "сложности" тегов, где "базовыми" ("Basic") считались такие теги, как пол говорящего, скорость речи и высота голоса, поддающиеся определению с помощью методов обработки сигналов, а "расширенными" ("Rich") считались такие теги, как выражаемая эмоция, акцент и другие, обычно требующие человеческой разметки.

Сфокусировавшись на масштабировании расширенных характеристик (R-тегов), авторы проанализировали существующие, преимущественно англоязычные, наборы данных (датасеты). Из находящихся в открытом доступе датасетов, 1 из 10 имел разметку исключительно присущих R-тегов [49], в 6 из 10 присутствовала разметка только ситуативных R-тегов [2, 50, 51, 52, 53, 54] и в 1 из 10 [55] – разметка как присущих, так и ситуативных R-тегов. При этом датасеты, разметка которых производилась автоматически, не имели разметки присущих R-тегов, а размеченные ситуативные R-теги ограничивались 4 [2] и 7 [54] тегами, соответственно. Таким образом, была выявлена необходимость автоматизации разметки R-тегов.

Дальнейшая разметка производилась двумя способами, отдельно для присущих R-тегов (IR-тегов) и ситуативных R-тегов (SR-тегов). Пайплайн для IR-тегов начинался с ручной разметки небольшого "стартового" датасета. Затем датасет масштабировался путем нахождения голосов, похожих на известные, и переноса на них имеющихся IR-тегов. Для каждого размеченного вручную голоса и каждого голоса из размечаемого датасета, авторы вычисляли медианные эмбединги на основе десяти случайно выбранных аудиофрагментов, используя модель VoxSim [56]. Особенностью данной модели является то, что она обучалась определять не то, насколько разные фрагменты голосов при-

надлежат одному и тому же голосу, но то, насколько те или иные фрагменты воспринимаются похожими с точки зрения человека ("perceptual speaker similarity"). Авторы статьи отметили, что если два голоса имеют высокое сходство по восприятию, то у них обычно совпадает большинство IR-тегов.

Для каждого говорящего из размеченного датасета находились говорящие из размечаемого датасета, косинусное сходство с которыми было больше или равно 0.8, а затем размечаемому говорящему копировались все IR-теги.

Вторая часть пайплайна, использовавшаяся для разметки SR-тегов, состояла из трех этапов. На первом из них авторы фильтровали наиболее эмоционально окрашенные высказывания, пользуясь готовым классификатором для трехфакторного моделирования эмоциональных состояний по шкалам степени контроля (dominance), интенсивности (arousal) и приятности (valence) [57]. Отметим, что в официальной документации модели описано ее применение для классификации эмоций в аудио-текстовом наборе данных "The Berlin Database of Emotional Speech" [58] на немецком языке, содержащем разметку аудио на пять эмоций, плюс одну "нейтральную" эмоцию. Обученный на основе эмбедингов данной модели классификатор, использующий метод опорных векторов (Support Vector Classifier, SVC), показал значение метрики Unweighted Average Recall (UAR) в 0.93, что говорит в пользу ее применимости для языков, отличных от английского.

На втором этапе текстовая расшифровка оценивалась на предмет соответствия семантики размечаемой характеристике. С помощью модели SFR-Embedding-Mistral [59], авторы вычисляли косинусное сходство между промптом: "Instruct: Given an emotion, retrieve relevant transcript lines whose overall style/emotions matches the provided emotion. Query: {emotion}" ["Инструкция: По заданной эмоции верни релевантные строки расшифровки, стиль/эмоция которых соответствует предложенной. Запрос: {Эмоция}"] и текстовыми расшифровками речи, полученными в результате первого этапа фильтрации. Чтобы избежать переоценки реплик, в которых просто упоминается эмоция (например, реплики, содержащие слово "ярость", но не имеющие соответствующего эмоционального окраса), реплики также фильтровались по ключевым словам. После получения косинусного сходства реплики ранжировались от наиболее до наименее подходящих под промпт.

Заключительным этапом было акустическое сопоставление. Для фильтрации ложноположительных срабатываний после второго этапа, авторы брали топ-100 тысяч реплик, наиболее подходивших под промпт той или иной эмоции. Выбранные реплики подавались на вход аудио-модели Gemini 1.5 Flash. Модель оценивала, насколько интонация соответствует заданной эмоции по шкале от 1 до 5, при этом в промпте содержалось указание не учитывать семантику высказывания. В результате оставлялись только реплики, получившие оценку 5. Результатом данной части пайплайна стал набор реплик, точно отражающих нужную эмоцию и по содержанию, и по звучанию. Проведенное исследование с удалением части компонент ("ablation study") пайплайна показало, что каждый из них необходим для получения более качественного результата.

В целях оценки возможности применимости такой модели фильтрации к русскоязычным аудио, авторами данной статьи также был проведен мини-эксперимент по ее использованию. Для сравнения было выбрано высказывание "Я так люблю эту жизнь, я самый счастливый человек на планете Земля", произнесенное с грустной эмоцией, прямо противоположной его ярко-выраженной положительной семантике. Нескольким моделям серии Gemini было предложено поставить оценку от 1 до 5, сравнив, насколько выражаемая в аудио эмоция соответствует эмоции "грусть". Используемый промпт соответствовал промпту из статьи [1], однако название эмоции указывалось на русском: "Analyze the provided speech clip to evaluate how effectively it conveys the emotion {emotion

→ Грусть}, focusing on tone of voice and delivery, rather than the spoken content... " ["Про-анализируй данный отрывок речи для того, чтобы оценить, насколько эффективно он передает эмоцию {эмоция → Грусть}, фокусируясь на тоне голоса и подаче, нежели на содержании высказывания...}"].

Модель "gemini-1.5-flash-002", применявшаяся в изначальной статье, не справилась с поставленной задачей, поставив оценку 1/5. Однако более новая модель "gemini-2.5-flash-preview-04-17" в режиме "Thinking mode" справилась с задачей, поставив оценку 5/5, что указывает на потенциал использования ее API для реализации аналогичного пайплайна для русскоязычных данных.

3. Результаты

В статье были проанализированы основные подходы к автоматизации разметки паралингвистических характеристик речи, таких как паузы, ударения, высота, а также "R-теги", включающие эмоции, акцент и другие особенности говорящих.

Наиболее исследованным из направлений автоматизации разметки является разметка ударений, в частности для домена поэтических текстов. Устоявшиеся алгоритмы к моделированию просодических границ и тональных контуров могут быть использованы без дополнительных модификаций. Существуют также нестандартные подходы, успешно применявшиеся для китайского и тамильского языков, однако требующие апробации для доказательства эффективности на русскоязычных данных. В силу различной природы происхождения текстовых данных, для повышения качества и обеспечения стабильности работы алгоритмов, рекомендуется использовать предобработку, включающую в себя "ёфикацию", капитализацию и восстановление знаков препинания.

Сводная информация о доступных методах автоматической разметки базовых характеристик на русскоязычных данных представлена в таблице.

Сводная таблица доступных программных решений для разметки базовых паралингвистических характеристик на русском языке

Размечаемая характеристика	Программное решение
Паузы	Для восстановления знаков препинания: ru_punct [8] ru-autopunctuation [9]
	Для транскрибации и временных меток: WhisperX [12]
	Для гибридного использования с моделями транскрибации: WebRTC [60] Silero VAD [61]
Ударения	Accenter [14] russ [17] RussianPoetryScansionTool [20] russtress [23] ru-accent-poet [26]
F0, Тональный контур	librosa [62] pYAAAPT [41] Parselmouth [45]

В целом можно сказать, что задача автоматизации разметки базовых паралингвистических характеристик в русском языке может быть успешно решена с использованием комбинации существующих программных решений. Выбор библиотек для реализации будет зависеть от существующих вычислительных мощностей, объема обрабатываемых данных, а также требуемой для конкретной задачи разметки точности.

Так, для быстрого прототипирования подойдут методы, основанные на правилах и эвристиках. В частности – разметка пауз на основе синтаксиса или словарный подход в случае разметки ударений. Для задач, требующих повышенной точности, подойдет использование нейросетевых и гибридных методов, таких как WhisperX для разметки пауз и *gu-accent-roet* для разметки ударений. Стоит отметить, что все описанные методы разметки ударений требуют наличия текстовой расшифровки аудио. В условиях отсутствия такой расшифровки, с учетом особенностей ударения в русском языке, разметка может быть произведена за счет выделения участков локального изменения длительности и тембральных характеристик гласных [63] непосредственно из аудио. На момент написания статьи авторам не известно о существовании готовых программных реализаций подобных алгоритмов, что предполагает дополнительные временные затраты на их разработку.

В случае разметки расширенных характеристик наиболее популярными являются подходы, основанные на нейросетевых моделях, опирающиеся, как правило, на решения, обученные преимущественно на англоязычных данных. Например, модель VoxSim, описанная в статье [36], была обучена исключительно на англоязычном материале и может оказаться непригодной для сравнения голосов в русскоязычной речи. С другой стороны, для моделей, не обучавшихся в явном виде использовать семантику высказывания при определении эмоций, существует потенциал применимости для языков, не входивших в обучающую выборку.

В совокупности результаты анализа показывают необходимость разработки программных решений, позволяющих автоматизировать разметку расширенных характеристик для русскоязычных данных.

4. Обсуждение

В перспективе разработки методов автоматизации разметки расширенных тегов для русскоязычных данных, актуальной задачей остается и разработка собственной системы тегов. Существующие аудио-текстовые наборы данных на русском языке, такие как Dusha, *resd_annotated*, CommonVoice 21.0, ПРuD, РИНКО [4, 5, 6, 64, 65] и другие, не имеют единой стандартизированной системы аннотации. Они также различаются по качеству записей и квалификации разметчиков (от собранных и размеченных пользователями сети Интернет до записанных и размеченных профессионально). Некоторые из них содержат отдельные базовые или расширенные характеристики: например, в CommonVoice аннотирован пол говорящих, в ПРuD – их диалекты. При этом не все из описанных корпусов находятся в открытом доступе. Для создания системы, охватывающей как можно более гибкий спектр характеристик, необходимо провести дополнительное исследование, которое позволит определить, какие данные доступны и пригодны для использования.

Исходя из существующих решений для разметки расширенных тегов можно отметить, что некоторые из них показывают возможность масштабирования на языки, не входившие в обучающую выборку. Однако для однозначных выводов требуется проведение тестов на наборах вариативных русскоязычных данных. При интерпретации результатов стоит учитывать такие факторы, как культурные отличия в выражении тех или иных эмоций, спонтанность размечаемой речи, а также, для анализа качества работы моделей на подготовленной речи – наличие навыков сценического мастерства у говорящих.

Дополнительным направлением для улучшения системы генерации естественной речи может стать разметка и синтез экстралингвистических компонент речи – таких как смех, кашель, цоканье и другие [66]. Кроме того, перспективным является задание стиля не только говорящего, но и окружающей среды, в том числе с помощью изображений [67, 68], а также расширение на ситуации с несколькими говорящими или переключением кодов.

Данная работа ограничена по охвату и не претендует на исчерпывающее рассмотрение всех существующих подходов и методов. Авторы стремились сделать акцент на практическом применении программных решений, что, как они надеются, может способствовать прикладному развитию в области создания и масштабирования аудио-текстовых наборов русскоязычных данных.

5. Заключение

В данной статье приведен обзор некоторых из существующих методов автоматической разметки паралингвистических характеристик, а также оценен потенциал их использования для русскоязычных аудио-текстовых наборов данных. Анализ выявил, что существующих программных решений достаточно для выделения базовых характеристик, таких как паузы, ударения, фундаментальная частота и тональные контуры. Описанные методы могут использоваться для унификации и масштабирования разметки наборов данных, находящихся в открытом доступе, в целях обучения моделей синтеза речи.

Вместе с тем обнаружена нехватка методов, позволяющих автоматизировать разметку расширенных характеристик для русскоязычных данных. По этой причине возникает необходимость в адаптации и проверке методов, продемонстрировавших эффективность на англоязычных данных, с использованием русскоязычных данных и моделей.

Для дальнейшего развития в этом направлении предлагается:

1. Разработать собственную систему тегов, учитывающую как базовые, так и расширенные характеристики;
2. Провести комплексный анализ существующих русскоязычных аудио-текстовых наборов данных с точки зрения доступности и характера аннотации;
3. Выполнить сравнительный анализ методов оценки голосовой схожести, исследовать потенциал использования моделей, обучавшихся на англоязычных данных, для сравнения русскоязычной речи.

Список источников

1. Diwan A., Zheng Z., Harwath D., Choi E. Rich Style-Prompted Text-to-Speech Datasets, 2025. URL: <http://arxiv.org/abs/2503.04713> (дата обращения: 31.03.2025).
2. Guo Z., Leng Y., Wu Y., et al. PromptTTS: Controllable Text-to-Speech with Text Descriptions // Proceedings of ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023. P. 1–5. DOI: 10.1109/ICASSP49357.2023.10096285.
3. Lacombe Y., Srivastav V., Gandhi S. Inference and training library for high-quality TTS models. URL: <https://github.com/huggingface/parler-tts> (дата обращения: 15.04.2025).
4. Kondratenko V., Sokolov A., Karpov N. et al. Large Raw Emotional Dataset with Aggregation Mechanism, 2022. URL: <http://arxiv.org/abs/2212.12266> (дата обращения: 31.03.2025).
5. Князев С. В., Мороз Г. А., Дьяченко С. В. Корпус Просодии Русских Дialectов (ПРyД). URL: <https://lingconlab.github.io/PRuD/> (дата обращения: 10.04.2025).

6. Кривнова О. Ф., Архипов А. В., Захаров Л. М., Кобозева И. М. Интонация устного дискурса: русский интонационный корпус РИНКО (RINCO) // Речевые Технологии. 2020. № 1–2. С. 113–120.
7. Речевые хезитации: формальный и функциональный аспекты / Яковлева Э. Б.: Институт научной информации по общественным наукам РАН, 2016. 74 с. URL: <https://elibrary.ru/item.asp?id=30706219> (дата обращения: 21.04.2025).
8. Гутник Г. Нейронная сеть для восстановления пунктуации на русском языке. URL: https://github.com/gleb-skobinsky/ru_punct (дата обращения: 20.04.2025).
9. Котик К. Punctuation and casing restoration for the Russian Language (BERT-based). URL: <https://github.com/kotikkonstantin/ru-autopunctuation> (дата обращения: 20.04.2025).
10. Hwang J.-S., Lee S.-H., Lee S.-W. PauseSpeech: Natural Speech Synthesis via Pre-trained Language Model and Pause-Based Prosody Modeling // Pattern Recognition / eds. H. Lu et al. Springer Nature, 2023. P. 415–427.
11. Dai Z., Yu J., Wang Y. et al. Automatic Prosody Annotation with Pre-Trained Text-Speech Model // Proceedings of Interspeech 2022. P. 5513–5517. DOI: 10.21437/Interspeech.2022-10005.
12. Bain M., Huh J., Han T., Zisserman A. WhisperX: Time-Accurate Speech Transcription of Long-Form Audio // Proceedings of Interspeech. 2023. P. 4489–4493. DOI: 10.21437/Interspeech.2023-78.
13. RapidFuzz 3.13.0 documentation. URL: <https://rapidfuzz.github.io/RapidFuzz/index.html> (дата обращения: 20.04.2025).
14. Поляков А. Accenter/Accenter.txt at master. Accenter created by Alexey Polyakov, GitHub. URL: <https://github.com/sStress/Accenter/blob/master/Accenter.txt> (дата обращения: 11.04.2025).
15. Reynolds R., Tyers F. Automatic word stress annotation of Russian unrestricted text // Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015). 2015. P. 173–180. URL: <https://aclanthology.org/W15-1822/> (дата обращения: 11.04.2025).
16. Vaswani A. et al. Attention is All you Need // Advances in Neural Information Processing Systems. Curran Associates, Inc., 2017. Vol. 30. URL: https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (дата обращения: 11.04.2025).
17. Гусев И. Package for word stress detection URL: <https://github.com/IlyaGusev/russ> (дата обращения: 11.04.2025).
18. Грамматический словарь русского языка / Зализняк А. А. М., 1977. URL: <https://gramdict.ru/> (дата обращения: 11.04.2025).
19. Гришина Е. А., Зеленков Ю. Г., Орехов Б. В. Наивная Поэзия В Акцентологическом Корпусе // Труды Института русского языка им. В. В. Виноградова. 2015. Т. 3, № 6. С. 257–272.
20. Koziev I. Detection of poetic meter, rhyme, and stress placement in the texts of Russian accentual-syllabic poems and songs URL: <https://github.com/RussianNLP/RussianPoetryScansionTool> (дата обращения: 13.04.2025).
21. Koziev I. Automated Evaluation of Meter and Rhyme in Russian Generative and Human-Authored Poetry, 2025. URL: <http://arxiv.org/abs/2502.20931> (дата обращения: 16.04.2025).
22. Ponomareva M., Milintsevich K., Chernyak E., Starostin A. Automated Word Stress Detection in Russian // Proceedings of the First Workshop on Subword and Character Level Models in NLP SCLeM. 2017. P. 31–35. URL: <https://aclanthology.org/W17-4104/> (дата обращения: 11.04.2025).

23. *Ponomareva M.* Python package russtress accentuates russian text. URL: <https://github.com/MashaPo/russtress> (дата обращения: 13.04.2025).
24. *Короткова Ю. О.* Комбинированный Словарно-Нейросетевой Акцентуатор Для Разметки Русского Поэтического Текста // Труды Института русского языка им. В. В. Виноградова. 2022. № 3. С. 181–190.
25. *Савчук С. О., Архангельский Т. А., Бонч-Осмоловская А. А. и др.* Национальный корпус русского языка 2.0: новые возможности и перспективы развития // Вопросы Языкознания. 2024. № 2. С. 7–34.
26. *Короткова Ю. О.*, ru-accent-poet. URL: https://github.com/yuliya1324/ru_accent (дата обращения: 11.04.2025).
27. *Педагогическое речеведение: словарь-справочник* под ред. Т. А. Ладыженской и А. К. Михальской / Князьков А. А. под ред. Т. А. Ладыженской и А. К. Михальской. 1998. URL: <http://rus-yaz.niv.ru/doc/pedagogical-speech/index.htm> (дата обращения: 11.04.2025).
28. *Pitch Extraction and Fundamental Frequency: History and Current Techniques.* Pitch Extraction and Fundamental Frequency / Gerhard D. Department of Computer Science, University of Regina, 2003. 44 p.
29. *Klofstad C. A. et al.* Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women // *Proceedings in Biological Sciences*. 2012. Vol. 279, № 1738. P. 2698–2704.
30. *O'Connor J. J. M. et al.* The influence of voice pitch on perceptions of trustworthiness across social contexts // *Evolution and Human Behavior*. 2017. Vol. 38, № 4. P. 506–512.
31. *Wang T.-Y., Kawaguchi I., Kuzuoka H., Otsuki M.* Effect of Manipulated Amplitude and Frequency of Human Voice on Dominance and Persuasiveness in Audio Conferences // *Proc. ACM Human-Computer Interaction*. 2018. Vol. 2. № CSCW. P. 177:1–177:18.
32. *Wu H. X., Li Y., Ching B. H.-H., Chen. T. T.* You are how you speak: The roles of vocal pitch and semantic cues in shaping social perceptions // *Perception*. 2023. Vol. 52, № 1. P. 40–55.
33. *Имамвердиев Я. Н., Сухостат Л. В.*, Подходы для оценки периода основного тона речевого сигнала в зашумлённой среде // *Речевые Технологии*. 2014. № 1–2. С. 84–103.
34. *Ross M., Shaffer H., Cohen A. et al.* Average magnitude difference function pitch extractor // *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1974. Vol. 22, № 5. P. 353–362.
35. *De Cheveigné A., Kawahara H.* YIN, a fundamental frequency estimator for speech and music // *The Journal of the Acoustical Society of America*. 2002. Vol. 111, № 4. P. 1917–1930.
36. *Mauch M., Dixon S.* PYIN: A fundamental frequency estimator using probabilistic threshold distributions // *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2014. P. 659–663. URL: <https://ieeexplore.ieee.org/document/6853678> (дата обращения: 15.04.2025).
37. *Morise M.* Harvest: A High-Performance Fundamental Frequency Estimator from Speech Signals // *Proceedings of Interspeech 2017*. P. 2321–2325. DOI:10.21437/Interspeech.2017-68.
38. *Kasi K., Zahorian S. A.* Yet Another Algorithm for Pitch Tracking // *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2002. Vol. 1. P. I-361-I-364. URL: <https://ieeexplore.ieee.org/document/5743729> (дата обращения: 15.04.2025).

39. *Hsu J. et al.* PyWorld: a Python wrapper for WORLD vocoder. URL: <https://github.com/JeremyCCHsu/Python-Wrapper-for-World-Vocoder> (дата обращения: 20.04.2025).
40. *Morise M., Kawahara H., Katayose H.* Fast and Reliable F0 Estimation Method Based on the Period Extraction of Vocal Fold Vibration of Singing Voice and Speech // CD-ROM Proceeding AES 35th International Conference: Audio for Games. London, United Kingdom, 2009.
41. *Schmitt B. J. B.* pYAAPT. AMFM_decompy 1.0.11 documentation. URL: https://bjbschmitt.github.io/AMFM_decompy/pYAAPT.html (дата обращения: 15.04.2025).
42. *Drugman T., Huybrechts G., Klimkov V., Moinet A.* Traditional Machine Learning for Pitch Detection // IEEE Signal Processing Letters. 2018. Vol. 25, № 11. P. 1745–1749.
43. *Kim J. W., Salamon J., Li P., Bello J. P.* CREPE: A Convolutional Representation for Pitch Estimation. CREPE // Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018. P. 161–165. DOI: 10.1109/ICASSP.2018.8461329.
44. *Boersma P., van Heuven P.* Praat, a system for doing phonetics by computer // Glot International. 2001. Vol. 5, № 9/10. P. 341–345.
45. *Jadoul Y., Thompson B., de Boer B.* Introducing Parselmouth: A Python interface to Praat // Journal of Phonetics. 2018. Vol. 71. P. 1–15.
46. *Boersma P.* Accurate Short-Term Analysis of The Fundamental Frequency and The Harmonics-To-Noise Ratio of a Sampled Sound // Proceedings of the institute of phonetic sciences. 1993. Vol. 17, № 1193. P. 97–110.
47. *Thinakaran P., Gladston A. R., Vijayalakshmi P. et al.* Utilizing POS-Driven Pitch Contour Analysis for Enhanced Tamil Text-to-Speech Synthesis // Proceedings of the 21st International Conference on Natural Language Processing (ICON). 2024. P. 269–273.
48. *Hai J., Thakkar K., Wang H. et al.* DreamVoice: Text-Guided Voice Conversion // Proceedings of Interspeech 2024. P. 4373–4377. DOI: 10.21437/Interspeech.2024-1432.
49. *Kawamura M., Yamamoto R., Shirahata Y. et al.* LibriTTS-P: A Corpus with Speaking Style and Speaker Identity Prompts for Text-to-Speech and Style Captioning. // Proceedings of Interspeech 2024. P. 1850–1854. DOI: 10.21437/Interspeech.2024-692.
50. *Nguyen T. A., Hsu W.-N., D’Avirro A. et al.* EXPRESSO: A Benchmark and Analysis of Discrete Expressive Speech Resynthesis // Proceedings of Interspeech 2023. P. 4823–4827. DOI: 10.21437/Interspeech.2023-1905.
51. *Richter J., Wu Y.-C., Krenn S., et al.* EARS: An Anechoic Fullband Speech Dataset Benchmarked for Speech Enhancement and Dereverberation // Proceedings of Interspeech 2024. P. 4873–4877. DOI: 10.21437/Interspeech.2024-153.
52. *Ji S., Zuo J., Fang M. et al.* TextrolSpeech: A Text Style Control Speech Corpus with Codec Language Text-to-Speech Models // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2024. P. 10301–10305. URL: <https://ieeexplore.ieee.org/document/10445879> (дата обращения: 31.03.2025).
53. *Guan W., Li Y., Li T. et al.* MM-TTS: Multi-modal Prompt based Style Transfer for Expressive Text-to-Speech Synthesis // Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence. P. 18117–18125. DOI: 10.1609/aaai.v38i16.29769.

54. Jin Z. *et al.* SpeechCraft: A Fine-Grained Expressive Speech Dataset with Natural Language Description // Proceedings of the 32nd ACM International Conference on Multimedia. 2024. P. 1255–1264. DOI: 10.1145/3664647.3681674.
55. Watanabe A., Takamichi S., Saito Y., *et al.* Coco-Nut: Corpus of Japanese Utterance and Voice Characteristics Description for Prompt-based Control // 2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU). P. 1–8. DOI: 10.1109/ASRU57964.2023.10389693.
56. Nagrani A., Chung J. S., Xie W., Zisserman A. Voxceleb: Large-scale speaker verification in the wild // Computer Speech & Language. 2020. Vol. 60. P. 101027. DOI: 10.1016/j.csl.2019.101027.
57. Wagner J., Triantafyllopoulos A., Wierstorf H. *et al.* Dawn of the transformer era in speech emotion recognition: closing the valence gap // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2023. Vol. 45, № 9. P. 10745–10759. DOI: 10.1109/TPAMI.2023.3263585.
58. Sendlmeier W., Burkhardt F., Kienast M., Paeschke A., Weiss B. The Berlin Database of Emotional Speech. URL: <http://emodb.bilderbar.info/docu/> (дата обращения: 09.06.2025).
59. Meng R. *et al.* SFR-Embedding-Mistral: Enhance Text Retrieval with Transfer Learning. URL: <https://huggingface.co/Salesforce/SFR-Embedding-Mistral> (дата обращения: 15.04.2025).
60. Wiseman J. Python interface to the WebRTC Voice Activity Detector. URL: <https://github.com/wiseman/py-webrtcvad> (дата обращения: 15.04.2025).
61. Silero VAD: pre-trained enterprise-grade Voice Activity Detector (VAD), Number Detector and Language Classifier. URL: <https://github.com/snakers4/silero-vad> (дата обращения: 15.04.2025).
62. McFee B. *et al.* librosa/librosa: 0.11.0. 2025. DOI: 10.5281/zenodo.15006942.
63. Кедрова Г. Е., Потанов В. В., Егоров А. М., Омелянова Е. Б. Акцентология. Ударение и фонетическое оформление слова. Фонетическая природа ударных гласных // Русская фонетика: учеб. материалы. 2002. URL: https://www.philol.msu.ru/~fonetica/akcent/phon_priroda/index.html (дата обращения: 09.06.2025).
64. Lubenets I., Davidchuk N., Amenets A. Emotions recognition from audio and text files URL: https://huggingface.co/datasets/Aniemore/resd_annotated (дата обращения: 15.04.2025).
65. Ardila R., Branson M., Davis K., *et al.* Common Voice: A Massively-Multilingual Speech Corpus. Common Voice. // Proceedings of the Twelfth Language Resources and Evaluation Conference. P. 4218–4222. ISBN: 979-10-95546-34-4.
66. Поволоцкая А. А., Карнов А. А. Аналитический обзор методов автоматического анализа экстралингвистических компонентов спонтанной речи // Информатика и автоматизация. 2024. Т. 23, № 1. С. 5–38.
67. Lee Y., Yeon I., Nam J., Chung J. S. VoiceLDM: Text-to-Speech with Environmental Context. VoiceLDM. 2023. URL: <http://arxiv.org/abs/2309.13664> (дата обращения: 15.04.2025).
68. Jung J., Ahn J., Jung C. VoiceDiT: Dual-Condition Diffusion Transformer for Environment-Aware Speech Synthesis. 2024. URL: <http://arxiv.org/abs/2412.19259> (дата обращения: 13.04.2025).

References

1. Diwan, A., Zheng, Z., Harwath, D. and Choi, E. (2025), "Scaling Rich Style-Prompted Text-to-Speech Datasets", DOI: 10.48550/arXiv.2503.04713.

2. Guo, Z., Leng, Y., Wu, Y., Zhao, S. and Tan, X. (2022), "PromptTTS: Controllable Text-to-Speech with Text Descriptions", *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, pp. 1–5, DOI: 10.1109/ICASSP49357.2023.10096285.
3. Lacombe, Y., Srivastav, V. and Gandhi, S., Inference and training library for high-quality TTS models, available at: <https://github.com/huggingface/parler-tts> (Accessed 4.15.25).
4. Kondratenko, V., Sokolov, A., Karpov, N., Kutuzov, O., Savushkin, N. and Minikin, F. (2022), "Large Raw Emotional Dataset with Aggregation Mechanism", DOI: 10.48550/arXiv.2212.12266.
5. Knyazev, S. V., Moroz, G. A. and Dyachenko, S. V. "Russian Speech Prosody Corpus (PRuD)", available at: <https://lingconlab.github.io/PRuD/> (Accessed 4.10.25).
6. Krivnova, O. F., Arkhipov, A. V., Zakharov, L. M. and Kobozeva, I. M. (2020), "Russian Discourse Intonation: RINCO – a Russian Speech Intonation Corpus", *Rechevye Tekhnologii*, no. 1-2, pp. 113–120, DOI: 10.58633/2305-8129_2020_1-2_113.
7. Yakovleva, E. B. (2016), *Rechevye khezitacii: formal'nyj i funkcional'nyj aspekty*. [Speech hesitations: formal and functional aspects], RAS Institute of Scientific Information for Social Sciences (INION), Moscow, Russia.
8. Gutnik, G. (2024), "Neural network for punctuation restoration in Russian Language", available at: https://github.com/gleb-skobinsky/ru_punct (Accessed 4.20.25).
9. Kotik, K. (2025), "Punctuation and casing restoration for the Russian Language (BERT-based)", available at: <https://github.com/kotikkonstantin/ru-autopunctuation> (Accessed 4.20.25).
10. Hwang, J.-S., Lee, S.-H. and Lee, S.-W. (2023), "PauseSpeech: Natural Speech Synthesis via Pre-trained Language Model and Pause-Based Prosody Modeling", *Pattern Recognition. Springer Nature*, pp. 415–427. DOI: 10.1007/978-3-031-47634-1_31.
11. Dai, Z. et al. (2022), "Automatic Prosody Annotation with Pre-Trained Text-Speech Model", *Proceedings of Interspeech 2022*, Incheon, Korea, pp. 5513–5517, DOI: 10.21437/Interspeech.2022-10005.
12. Bain, M., Huh, J., Han, T. and Zisserman, A. (2023), "WhisperX: Time-Accurate Speech Transcription of Long-Form Audio", *Proc. of Interspeech 2023*, Dublin, Ireland, pp. 4489–4493, DOI: 10.21437/Interspeech.2023-78.
13. RapidFuzz 3.13.0 documentation, available at: <https://rapidfuzz.github.io/RapidFuzz/index.html> (Accessed 4.20.25).
14. Polyakov, A., *Accenter/Accenter.txt at master. Accenter created by Alexey Polyakov*, GitHub, available at: <https://github.com/sStress/Accenter/blob/master/Accenter.txt> (Accessed 4.11.25).
15. Reynolds, R. and Tyers, F. (2015), "Automatic word stress annotation of Russian unrestricted text", *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)*, pp. 173–180.
16. Vaswani, A. et al. (2017). "Attention is All you Need", *Advances in Neural Information Processing Systems 2017*.
17. Gusev, I. (2019), Package for word stress detection, available at: <https://github.com/IlyaGusev/russ> (Accessed 4.11.25).
18. Zaliznyak, A. A. (1977), *Grammaticheskij slovar' russkogo yazyka* [Grammatical dictionary of Russian language], Moscow, available at: <https://gramdict.ru/> (Accessed 11.04.2025).
19. Grishina, E. A., Zelenkov, Y. G. and Orekhov, B. V. (2015), "Naive Poetry in an Accentological Corpus", *Proceedings of the V. V. Vinogradov Russian Language Institute*, Vol. 3, no. 6, pp. 257–272.

20. Koziev, I. (2025), Detection of poetic meter, rhyme, and stress placement in the texts of Russian accentual-syllabic poems and songs, available at: <https://github.com/RussianNLP/RussianPoetryScansionTool> (accessed 4.13.25).
21. Koziev, I. (2025), "Automated Evaluation of Meter and Rhyme in Russian Generative and Human-Authored Poetry", DOI: 10.48550/arXiv.2502.20931.
22. Ponomareva, M., Milintsevich, K., Chernyak, E. and Starostin, A. (2017), "Automated Word Stress Detection in Russian", *Proceedings of the First Workshop on Subword and Character Level Models in NLP*, pp. 31–35, DOI: 10.18653/v1/W17-4104.
23. Ponomareva, M. (2025), Python package russtress accentuates russian text, available at: <https://github.com/MashaPo/russtress> (Accessed 4.13.25).
24. Korotkova, Y. O. (2022), "Combined Dictionary-Neural Network Accentuator for Annotation of Russian Poetic Text", *Proceedings of the V. V. Vinogradov Russian Language Institute*, Vol. 3, no. 33, pp. 181–190, DOI: 10.31912/pvrli-2022.3.11.
25. Savchuk, S. O., Arkhangel'skij, T. A., Bonch-Osmolovskaya, A. A., Donina, O. V., Kuznecova, Y. N., Lyashevskaya, O. N., Orekhov, B. V. and Podryadchikova, M. V. (2024), "National Russian Language Corpus 2.0: new possibilities and developmental prospects", *Topics in the study of language*, pp. 7–34, DOI: 10.31857/0373-658X.2024.2.7-34.
26. Korotkova, Y. O. (2022), ru-accent-poet, a tool for putting stress marks in russian poetic texts, available at: https://github.com/yuliya1324/ru_accent (Accessed 4.11.25).
27. Knyaz'kov, A. A. (1998), *Pedagogicheskoe rechevedenie. Slovar'-spravochnik* [Pedagogical speech studies. Reference dictionary] in T.A. Ladyzhenskoy, and A.K. Mikhail'skoj (ed), available at: <http://rus-yaz.niv.ru/doc/pedagogical-speech/index.htm> (Accessed 11.04.2025).
28. Gerhard, D. (2003), *Pitch Extraction and Fundamental Frequency: History and Current Techniques*, Department of Computer Science, University of Regina.
29. Klofstad, C. A., Anderson, R. C. and Peters, S. (2012), "Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women", *Proceedings in Biological Sciences*, Vol. 279, № 1738, pp. 2698–2704, DOI: 10.1098/rspb.2012.0311.
30. O'Connor, J. J. M. and Barclay, P. (2017), "The influence of voice pitch on perceptions of trustworthiness across social contexts", *Evolution and Human Behavior*, 2017, Vol. 38, no. 4, pp. 506–512, DOI: 10.1016/j.evolhumbehav.2017.03.001.
31. Wang, T.-Y., Kawaguchi, I., Kuzuoka, H. and Otsuki, M. (2018), "Effect of Manipulated Amplitude and Frequency of Human Voice on Dominance and Persuasiveness in Audio Conferences", *Proc. ACM Human-Computer Interaction*, Vol. 2, pp. 177:1–177:18, DOI: 10.1145/3274446.
32. Wu, H. X., Li, Y., Ching, B. H.-H. and Chen, T. T. (2023), "You are how you speak: The roles of vocal pitch and semantic cues in shaping social perceptions", *Perception*, Vol. 52, no. 1, pp. 40–55, DOI: 10.1177/03010066221135472.
33. Imamverdiev, Y. N. and Sukhostat, L. V. (2014), "Approaches for estimating fundamental pitch period of speech signal in a noisy environment", *Rechevye Tekhnologii*, no. 1-2, pp. 84–103.
34. Ross, M. et al. (1974), "Average magnitude difference function pitch extractor", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 22, no. 5, pp. 353–362, DOI: 10.1109/TASSP.1974.1162598.
35. De Cheveigné, A. and Kawahara, H. (2002), "YIN, a fundamental frequency estimator for speech and music", *The Journal of the Acoustical Society of America*, Vol. 111, no. 4, pp. 1917–1930, DOI: 10.1121/1.1458024.

36. Mauch, M. and Dixon, S. (2014), "PYIN: A fundamental frequency estimator using probabilistic threshold distributions", *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, 2014, pp. 659–663, DOI: 10.1109/ICASSP.2014.6853678.
37. Morise, M. (2017), "Harvest: A High-Performance Fundamental Frequency Estimator from Speech Signals", *Proceedings of Interspeech 2017*, pp. 2321–2325, DOI: 10.21437/Interspeech.2017-68.
38. Kasi, K. and Zahorian, S. A. (2002), "Yet Another Algorithm for Pitch Tracking", *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. I-361–I-364, DOI: 10.1109/ICASSP.2002.5743729.
39. Hsu J. et al., PyWorld: a Python wrapper for WORLD vocoder, available at: <https://github.com/JeremyCCHsu/Python-Wrapper-for-World-Vocoder> (Accessed 4.20.25).
40. Morise, M., Kawahara, H. and Katayose, H. (2009), "Fast and Reliable F0 Estimation Method Based on the Period Extraction of Vocal Fold Vibration of Singing Voice and Speech", *Proceedings of the AES 35th International Conference: Audio for Games*, London, United Kingdom.
41. Schmitt, B. J. B., pYAAPT — AMFM_decompy 1.0.11 documentation, available at: https://bjbschmitt.github.io/AMFM_decompy/pYAAPT.html (Accessed 4.15.25).
42. Drugman, T., Huybrechts, G., Klimkov, V. and Moinet, A. (2018), "Traditional Machine Learning for Pitch Detection", *IEEE Signal Processing Letters*, Vol. 25, no. 11, pp. 1745–1749, DOI: 10.1109/LSP.2018.2874155
43. Kim, J. W., Salamon, J., Li, P. and Bello, J. P. (2018), "CREPE: A Convolutional Representation for Pitch Estimation", *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 161–165, DOI: 10.1109/ICASSP.2018.8461329.
44. Boersma, P., and van Heuven, V. (2001), "Praat, a system for doing phonetics by computer", *Glott International*, Vol. 5, no. 9/10, pp. 341–345.
45. Jadoul, Y., Thompson, B. and de Boer, B. (2018), "Introducing Parselmouth: A Python interface to Praat", *Journal of Phonetics*, Vol. 71, pp. 1–15, DOI: 10.1016/j.wocn.2018.07.001.
46. Boersma, P. (1993), "Accurate Short-Term Analysis of The Fundamental Frequency and The Harmonics-To-Noise Ratio of a Sampled Sound", *Proceedings of the institute of phonetic sciences*, Vol. 17, no. 1193, pp. 97–110.
47. Thinakaran, P. et al. (2024), "Utilizing POS-Driven Pitch Contour Analysis for Enhanced Tamil Text-to-Speech Synthesis", *Proceedings of the 21st International Conference on Natural Language Processing (ICON)*, 2024, pp. 269–273.
48. Hai, J., Thakkar, K., Wang, H., Qin, Z. and Elhilali, M. (2024), "DreamVoice: Text-Guided Voice Conversion", *Proceedings of Interspeech 2024*, pp. 4373–4377, DOI: 10.21437/Interspeech.2024-1432.
49. Kawamura, M., Yamamoto, R., Shirahata, Y., Hasumi, T. and Tachibana, K. (2024), "LibriTTS-P: A Corpus with Speaking Style and Speaker Identity Prompts for Text-to-Speech and Style Captioning", *Proceedings of Interspeech 2024*, pp. 1850–1854, DOI: 10.21437/Interspeech.2024-692.
50. Nguyen, T. A. et al. (2023), "EXPRESSO: A Benchmark and Analysis of Discrete Expressive Speech Resynthesis", *Proceedings of Interspeech 2023*, pp. 4823–4827. DOI: 10.21437/Interspeech.2023-1905.
51. Richter, J. et al. (2024), "EARS: An Anechoic Fullband Speech Dataset Benchmarked for Speech Enhancement and Dereverberation", *Proceedings of Interspeech 2024*, pp. 4873–4877, DOI: 10.21437/Interspeech.2024-153.

52. Ji, S., Zuo, J., Fang, M. *et al.* (2024), "TextrolSpeech: A Text Style Control Speech Corpus with Codec Language Text-to-Speech Models", *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 10301–10305, DOI: 10.1109/ICASSP48485.2024.10445879.
53. Guan, W., Li, Y., Li, T., Huang, H. *et al.* (2024), "MM-TTS: Multi-modal Prompt based Style Transfer for Expressive Text-to-Speech Synthesis", *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence*, pp. 18117–18125, DOI: 10.1609/aaai.v38i16.29769.
54. Jin, Z. *et al.* (2024), "SpeechCraft: A Fine-Grained Expressive Speech Dataset with Natural Language Description", *Proceedings of the 32nd ACM International Conference on Multimedia, MM '24*, pp. 1255–1264, DOI: 10.1145/3664647.3681674.
55. Watanabe, A. *et al.* (2023), "Coco-Nut: Corpus of Japanese Utterance and Voice Characteristics Description for Prompt-based Control", *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Taipei, Taiwan, pp. 1–8, DOI: 10.1109/ASRU57964.2023.10389693.
56. Nagrani, A., Chung, J. S., Xie, W. and Zisserman, A. (2020), "Voxceleb: Large-scale speaker verification in the wild", *Computer Speech & Language*, Vol. 60, pp. 101027, DOI: 10.1016/j.csl.2019.101027.
57. Wagner, J., Triantafyllopoulos, A. *et al.* (2023), "Dawn of the transformer era in speech emotion recognition: closing the valence gap", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, no. 9, pp. 10745–10759, DOI: 10.1109/TPAMI.2023.3263585.
58. Sendlmeier, W., Burkhardt, F., Kienast, M., Paeschke, A. and Weiss, B., *The Berlin Database of Emotional Speech*, available at: <http://emodb.bilderbar.info/docu/> (Accessed 09.06.25).
59. Meng, R. *et al.*, *SFR-Embedding-Mistral: Enhance Text Retrieval with Transfer Learning*, available at: <https://huggingface.co/Salesforce/SFR-Embedding-Mistral> (Accessed 4.15.25).
60. Wiseman, J., *Python interface to the WebRTC Voice Activity Detector*, available at: <https://github.com/wiseman/py-webrtcvad> (accessed 4.15.25).
61. Silero VAD: pre-trained enterprise-grade Voice Activity Detector (VAD), *Number Detector and Language Classifier*, available at: <https://github.com/snakers4/silero-vad> (Accessed 15.04.2025).
62. McFee B. *et al.*, *librosa/librosa: 0.11.0*. 2025, DOI: 10.5281/zenodo.15006942.
63. Kedrova G. E., Potapov V. V., Egorov A. M. and Omelyanova E. B. (2002), *Akcentologiya. Udarenie i foneticheskoe oformlenie slova. Foneticheskaya priroda udarnyh glasnyh. Russkaya fonetika. Uchebnye materialy*. [Accentology. Word Stress and Phonetic Word Framing. Phonetic Nature of Stressed Vowels. Russian Phonetics. Learning materials.], available at: https://www.philol.msu.ru/~fonetika/akcent/phon_priroda/index.html (Accessed: 09.06.2025).
64. Lubenets I., Davidchuk N. and Amenets A., *Emotions recognition from audio and text files*, available at: https://huggingface.co/datasets/Aniemore/resd_annotated (Accessed 4.15.25).
65. Ardila, R., Branson, M., Davis, K. *et al.* (2020), "Common Voice: A Massively-Multilingual Speech Corpus", *Proceedings of the Twelfth Language Resources and Evaluation Conference*, Marseille, France, pp. 4218–4222, ISBN: 979-10-95546-34-4.

66. Povolockaya, A. A. and Karpov, A. A. (2024), "Analysis of methods for automating extralinguistic component analysis in spontaneous speech. ", *Informatika I Avtomatizaciya*, Vol. 23, no. 1., pp. 5–38, DOI: 10.15622/ia.23.1.1.
67. Lee, Y., Yeon, I., Nam, J. and Chung, J. S. (2023), "VoiceLDM: Text-to-Speech with Environmental Context", DOI: 10.48550/arXiv.2309.13664.
68. Jung, J., Ahn, J., Jung, C., Nguyen, T. D., Jang, Y. and Chung, J. S. (2024), "VoiceDiT: Dual-Condition Diffusion Transformer for Environment-Aware Speech Synthesis", DOI: 10.48550/arXiv.2412.19259.

Информация об авторах:

Е. Н. Радченко – студент 1-го курса магистратуры, Национальный исследовательский технологический университет "МИСИС" (119049, Россия, г. Москва, Ленинский пр-кт, д. 4, стр. 1);

Е. В. Исаева – кандидат филологических наук, доцент, зав. кафедрой английского языка профессиональной коммуникации, Пермский государственный национальный исследовательский университет (614068, Россия, г. Пермь, ул. Букирева, 15), ScopusAuthorID: 57204498718, ResearcherID: O-6777-2015.

Information about the authors:

E. N. Radchenko – 1st year Master's student, National University of Science and Technology "MISIS" (4, B. 1, Leninsky pr. Moscow, Russia, 119049).

E. V. Isaeva – Candidate of Science (in philology), Associate Professor, Head of the Department of English Language of Professional Communication, Perm State University (15, Bukireva St., Perm, Russia, 614068), Scopus Author ID: 57204498718, ResearcherID: O-6777-2015.

Научная статья

УДК 681.51, 517.977, 004.94

DOI: 10.17072/1993-0550-2025-2-123-136

<https://elibrary.ru/lasrhw>



Решение задачи расстановки обнаружителей для охраны периметра градиентным методом

Марина Александровна Самохина ¹, Александр Сергеевич Самохин ²

^{1,2}Институт проблем управления имени В.А. Трапезникова Российской академии наук,
г. Москва, Россия

¹ph@ipu.ru

²samokhin@ipu.ru

Аннотация. Работа посвящена математическим аспектам создания современной охранной системы. Разработан метод численного решения задачи оптимизации расстановки обнаружителей с целью противодействия прорыву защищаемого периметра. Защищаемый периметр представляет собой отрезок на плоскости, который пытается за заданное время пересечь уклоняющийся от обнаружения подвижный объект. Данный объект рассматривается как материальная точка, управляемая с целью минимизации функционала риска обнаружения по первичному гидроакустическому полю. Задача формализуется как задача максимина, обнаружители следует расставить в допустимой области таким образом, чтобы минимально возможное значение функционала подвижного объекта было наибольшим. Для решения авторами был разработан программный комплекс на языке C++. В работе приводятся результаты численного моделирования, полученные с использованием градиентного метода и решения краевой задачи принципа максимума Л.С. Понтрягина для нахождения локально оптимальных траекторий во вспомогательной задаче поиска пути подвижным объектом. Краевая задача решалась методом стрельбы, соответствующие задачи Коши изначально формировались заданием значений параметров пристрелки на сетке в пространстве их возможных значений, и интегрировались численно методом Рунге–Кутты с автоматическим выбором шага. Далее значения параметров пристрелки уточнялись модифицированным методом Ньютона. В результате численного моделирования расстановки пяти сенсоров оказалось, что выгоднее всего расставлять их таким образом, чтобы первые шесть лучших локально оптимальных траекторий уклоняющегося объекта совпадали по функционалу.

Ключевые слова: *расстановка обнаружителей; конфликтная среда; противодействие уклонению; планирование пути; оптимизация; градиентный метод; численное моделирование; защита периметра; предотвращение прорыва; максимин*

Для цитирования: Самохина М. А., Самохин А. С. Решение задачи расстановки обнаружителей для охраны периметра градиентным методом // Вестник Пермского университета. Математика. Механика. Информатика. 2025. Вып. 2(69). С. 123–136. DOI: 10.17072/1993-0550-2025-2-123-136. <https://elibrary.ru/lasrhw>

Благодарности: работа выполнена при частичной поддержке Российского научного фонда, проект № 23-19-00134; авторы выражают благодарность Галяеву Андрею Алексеевичу за плодотворное обсуждение работы.



Эта работа © 2025 Самохина М. А., Самохин А. С. распространяется по лицензии CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <https://creativecommons.org/licenses/by/4.0/>.

Статья поступила в редакцию 30.04.2025; одобрена после рецензирования 21.05.2025; принята к публикации 11.07.2025.

Research article

Solution of the Sensors Placement Problem for Perimeter Defense Using the Gradient Method

Marina A. Samokhina¹, Alexander S. Samokhin²

^{1,2}V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

¹ph@ipu.ru

²samokhin@ipu.ru

Abstract. The research is devoted to the mathematical aspects of creating a modern security system. The paper develops a method that numerically solves the problem of optimizing the deployment of sensors to counteract the breach of a protected perimeter. The protected perimeter is a segment on the plane, and a mobile object evading detection tries to cross it in a given time. This object is considered as a material point, controlled to minimize the risk functional of detection by the primary hydroacoustic field. The problem is formalized as a maximin problem, the sensors should be placed in the admissible area so that the minimum possible value of the functional of the mobile object was as large as possible. For the solution the authors have developed a program in C++ language. The paper presents the results of numerical modeling obtained using the gradient method and the solution of the boundary value problem of the maximum principle of L.S. Pontryagin to find locally optimal trajectories in the auxiliary problem of pathfinding by a mobile object. The boundary value problem was solved by the shooting method, the corresponding Cauchy problems were initially formed by setting the values of the shooting parameters on the grid of their possible values, and were integrated numerically by Runge-Kutta method with automatic step selection. Further, the values of shooting parameters were refined by the modified Newton's method. As a result of numerical modeling of 5 sensors placement, it turned out that it is most advantageous to place them in such a way that the first 6 best locally optimal trajectories of the evading object coincide in functional value.

Keywords: sensors placement; threat environment; evasion countering; path planning; optimization; gradient method; numerical simulations; perimeter defense; breach prevention; maximin

For citation: Samokhina, M. A. and Samokhin, A. S. (2025), "Solution of the Sensors Placement Problem for Perimeter Defense Using the Gradient Method", *Bulletin of Perm University. Mathematics. Mechanics. Computer Science*, no. 2(69), pp. 123–136. (In Russ.). DOI: 10.17072/1993-0550-2025-2-123-136. <https://elibrary.ru/lasrhw>

Acknowledgments: the work was supported in part by the Russian Science Foundation, Project № 23-19-00134; the authors would like to thank Andrei A. Galyaev for fruitful discussion of the research.

The article was submitted 30.04.2025; approved after reviewing 21.05.2025; accepted for publication 11.07.2025.

Введение

В настоящее время все более широкое применение находят автономные необитаемые подводные аппараты. Такие устройства используются в научном, коммерческом, политическом секторах [1–3]. В том числе они могут использоваться для несанкционированного доступа к охраняемому объекту. Работа посвящена созданию современной

охранной системы противодействия такому доступу к протяженному защищаемому объекту, например газопроводу.

Предполагается, что система охраны строится из неподвижных обнаружителей, совместная задача которых заключается в противодействии скрытному перемещению автономного аппарата [4–9]. Настоящее исследование является развитием работ [10–12].

1. Постановка задачи

Считается, что обнаружители обладают равномерно распространяющимися бесконечными круговыми полями обнаружения, и их расположение, формирующее карту угроз [13], уклоняющемуся подвижному объекту (УПО) известно. Совместно УПО и обнаружители формируют конфликтную среду [14]. При заданной конфигурации обнаружителей УПО перемещается таким образом, чтобы минимизировать интегральный функционал – риск обнаружения [15].

Задача заключается в оптимизации расстановки неподвижных обнаружителей с целью максимизации минимально возможного значения функционала, оптимизируемого УПО. Рассматривается плоская модель, в которой УПО и обнаружители представляют собой материальные точки.

2. Формализация задачи

На плоскости вводится декартова система координат, без ограничения общности считается, что УПО перемещается из начала координат в конечный отрезок с ординатой 1 за фиксированное время T с максимальной скоростью v_{\max} и управляется модулем вектора скорости v и углом φ , задающим его направление, так, чтобы минимизировать функционал:

$$I = \int_0^T \left(\sum_{i=1}^N q_i \frac{v^2}{(x - a_i)^2 + (y - b_i)^2} \right) dt, \quad (2.1)$$

зависящий от величины вектора скорости УПО v , координат УПО (x, y) , расстояний $r_i = \sqrt{(x - a_i)^2 + (y - b_i)^2}$ от УПО до каждого из обнаружителей S_i , заданных координатами (a_i, b_i) , весовых коэффициентов q_i влияния обнаружителей S_i и количества обнаружителей $N \geq 1$. Конечный отрезок выбирается достаточно длинным, таким, что его края недостижимы при заданных ограничениях на время и скорость передвижения. Нужно расположить точки S_i в области $y \geq Y$ так, чтобы по всем возможным наборам фиксированных расположений обнаружителей максимизировать: J – глобальный минимум функционала (2.1) УПО $J = \min_{(x(\cdot), y(\cdot))} I$, определяемого для каждого заданного расположения S_i , т.е. решить задачу максимизации минимума функционала I :

$$\hat{J} = \max_{\{(a_1, b_1), \dots, (a_N, b_N)\}} \min_{(x(\cdot), y(\cdot))} I.$$

3. Метод решения

Считается, что ни в один момент времени $t \in [0, T]$ траектория УПО не проходит ни через один обнаружитель, $r_i(t) > 0$. Тогда задача поиска пути для фиксированного расположения обнаружителей на основе принципа максимума Понтрягина сводится к краевой задаче [16]:

$$\begin{cases} \dot{x} = \hat{v} \cos \hat{\phi}, \\ \dot{y} = \hat{v} \sin \hat{\phi}, \\ \dot{p}_x = -\sum_{i=1}^N q_i \frac{\hat{v}^2}{((x-a_i)^2 + (y-b_i)^2)^2} \cdot (x-a_i), \\ \dot{p}_y = -\sum_{i=1}^N q_i \frac{\hat{v}^2}{((x-a_i)^2 + (y-b_i)^2)^2} \cdot (y-b_i), \end{cases} \quad (3.1)$$

$$\text{где } \hat{v} = \min \left(\frac{\|p\|}{\sum_{i=1}^N \frac{q_i}{(x-a_i)^2 + (y-b_i)^2}}, v_{\max} \right), \quad (3.2)$$

$$\hat{\phi} : \cos \hat{\phi} = \frac{p_x}{\|p\|}, \sin \hat{\phi} = \frac{p_y}{\|p\|}, \|p\| = \sqrt{p_x^2 + p_y^2}, \quad (3.3)$$

$$\begin{cases} x(0) = 0, & p_x(T) = 0, \\ y(0) = 0, & y(T) = 1. \end{cases} \quad (3.4)$$

Полученная краевая задача решается численно методом стрельбы. Для этого недостающие для формирования задач Коши параметры пристрелки при задании начального приближения перебираются на сетке своих возможных значений, а далее уточняются модифицированным многомерным методом Ньютона. Сами задачи Коши интегрируются методом Рунге–Кутты с автоматическим выбором шага на основе контроля максимально допустимой локальной погрешности [17]. Найденные таким образом экстремали Понтрягина удовлетворяют необходимым условиям оптимальности. Среди всех построенных для фиксированного расположения обнаружителей траекторий выбирается траектория, соответствующая минимальному значению функционала I .

Для нахождения максимума функционала J по всем возможным расположениям обнаружителей S_i запускается метод проекции градиента [18] со стандартным условием остановки счета. Опишем его работу. Обозначим за $z_k = (a_1^k, b_1^k, a_2^k, b_2^k, \dots, a_N^k, b_N^k)^T$ вектор оптимизируемых параметров на k -м шаге. Тогда переход от z_k к z_{k+1} осуществляется по формуле

$$z_{k+1} = P(z_k + \theta_k \nabla J(z_k)), \quad (3.5)$$

где P – оператор проектирования на полуплоскость $y \geq Y$, а параметры $\theta_k > 0$ подбираются на каждом шаге решения задачи, так, чтобы сохранялась монотонность: $J(z_{k+1}) > J(z_k)$. При этом компоненты градиента функционала

$$\nabla J(z_k) = \left(\frac{\partial J(z_k)}{\partial a_1}, \frac{\partial J(z_k)}{\partial b_1}, \dots, \frac{\partial J(z_k)}{\partial a_N}, \frac{\partial J(z_k)}{\partial b_N} \right) \text{ в точке } z_k \text{ находятся численно для всех}$$

$$j=1, \dots, N \text{ по формулам } \frac{\partial J(z_k)}{\partial a_j} = \frac{J(z_k + h_{2j-1}) - J(z_k)}{h}, \frac{\partial J(z_k)}{\partial b_j} = \frac{J(z_k + h_{2j}) - J(z_k)}{h}, \text{ где}$$

h_l – вектор, у которого l -ная координата равна h , а все остальные координаты равны 0,

таким образом параметр h – шаг численного дифференцирования. Заметим, что для вычисления градиента $\nabla J(z_k)$ на шаге требуется $2N$ дополнительных вычислений функционала J , для чего в свою очередь решается $2N$ краевых задач (3.1)-(3.4).

Отметим, что при проектировании (3.5) на полуплоскость $y \geq Y$ точки (a_i, b_i) с координатой $b_i \geq Y$ сохраняются, а точки (a_i, b_i) с координатой $b_i < Y$ смещаются на прямую $y = Y$, переходя в точки с координатами (a_i, Y) .

4. Результаты численного моделирования

При моделировании в качестве одного из примеров для расчетов были выбраны следующие значения параметров задачи: $N = 5$, $q_i = 1$, $T = 2.6$, $v_{\max} = 1$, $Y = 0.6$ и отрезок финиша AB , $A = (-3, 1)$, $B = (3, 1)$, см. рис. 1. Для каждого фиксированного на шаге градиентного метода расположения обнаружителей пути УПО строились на основе решения краевой задачи принципа максимума [16]. Обнаружители разрешалось размещать в полуплоскости $y \geq Y$.

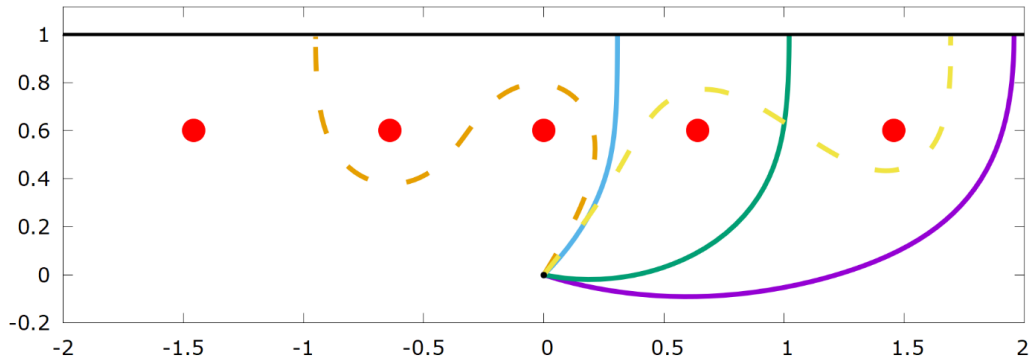


Рис. 1. Оптимальная конфигурация обнаружителей и соответствующая выходу УПО в правую полуплоскость из точки старта $(0,0)$ карта классификации локально оптимальных путей УПО для $N = 5$, $q_i = 1$, $T = 2.6$, $v_{\max} = 1$. Сплошными и прерывистыми линиями изображены локально оптимальные траектории УПО, красными кругами – места расположения обнаружителей S_i

Метод проекции градиента (3.5), запущенный из разных стартовых конфигураций точек по 10 параметрам – координатам S_i , каждый раз уменьшал b_i до выхода на ограничение $Y = 0.6$. При одинаковых весовых коэффициентах влияния q_i точки (a_i, b_i) располагались на прямой $y = Y$ симметрично перпендикуляру, опущенному из точки старта на конечный отрезок: $a_1 = -\alpha_1$, $a_2 = -\alpha_2$, $a_3 = 0$, $a_4 = \alpha_2$, $a_5 = \alpha_1$, при некоторых $\alpha_1, \alpha_2 \in \mathbb{R}$. Таким образом, в случае одинаковых q_i для различных T , v_{\max} , Y изначальная 10-параметрическая задача в итоге сводилась к 2-параметрической задаче поиска величин α_1, α_2 , см. рис. 2.

На рис. 1 показаны траектории УПО, выходящие из точки старта под углом $\varphi(0) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Также локально оптимальными являются траектории, симметричные им относительно перпендикуляра из точки старта на конечный отрезок.

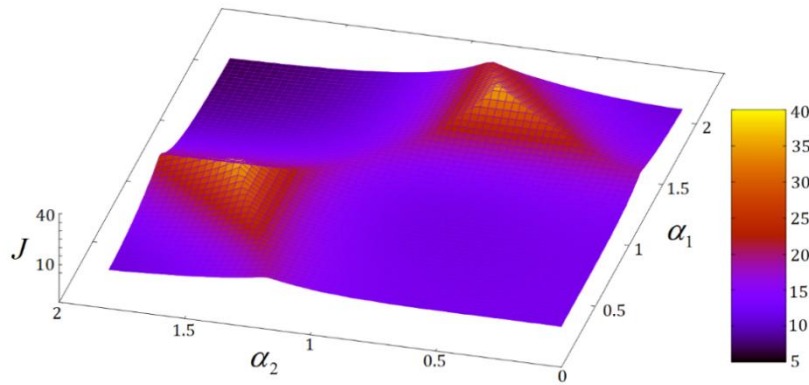


Рис. 2. Зависимость функционала задачи расстановки обнаружителей J от параметров α_1, α_2 при $N=5$, $q_i=1$, $T=2.6$, $v_{\max}=1$

В результате проведенного моделирования оказалось, что при оптимальном расположении пяти обнаружителей $\alpha_1 = 0.64$, $\alpha_2 = 1.46$, и получилась конфигурация, при которой на лучших шести локально оптимальных траекториях УПО достигается значение функционала $I = 38.3$, и, следовательно, $\hat{J} = 38.3$. Для выбранных параметров задачи это оптимальное значение достигается на трех траекториях, изображенных сплошными линиями на рис. 1 и трех симметричных им относительно перпендикуляра из точки старта на конечный отрезок.

График функционала J на рис. 2 получился ожидаемо симметричный по параметрам α_1, α_2 . По нему можно сделать вывод о единственности найденного решения задачи.

5. Аналитическое исследование

Помимо проведения численного моделирования были получены следующие теоретические результаты.

Лемма. При $v_{\max}T = 1 + \delta_1$, $\delta_1 > 0$ для любой расстановки N обнаружителей в полуплоскости $y \geq \delta_2$, $\delta_2 > 0$ существует решение задачи поиска пути УПО из точки $O(0,0)$ в отрезок E_1E_2 , $E_1 = (-c,1)$, $E_2 = (c,1)$, $c > 0$, на котором функционал I ограничен.

Доказательство. Построим траекторию $(x(t), y(t))$, не проходящую через N фиксированных обнаружителей с координатами (a_i, b_i) , $i \in \{1, \dots, N\}$.

Обозначим $\varepsilon = \min\left(\frac{\delta_1}{2\pi N}, \frac{\delta_2}{2}, \frac{c}{N}, 1\right)$ и вокруг каждого обнаружителя $S_i(a_i, b_i)$, $i \in \{1, \dots, N\}$ построим открытый шар $B(S_i, \varepsilon)$ с центром $S_i(a_i, b_i)$ и радиусом ε . Обозначим $E(0,1)$ ближайшую к точке старта O точку на конечном отрезке E_1E_2 . Построим пошагово плоское множество Φ , граничные точки которого $\partial\Phi$ будут содержать допустимые траектории УПО, следующим образом. Положим $\Phi_0 = OE$, $\Theta_0 = \{1, \dots, N\}$.

Пусть после k шагов построены плоское множество Φ_k и Θ_k – множество индексов еще не включенных в Φ_k шаров. Сделаем $(k+1)$ -й шаг построения:

$\mathcal{D}_k = \{i: i \in \Theta_k, B(S_i, \varepsilon) \cap \Phi_k \neq \emptyset\}$, тогда если $\mathcal{D}_k = \emptyset$, то на этом построение заканчивается, и $\Phi = \Phi_k$, $\Theta = \Theta_k$, иначе $\Phi_{k+1} = \Phi_k \cup \left(\bigcup_{i \in \mathcal{D}_k} \bar{B}(S_i, \varepsilon) \right)$ и $\Theta_{k+1} = \Theta_k \setminus \mathcal{D}_k$. Заметим, что в множестве Θ_0 изначально N элементов, и на каждом шаге из него удаляется хотя бы один элемент, поэтому не более, чем за $N < \infty$ шагов множества Φ и Θ будут построены, см. рис. 3 а).

Для $\partial\Phi$ будут выполнены следующие свойства:

- 1) $\partial\Phi$ состоит из конечного числа отрезков и конечного числа частей окружностей (по построению);
- 2) любая точка $\partial\Phi$ удалена от каждого из обнаружителей S_i не менее, чем на ε (по построению);
- 3) $O \in \partial\Phi$, так как все $b_i \geq \delta_2$, $i \in \{1, \dots, N\}$ и радиусы шаров $\bar{B}(S_i, \varepsilon)$ не более $\frac{\delta_2}{2}$,

поэтому ордината любой точки каждого из шаров не менее $\delta_2 - \frac{\delta_2}{2} = \frac{\delta_2}{2} > 0$, и

$O \notin \bar{B}(S_i, \varepsilon)$, $i \in \{1, \dots, N\}$, но $O \in \partial\Phi$ в начале построения;

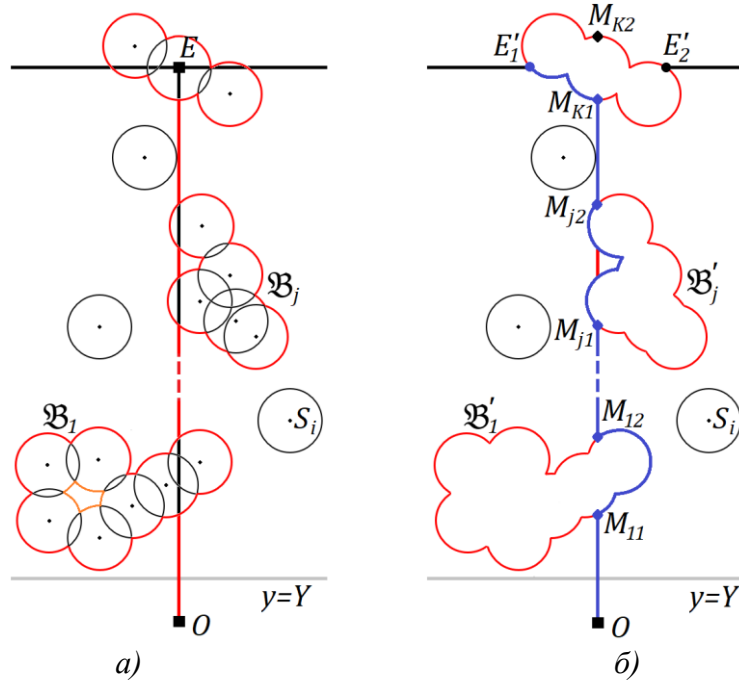


Рис. 3 а) Плоское множество Φ , часть границы которого $\partial\Phi$, обозначенная красной линией, содержит допустимые траектории УПО; б) Плоское множество Φ' и траектория УПО, показанная синей линией

- 4) отрезок финиша $E_1E_2 \cap \partial\Phi \neq \emptyset$, так как $E \in E_1E_2$, $E \in \Phi$ вначале и благодаря выбору ε ;
- 5) объединение всех шаров не покрывает отрезок E_1E_2 целиком, то есть

$$E_1E_2 \not\subset \bigcup_{i=1}^N B(S_i, \varepsilon) \text{ благодаря выбору } \varepsilon;$$

б) по построению, общая длина границы $\ell(\partial\Phi) \leq OE + 2\pi\varepsilon \cdot N \leq 1 + 2\pi\varepsilon \cdot N \leq 1 + \delta_1$.

Покажем, что можно выбрать траекторию движения из начальной точки O в какую-то точку конечного интервала, как часть $\partial\Phi$.

Для этого покажем, что можно разбить все шары $\bar{B}(S_i, \varepsilon) \subset \Phi$ на конечное число непересекающихся объединений \mathcal{B}_j . Обозначим $\mathcal{G}_0 = \{1, \dots, N\} \setminus \Theta$ – множество индексов шаров, лежащих в Φ .

Построим пошагово плоскую фигуру \mathcal{B}_1 . Выберем какой-нибудь индекс $i \in \mathcal{G}_0$. Положим $\mathcal{B}_1^1 = B(S_i, \varepsilon)$, $\mathcal{G}_1^1 = \mathcal{G}_0 \setminus \{i\}$. Пусть после k шагов построены фигура \mathcal{B}_1^k и \mathcal{G}_1^k – множество индексов шаров, входящих в фигуру Φ , но еще не включенных в \mathcal{B}_1^k . Сделаем $(k+1)$ -й шаг построения \mathcal{B}_1 : $\mathcal{D}_1^k = \{i : i \in \mathcal{G}_1^k, \bar{B}(S_i, \varepsilon) \cap \mathcal{B}_1^k \neq \emptyset\}$, тогда, если $\mathcal{D}_1^k = \emptyset$, то на этом построение заканчивается, и $\mathcal{B}_1 = \mathcal{B}_1^k$, $\mathcal{G}_1 = \mathcal{G}_1^k$, иначе

$\mathcal{B}_1^{k+1} = \mathcal{B}_1^k \cup \left(\bigcup_{i \in \mathcal{D}_1^k} \bar{B}(S_i, \varepsilon) \right)$ и $\mathcal{G}_1^{k+1} = \mathcal{G}_1^k \setminus \mathcal{D}_1^k$. Заметим, что в множестве \mathcal{G}_0 изначально не

более N элементов, и каждый шаг из него удаляется хотя бы один элемент, поэтому не более чем за $N < \infty$ шагов фигура \mathcal{B}_1 и множество \mathcal{G}_1 будут построены. Остальные плоские фигуры \mathcal{B}_j и множества индексов \mathcal{G}_j строятся последовательно по аналогии. Так как каждая из построенных фигур \mathcal{B}_j включает в себя хотя бы один шар, то количество элементов в \mathcal{G}_j с каждым новым построенным множеством уменьшается хотя бы на 1, поэтому за $K \leq N < \infty$ шагов, множество \mathcal{G}_K станет пустым, и приведенный выше алгоритм завершится.

Заметим, что все $\mathcal{B}_j \setminus \partial\mathcal{B}_j$ являются конечносвязными областями, получающимися из некоторой односвязной области D_j^0 исключением лежащих в ней односвязных областей $D_j^1, D_j^2, \dots, D_j^{N_j}$, а также их границ. Положим $\mathcal{B}_j' = D_j^0 \cup \partial D_j^0, j = 1, \dots, K$ и $\Phi' = OE \cup \left(\bigcup_{j=1}^K \mathcal{B}_j' \right)$, тогда $\Phi \subseteq \Phi'$, Φ' является односвязным замкнутым ограниченным множеством и $\partial\Phi'$ совпадает с внешней границей фигуры Φ , т.е. $\partial\Phi' \subseteq \partial\Phi$.

Теперь построим путь из точки O до конечного отрезка E_1E_2 , проходящий по $\partial\Phi'$.

Для каждого $j \in \{1, \dots, K\}$ множество точек пересечения $\partial\mathcal{B}_j'$ с лучом OE будет содержать от 1 до $2N$ точек, причем у всех таких точек будут разные ординаты. Обозначим точки пересечения $\partial\mathcal{B}_j'$ с отрезком OE с наименьшей и с наибольшей ординатами

M_{j1} и M_{j2} соответственно. Тогда M_{j1} лежит в пределах от $\frac{\delta_2}{2}$ до 1, M_{j2} лежит в пределах от $\frac{\delta_2}{2} + 2\varepsilon$ до $1 + 2\varepsilon N < \infty$ и при движении от точки M_{j1} к точке M_{j2} по $\partial\mathcal{B}_j'$ ордината УПО увеличится. M_{j1} и M_{j2} делят $\partial\mathcal{B}_j$ на две части, при движении УПО от M_{j1} к M_{j2} по одной из них область будет находиться слева от УПО, а по другой – справа. В качестве пути выберем ту часть границы $\partial\mathcal{B}_j$, длина которой меньше, и любую, если длины равны.

Вначале от точки O будем двигаться по отрезку OE "вверх" (с увеличением ординаты) до первой точки M_{k_1} с наименьшей ординатой среди всех точек $M_{j_1}, j \in \{1, \dots, K\}$, либо до точки E , если $K = 0$. Построение пути заканчивается, когда УПО впервые оказывается в точке с ординатой 1. Далее УПО движется по границе $\partial\mathcal{B}_k'$ до точки M_{k_2} . Если ордината точки M_{k_2} больше 1, то по непрерывности кривой $\partial\mathcal{B}_k'$, на ней найдется точка с ординатой 1, а конечной точкой пути будет ближайшая к M_{k_1} такая точка. Если ордината точки M_{k_2} равна 1, конечной точкой пути будет M_{k_2} . Если же ордината точки M_{k_2} меньше 1, то далее опять чередуются движение по отрезку OE с движением по границе $\partial\mathcal{B}_j'$, причем M_{j_1} каждый раз, кроме первого, выбирается из множества $M_{j_1}, j \in \{1, \dots, K\}$, ордината которых больше ординаты M_{j_2} на предыдущем шаге. Заметим, что чередоваться движения УПО по отрезку и по $\partial\mathcal{B}_j', j \in \{1, \dots, K\}$ могут не более $K \leq N < \infty$ раз. Если построение пути не закончено, то далее УПО движется по отрезку OE до точки $E(0,1)$, которая будет конечной точкой пути.

Таким образом, если УПО будет двигаться из точки O в направлении к конечному интервалу по $\partial\Phi' \subseteq \partial\Phi$ указанным выше образом, то траектория движения $(x(t), y(t))$, не будет содержать ни одного обнаружителя, ее длина не будет превосходить $1 + \delta_1$, а значение функционала I^* на такой траектории можно оценить следующим образом:

$$I^* \leq N \frac{q_{\max} v_{\max}^2 T}{r_{\min}^2} = N \frac{q_{\max} v_{\max} (1 + \delta_1)}{\varepsilon^2} < \infty.$$

Замечание 1. Выбор $\varepsilon \leq \frac{c}{N}$ гарантирует наличие точек на конечном отрезке, не покрытых шарами, построенными вокруг S_i , выбор $\varepsilon \leq \frac{\delta_2}{2}$ гарантирует, что точка старта УПО не покрыта ни одним из построенных шаров. Выбор $\varepsilon \leq \frac{\delta_1}{N}$ обеспечивает то, что множество проекций построенных шаров на отрезок $E_1 E_2 \cap [(1 - \delta_1, 1), (1 + \delta_1, 1)]$ не покрывает его целиком, а значит на нем есть некоторая непокрытая точка $P(x_p, 1)$, и тогда для УПО есть более простой по конфигурации путь, состоящий из горизонтального участка $\overline{OP_0}$, где $P_0 = (x_p, 0)$, затем вертикального $\overline{P_0 P}$, проходящий на удалении не менее ε от каждого из обнаружителей. Такой путь (изображен зеленой линией на рис. 4) может значительно отклоняться от отрезка OE по сравнению с путем, построенным в лемме, изображенным синей линией на рис. 4.

Замечание 2. Приведенный в лемме способ построения траектории движения применим также при условии заранее неизвестного УПО положения обнаружителей, но при известном их количестве и возможности их обнаружения и изменения направления движения УПО на расстоянии не менее ε . Оценка сверху для риска обнаружения I^* на построенной таким образом траектории не изменится.

Теорема. При $v_{\max} T = 1 + \delta_1$, $\delta_1 > 0$ и $y \geq \delta_2 > 0$ функционал J задачи расстановки обнаружителей с целью противодействия УПО, перемещающемуся из точки $O(0,0)$ в отрезок $E_1 E_2$, $E_1 = (-c, 1)$, $E_2 = (c, 1)$, $c > 0$, ограничен.

Доказательство. Доказанная в лемме оценка сверху для функционала I не зависит от расположения обнаружителей $S_i(a_i, b_i)$, $i \in \{1, \dots, N\}$. Поэтому

$$J = \min_{(x(\cdot), y(\cdot))} I \leq I^* \leq N \frac{q_{\max} v_{\max} (1 + \delta_1)}{\varepsilon^2} < \infty, \text{ где } \varepsilon = \min \left(\frac{\delta_1}{2\pi N}, \frac{\delta_2}{2}, \frac{c}{N}, 1 \right).$$

Замечание 3. Лемма позволяет УПО добраться до некоторой точки любого отрезка длины $2c > 0$. Если же в соответствии с изначальной постановкой c таково, что с заданным ограничением на скорость УПО фиксированного времени не хватит, чтобы добраться до концов отрезка $E_1 E_2$, то член $\frac{c}{N}$ в определении ε в доказательстве леммы и теоремы избыточен, так как в этом случае из теоремы Пифагора следует, что $c > \delta_1$.

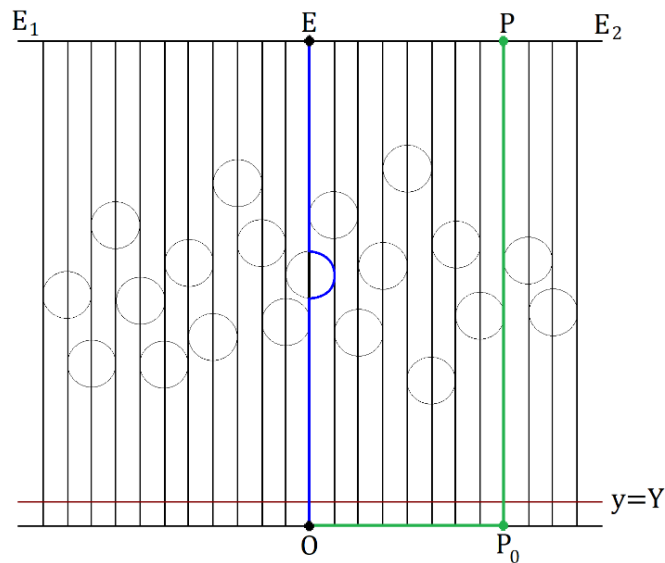


Рис. 4. Сравнение синего пути УПО, построенного по лемме, и зеленого, построенного по замечанию 1. Вертикальными линиями отмечены проекции шаров вокруг обнаружителей на отрезок финиша

Заключение

В статье разработан метод расстановки обнаружителей на плоскости с целью противодействия скрытному перемещению УПО для охраны периметра, математически представляющего собой отрезок. Задача расстановки решается градиентным методом, в котором оптимизируемыми параметрами являются координаты обнаружителей. При этом для каждой заданной их конфигурации, рассматриваемой в ходе решения, численно решается задача классификации путей УПО в конфликтной среде на основе принципа максимума Понтрягина. Авторами реализован программный комплекс на языке С, в работе приводятся результаты численного моделирования.

В качестве примера рассмотрена расстановка пяти обнаружителей при равномерно распространяющихся бесконечных круговых полях обнаружения. Интерес представляет теоретическое обоснование полученного результата – того, что при оптимальной конфигурации обнаружителей функционал I на лучших для УПО $N+1$ локально оптимальных путях УПО совпадает.

Ясно что функционал расстановки \hat{J} ограничен нулем снизу. В работе теоретически показана также ограниченность минимизируемого УПО функционала сверху для любой расстановки обнаружителей, приведены способы построения траекторий, для которых ограничение будет выполнено независимо от расположения обнаружителей.

Развитием задачи является рассмотрение более сложных диаграмм направленности обнаружителей. Возможно также рассмотрение задачи в трехмерной постановке и решение другими методами численной оптимизации. Постановка может быть обобщена на случай прорыва периметра группой объектов [19]–[20].

Список источников

1. *Gafurov S., Klochkov E.* Autonomous Unmanned Underwater Vehicles Development Tendencies // *Procedia Engineering*. 2015. Vol. 106. P. 141–148. 10.1016/j.pro-eng.2015.06.017 (дата обращения: 23.04.2025). DOI: 10.1016/j.pro-eng.2015.06.017 EDN: VAJWBB
2. *Mondal K., Banerjee T., Panja A.* Autonomous Underwater Vehicles: Recent Developments and Future Prospects // *International Journal for Research in Applied Science and Engineering Technology*. 2019. Vol. 7 (XI). P. 215–222. 10.22214/ijraset.2019.11036 (дата обращения: 23.04.2025). DOI: 10.22214/ijraset.2019.11036
3. *Wynn R. B., Huvenne V. A. I., Le Bas T. P., Murton B. J., et al.* Autonomous Underwater Vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience // *Marine Geology*. 2014. Vol. 352. P. 451–468. 10.1016/j.margeo.2014.03.012 (дата обращения: 23.04.2025). DOI: 10.1016/j.margeo.2014.03.012
4. *Railey K., Dibiaso D., Schmidt H.* Passive acoustic detection and tracking of an unmanned underwater vehicle from motor noise // *J. Acoust. Soc. Am.* 2021. Vol. 149. A34–A35. 10.1121/10.0004444 (дата обращения: 23.04.2025). DOI: 10.1121/10.0004444
5. *Abbasov I. B.* Hydroacoustic Ocean Exploration: Theories and Experimental Application. Beverly, MA, USA: Wiley & Sons and Scrivener Publishing. 2016. 192 p. 10.1002/9781119323723 (дата обращения: 23.04.2025). DOI: 10.1002/9781119323723
6. *Abraham D. A.* Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation. Springer Nature: Cham, Switzerland. 2019. 834 p. 10.1007/978-3-319-92983-5 (дата обращения: 23.04.2025). DOI: 10.1007/978-3-319-92983-5
7. *Ivanov P. V., Kolev N. Zh.* Acoustical system for synthetic imaging of an underwater object in a hydroacoustical test tank. National conference "Acoustics 2014". Vol. 16. P. 146–150.
8. *Buszman K., Gloza M.* Detection of Floating Objects Based on Hydroacoustic and Hydrodynamic Pressure Measurements in the Coastal Zone // *Polish Maritime Research, Sciendo*. 2020. Vol. 27 (2). P. 168–175. 10.2478/pomr-2020-0038 (дата обращения: 23.04.2025). DOI: 10.2478/pomr-2020-0038 EDN: XEDKIE
9. *Shirokov V., Bazhenova A., Milich V.* A set of instruments for tracking moving objects in the underwater environment // *AIP Conference Proceedings*. 2023. 2605 (1), 020026. 8 p. 10.1063/5.0111555 (дата обращения: 23.04.2025). DOI: 10.1063/5.0111555 EDN: PAYCHY

10. Galyaev A., Samokhin A., Samokhina M. Application of the Gradient Projection Method to the Problem of Sensors Arrangement for Counteraction to the Evasive Object // 28th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS). IEEE. 2021. 3 p. URL: 10.23919/icins43216.2021.9470857 (дата обращения: 23.04.2025). DOI: 10.23919/icins43216.2021.9470857 EDN: IAIIMC
11. Галяев А. А., Самохин А. С., Самохина М. А. Оптимизация расстановки обнаружителей градиентным методом // Проблемы управления безопасностью сложных систем: материалы XXVIII междунар. конф. 16 декабря 2020 г. М.: ИПУ РАН. 2020. С. 359–363. URL: 10.25728/iccss.2020.90.62.064 (дата обращения: 23.04.2025). DOI: 10.25728/iccss.2020.90.62.064(EDN: MKWRWP
12. Galyaev A. A., Samokhin A. S., Samokhina M. A. On problem of optimal observers' placement on plane // Journal of Physics: Conference Series. 2021. Vol. 1864. № 1. P. 1–7 URL: (дата обращения: 23.04.2025). DOI: 10.1088/1742-6596/1864/1/012075 EDN: FCIQWC
13. Dogan A., Zengin U. Unmanned Aerial Vehicle Dynamic-Target Pursuit by Using Probabilistic Threat Exposure Map // Journal of Guidance, Control and Dynamics. 2006. Vol. 29, № 4. P. 723–732. DOI: 10.2514/1.18386
14. Абрамянц Т. Г., Галяев А. А., Маслов Е. П., Рудько И. М., Яхно В. П. Уклонение подвижного объекта в конфликтной среде от обнаружения системой разнородных наблюдателей // Проблемы управления. 2015. № 2. С. 31–37. EDN: TPWZQH
15. Сысоев Л. П. Критерий вероятности обнаружения на траектории в задаче управления движением объекта в конфликтной среде // Проблемы управления. 2010. № 6. С. 65–72. EDN: MWLGRL
16. Самохина М. А., Галяев А. А. Построение карты локально оптимальных путей управляемого подвижного объекта в конфликтной среде при переходе из точки в точку // Проблемы управления. 2024. № 1. С. 90–102. URL: 10.25728/PU.2024.1.8 (дата обращения: 23.04.2025). DOI: 10.25728/PU.2024.1.8 EDN: KAJQCZ
17. Свидетельство о государственной регистрации программы для ЭВМ № 2020611811 Российская Федерация. Решение задачи Коши многомерным методом Рунге-Кутты, основанным на расчётных формулах Дормана-Принса 8(7), с автоматическим выбором шага: № 2020610774: заявл. 30.01.2020: опубл. 11.02.2020 / А. С. Самохин, М. А. Самохина.
18. Васильев Ф. П. Методы оптимизации. М.: Факториал Пресс, 2002. 824 с.
19. Корепанов В. О., Новиков Д. А., Задача о диффузной бомбе // Проблемы управления. 2011. Вып. 5. С. 6–73. URL: 10.1134/S000511791305010X (дата обращения: 23.04.2025). DOI: 10.1134/S000511791305010X EDN: OEESCPZ
20. Корепанов В. О., Новиков Д. А. Модели стратегического поведения в задаче о диффузной бомбе // Проблемы управления. 2015. № 2. С. 38–44. EDN: TPWZQR

References

1. Gafurov, S. and Klochkov, E. (2015), "Autonomous Unmanned Underwater Vehicles Development Tendencies", *Procedia Engineering*, vol. 106, pp. 141–148.

2. Mondal, K., Banerjee, T. and Panja, A. (2019), "Autonomous Underwater Vehicles: Recent Developments and Future Prospects", *International Journal for Research in Applied Science and Engineering Technology*, vol. 7 (XI), pp. 215–222.
3. Wynn, R. B., Huvenne, V. A. I., Le Bas, T. P., Murton, B. J., et al. (2014), "Autonomous Underwater Vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience", *Marine Geology*, vol. 352, pp. 451–468.
4. Railey, K., Dibiaso, D. and Schmidt, H. (2021), "Passive acoustic detection and tracking of an unmanned underwater vehicle from motornoise", *J. Acoust. Soc. Am.* 2021, vol. 149. A34–A35.
5. Abbasov, I. B. (2016), *Hydroacoustic Ocean Exploration: Theories and Experimental Application*. Wiley & Sons and Scrivener Publishing, Beverly, MA, USA.
6. Abraham, D. A. (2019), *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation*. Springer Nature: Cham, Switzerland.
7. Ivanov, P. V. and Kolev, N. Zh. (2014), "Acoustical system for synthetic imaging of an underwater object in a hydroacoustical test tank", *National conference "Acoustics 2014"*, vol. 16, pp. 146–150.
8. Buszman, K. and Gloza, M. (2020), "Detection of Floating Objects Based on Hydroacoustic and Hydrodynamic Pressure Measurements in the Coastal Zone", *Polish Maritime Research*, Sciendo, vol. 27 (2), pp. 168–175.
9. Shirokov, V., Bazhenova, A. and Milich, V. (2023), "A set of instruments for tracking moving objects in the underwater environment", *AIP Conference Proceedings*, vol. 2605, no 1:020026, 8 p.
10. Galyaev, A., Samokhin, A. and Samokhina, M. (2021), "Application of the Gradient Projection Method to the Problem of Sensors Arrangement for Counteraction to the Evasive Object", *28th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*, IEEE, 3 p.
11. Galyaev, A.A., Samokhin, A. S. and Samokhina, M. A. (2020), "Optimization of detector placement using the gradient method", *Problemy upravleniya bezopasnost'yu slozhnykh sistem: materialy XXVIII mezhdunarodnoj konferencii* [Problems of Complex Systems Safety Management: Proceedings of the XXVIII International Conference], Moscow, Russia, December 16, 2020, pp. 359–363.
12. Galyaev, A. A., Samokhin, A. S. and Samokhina, M. A. (2021), "On problem of optimal observers' placement on plane", *Journal of Physics: Conference Series*, vol. 1864, no. 1, pp. 1–7.
13. Dogan, A. and Zengin, U. (2006), "Unmanned Aerial Vehicle Dynamic-Target Pursuit by Using Probabilistic Threat Exposure Map", *Journal of Guidance, Control and Dynamics*, vol. 29, no. 4, pp. 723–732.
14. Abramyan, T. G., Galyaev, A. A., Maslov, E. P., Rudko, I. M. and Yakhno, V.P. (2015), "Mobile Object Evasion from Detection by a System of Heterogeneous Observers in Threat Environment", *Control Sciences*, no. 1, pp. 31–37.
15. Sysoev, L. P. (2010), "The Criterion of Probability Detecting on the Trajectory in the Problem of Movement Control in Threat Environment", *Control Sciences*, no. 6, pp. 64–70.
16. Samokhina, M. A. and Galyaev, A. A. (2024), "Constructing a Map of Locally Optimal Paths for a Controlled Moving Object in a Threat Environment", *Control Sciences*, no. 1, pp. 75–85.
17. Samokhin, A. S. and Samokhina, M. A., RUDN University (2021), *Reshenie zadachi Koshi mnogomernym metodom Runge-Kutty, osnovannym na raschyotnykh formulakh Dormana-Prinsa 8(7), s avtomaticheskim vyborom shaga* [Solving the Cauchy problem by the

multivariate Runge-Kutta method based on the Dorman-Prince 8(7) formulas with automatic step selection], Moscow, Russian Federation, Certificate of state registration of computer program no. 2020611811.

18. Vasiliev, F. P. (2002), *Metody optimizacii* [Optimization methods], Factorial Press, Moscow, Russia.
19. Korepanov, V. O. and Novikov, D. A. (2013), "The diffuse bomb problem", *Automation and Remote Control*, vol. 74, no. 5, pp. 863–874.
20. Korepanov, V. O. and Novikov, D. A. (2015), "Models of Strategic Behavior in the Diffuse Bomb Problem", *Control Sciences*, no. 2, pp. 38–44.

Информация об авторах:

М. А. Самохина – научный сотрудник 38 лаборатории «Управление по неполным данным» Института проблем управления имени В. А. Трапезникова Российской академии наук (117997, Россия, г. Москва, ул. Профсоюзная, 65), WoS Researcher ID: AAL-3864-2020, Scopus Author ID: 57201875555, РИНЦ Author ID: 1029494, Istina Researcher ID: 5803904;

А. С. Самохин – кандидат физико-математических наук, член Академии навигации и управления движением, старший научный сотрудник 38 лаборатории «Управление по неполным данным» Института проблем управления имени В. А. Трапезникова Российской академии наук (117997, Россия, г. Москва, ул. Профсоюзная, 65), WoS Researcher ID: N-1400-2019, Scopus Author ID: 56150017500, РИНЦ Author ID: 1029494. Istina Researcher ID: 3497959.

Information about the authors:

М. А. Samokhina – researcher, 38 laboratory «Control based on incomplete data», V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences (65 Profsoyuznaya St., Moscow, Russia, 117997), WoS Researcher ID: AAL-3864-2020, Scopus Author ID: 57201875555, RSCI Author ID: 1029494. Istina Researcher ID: 5803904;

A. S. Samokhin – candidate of physical and mathematical sciences, member of Academy of Navigation and Motion Control, senior researcher, 38 laboratory «Control based on incomplete data», V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences (65 Profsoyuznaya St., Moscow, Russia, 117997), WoS Researcher ID: N-1400-2019, Scopus Author ID: 56150017500, RSCI Author ID: 1029494. Istina Researcher ID: 3497959.

Научный периодический журнал "Вестник Пермского университета. Математика. Механика. Информатика" зарегистрирован в 2007 г. как самостоятельное издание, являющееся продолжением серии журнала "Вестник Пермского университета", издаваемого с 1994 г. В каждом номере научного журнала автору разрешается опубликование в любом разделе не более одной статьи, в том числе в соавторстве.

Статьи, оформленные в соответствии с нижеизложенными правилами, должны быть отправлены через **личный кабинет автора на сайте журнала**. Статьи отправляются в двух форматах: текстовый документ и pdf-файл без данных об авторе (для организации слепого рецензирования). Вместе с рукописью отправляются контактные данные авторов в свободной форме в текстовом документе (ФИО, должность, место работы с указанием полного адреса, e-mail, номер телефона), сопроводительное письмо и экспертное заключение о возможности открытого опубликования. Для публикации статьи необходимо заполнить и подписать согласие на обработку персональных данных и лицензионный договор.

Аннотация статьи (на русском и английском языках) должна содержать основные краткие сведения: описание проблемы, включая цели и задачи исследования; методы исследования с указанием новизны, если имеется; полученные результаты и выводы. Объем аннотации примерно 250 слов

Текст статьи должен содержать разделы: Введение, Постановка задачи, Методы исследования, Результаты, Обсуждения, Заключение, **Благодарности** для русскоязычной статьи; Introduction, Materials and Methods, Results, Discussion, Conclusion, Acknowledgements для англоязычной статьи. Объем рукописи – не менее 6 страниц машинописного текста.

Оформление. Текст статьи необходимо оформить в редакторе **Word**. Формат листа, используемый для написания статьи – А4. Размеры верхнего и нижнего полей – 2.6 см, правого и левого – 2.5 см. Расстояние до верхнего и нижнего колонтитулов 1.25 см. Шрифтом Times New Roman Cyr, размер – 12 пт, межстрочный интервал – множитель 1.1. Подписи к рисункам и таблицам набираются курсивом, в конце заголовков и ячеек таблицы точка не ставится. Формулы набираются в редакторе формул.

Структура статьи

Вид статьи (научная, обзорная или др.).

УДК

DOI:

EDN

Название раздела (математика, механика, компьютерные науки и информатика).

Название статьи на русском языке (шрифт Times New Roman – 16 пт, полужирный).

Ф.И.О. автора (-ов) полностью (Times New Roman – 12 пт, полужирный).

Место работы (с указанием подразделения), город и страна (шрифт Times New Roman – 11 пт), **e-mail**.

Аннотация на русском языке (шрифт Times New Roman – 11 пт).

Ключевые слова на русском языке: 5–7 слов или словосочетаний (шрифт Times New Roman – 11 пт, курсив).

Информация для цитирования (шрифт Times New Roman – 11 пт).

Благодарности (шрифт Times New Roman – 11 пт).

Даты поступления, одобрения и принятия рукописи (шрифт Times New Roman – 11 пт).

Название статьи на английском языке (шрифт Times New Roman – 16 пт, полужирный).

На английском языке Ф.И.О. автора (-ов) полностью (шрифт Times New Roman – 12 пт, полужирный).

Место работы (с указанием подразделения), город и страна на английском языке (шрифт Times New Roman – 11 пт), **e-mail**.

Аннотация на английском языке (шрифт Times New Roman – 11 пт).

Ключевые слова на английском языке: 5–7 слов или словосочетаний (шрифт Times New Roman – 11 пт, курсив).

Информация для цитирования на английском языке (шрифт Times New Roman – 11 пт).

Благодарности на английском языке (шрифт Times New Roman – 11 пт).

Даты поступления, одобрения и принятия рукописи на английском языке (шрифт Times New Roman – 11 пт).

Основной текст статьи набирается в ширину страницы, абзацный отступ основного текста – 1.0 см (задается автоматически, не пробелами). Другие отступы не допускаются. Выравнивание текста по ширине.

Список источников оформляется в едином формате, в соответствии с ГОСТ Р 7.0.100–2018, без использования тире. Если в списке литературы используется источник из электронного ресурса, необходимо указать дату обращения. Ссылка на литературный источник указывается в тексте статьи в квадратных скобках после упоминания цитаты. Список использованных литературных источников необходимо представить на русском и английском языках (**References**). **Список литературы на латинице**, называемый References, готовится отдельно от Списка литературы и размещается сразу за указанным списком. Ссылки на зарубежные источники необходимо повторять в Списке литературы и References. В References недопустимо использовать российские ГОСТы. Необходимо следовать требованиям международных стандартов. Инструкция по оформлению списка литературы в латинице – References, стандарт "Harvard" находится по ссылке http://www.psu.ru/files/docs/ob-universitete/smi/nauchnyj-jurnal/metodicheskie_materialy/standart_Harvard.doc.

ПОРЯДОК РЕЦЕНЗИРОВАНИЯ И ПУБЛИКАЦИИ СТАТЕЙ

Все рукописи проходят двойное слепое рецензирование двумя специалистами в научной области исследования. При возникновении спорной ситуации назначается третий рецензент. Окончательное решение о публикации рукописи принимает главный редактор журнала.

Датой поступления статьи считается день ее отправки через личный кабинет на сайте журнала или по электронной почте журнала (главного редактора). Рукописи рассматриваются в порядке их поступления в течение одного месяца в зависимости от сложности ситуации и объема работы. Редакция оставляет за собой право без согласования с автором проводить литературную правку текстов статей, не изменяющую их основного смысла.

Одобрённые статьи публикуются бесплатно. Полнотекстовая версия каждой рукописи выставляется на сайте журнала и в системе РИНЦ (e-library).

Scientific periodical journal "Bulletin of Perm University. Mathematics. Mechanics. Computer Science" is published since 1994 in the series of the journal "Bulletin of Perm University". Then it was registered as an independent publication in 2007.

An author is allowed to publish no more than one article (also as co-author) in any section for each issue of the journal.

Article's structure should be accordant with the rules below and should be sent through the **author's personal account on the journal's website**. Articles should be sent in two formats: a text document and a pdf-file without information about the author(s) (for review). Authors contact data (name, position and place of work with address, e-mail, phone number in a text document), a cover letter and an expert letter on the open publication possibility should be sent with the manuscript. Author(s) must make and sign consent to the personal data processing and an agreement.

Abstract (in Russian and English) should contain main information of research: a problem including the goals and objectives of the study; research methods with novelty indication, if any; significant results and findings. The abstract size is approximately 250 words.

Text should contain sections: Introduction, Materials and Methods, Results, Discussion, Conclusion, Acknowledgments. The manuscript size is at least 6 typewritten pages.

An article text must be formatted in the editor Word. The sheet format is A4. The top and bottom margins are 2.6 cm, the right and left margins are 2.5 cm. The distances to the header and footer are 1.25 cm. Times New Roman Cyr font, size – 12 pt, single spacing. Legends for figures and tables are typed in italics style. Formulas are typed in the Equation editor.

The structure of the article

Article's type

UDC

DOI

EDN

TITLE in Russian (Times New Roman font – 16 pt, bold).

Name: author (s) name in full (Times New Roman font – 12 pt).

Place of work (with departments), city, country (Times New Roman font – 11 pt), **e-mail**

Abstract in Russian (Times New Roman font – 11 pt).

Keywords in Russian: 5–7 words or phrases (Times New Roman font – 11 pt, italic).

Information for citation in English (Times New Roman font – 11 pt, italic).

Acknowledgments in English (Times New Roman font – 11 pt, italic).

Submitted, approved, accepted dates of receipt in English (Times New Roman font – 11 pt, italic).

TITLE in English (Times New Roman font – 16 pt, bold).

In English author (s) name in full, place of work (with departments), city, country (Times New Roman font – 11 pt), **e-mail**.

Abstract in English (Times New Roman font – 11 pt).

Keywords in English: 5–7 words or phrases (Times New Roman font – 11 pt).

Information for citation in English (Times New Roman font – 11 pt).

Acknowledgments in English (Times New Roman font – 11 pt).

Submitted, approved, accepted dates of receipt in English (Times New Roman font – 11 pt).

The article text is typed the paragraph indent is 1.0 cm (set automatically, not spaces). No other indents are allowed. Justify the text in width (Times New Roman font – 16 pt)..

References are formed in a single format with using GOST R 7.0.100–2018 without a dash. If a source from an electronic resource is used, the author should indicate the request date. To indicate a source in text author uses square brackets after the quotation mention. References must be submitted **in Russian and English**.

The References list in Latin, called References, is prepared separately from the References List in Russian and placed below. Foreign sources references should be repeated in the References Lists in Russian and English. It is inadmissible to use Russian GOSTs in References. It is necessary to follow the requirements of international standards. Instructions on the References in Latin script design – References, Harvard standard can be found at http://www.psu.ru/files/docs/ob-universitete/smi/nauchnyj-zhurnal/metodicheskie_materialy/standart_Harvard.doc

REVIEW AND PUBLICATION OF ARTICLES

All manuscripts are double-blind peer-reviewed by two research scientists. If a disputable situation arises, a third reviewer is appointed. The final decision on the manuscript publication is made by the editor-in-chief of the journal.

The receipt date of the article is the day of its submission through the personal account on the journal's website or by e-mail of the journal (or editor-in-chief). Manuscripts are reviewed in the order they are received within one month, review period may be changed depending on a situation complexity and a works amount. The journal editors can to edit the article text without basic meaning change.

Approved articles are published free. The full-text version of each manuscript is posted on the journal's website and in the RSCI system (e-library).

Научное издание

ВЕСТНИК ПЕРМСКОГО УНИВЕРСИТЕТА
Математика. Механика. Информатика

Выпуск 2(69)

2025

Адрес учредителя и издателя

614068, Пермский край, г. Пермь, ул. Букирева, д. 15, ПГНИУ

Адрес редакции

614068, Пермский край, г. Пермь, ул. Букирева, д. 15, ПГНИУ

Физико-математический институт

e-mail: vestnik_mmi@psu.ru

Редактор Е. В. Шумилова

Корректор В. Е. Пирожкова

Компьютерная верстка: Г. И. Кушнина

Подписано в печать 14.07.2025. Выход в свет 15.07.2025

Формат 60×84/8. Усл. печ. л. 16,16. Тираж 500 экз. Заказ 83

Пермский государственный национальный исследовательский университет

Управление издательской деятельности

614068, г. Пермь, ул. Букирева, 15. Тел. (342) 239-66-36

Отпечатано в типографии ПГНИУ.

Тел. (342) 239-65-47

Подписной индекс журнала

«Вестник Пермского университета. Математика. Механика. Информатика»

в онлайн-каталоге «Урал-Пресс» –015010

https://www.ural-press.ru/catalog/97266/8708250/?sphrase_id=396133

Распространяется бесплатно и по подписке