

ISSN 2713-3192
DOI 10.15622/ia.2025.24.4
<http://ia.spcras.ru>

TOM 24 № 4

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

INFORMATICS AND AUTOMATION



СПб ФИЦ РАН

Санкт-Петербург
2025



INFORMATICS AND AUTOMATION

Volume 24 № 4, 2025

Scientific and educational journal primarily specialized in computer science, automation, robotics, applied mathematics, interdisciplinary research

Founded in 2002

Founder and Publisher

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)

Editor-in-Chief

A. L. Ronzhin, Prof., Dr. Sci., St. Petersburg, Russia

Editorial Council

A. A. Ashimov	Prof., Dr. Sci., Academician of the National Academy of Sciences of the Republic of Kazakhstan, Almaty, Kazakhstan
I. A. Kalyaev	Prof., Dr. Sci., Academician of RAS, Taganrog, Russia
Yu. A. Merkuryev	Prof., Dr. Sci., Academician of the Latvian Academy of Sciences, Riga, Latvia
A. I. Rudskoi	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
V. Sgurev	Prof., Dr. Sci., Academician of the Bulgarian Academy of Sciences, Sofia, Bulgaria
B. Ya. Sovetov	Prof., Dr. Sci., Academician of RAE, St. Petersburg, Russia
V. A. Soyfer	Prof., Dr. Sci., Academician of RAS, Samara, Russia

Editorial Board

E. Azarov	Prof., Dr. Sci., Minsk, Belarus
O. Yu. Gusikhin	Ph. D., Dearborn, USA
V. Delic	Prof., Dr. Sci., Novi Sad, Serbia
A. Dolgui	Prof., Dr. Sci., St. Etienne, France
M. N. Favorskaya	Prof., Dr. Sci., Krasnoyarsk, Russia
M. Zelezny	Assoc. Prof., Ph.D., Plzen, Czech Republic
H. Kaya	Assoc. Prof., Ph.D., Utrecht, Netherlands
A. A. Karpov	Assoc. Prof., Dr. Sci., St. Petersburg, Russia
S. V. Kuleshov	Dr. Sci., St. Petersburg, Russia
A. D. Khomonenko	Prof., Dr. Sci., St. Petersburg, Russia
D. A. Ivanov	Prof., Dr. Habil., Berlin, Germany
K. P. Markov	Assoc. Prof., Ph.D., Aizu, Japan
R. V. Meshcheryakov	Prof., Dr. Sci., Moscow, Russia
N. A. Moldovian	Prof., Dr. Sci., St. Petersburg, Russia
V. V. Nikulin	Prof., Ph.D., New York, United States
V. Yu. Osipov	Prof., Dr. Sci., Deputy Editor-in-Chief, St. Petersburg, Russia
V. K. Pshikhopov	Prof., Dr. Sci., Taganrog, Russia
H. Samani	Assoc. Prof., Ph.D., Plymouth, UK
J. Savage	Assoc. Prof., Ph.D., Mexico City, Mexico
M. Secujski	Assoc. Prof., Ph.D., Novi Sad, Serbia
A. V. Smirnov	Prof., Dr. Sci., St. Petersburg, Russia
B. V. Sokolov	Prof., Dr. Sci., St. Petersburg, Russia
L. V. Utkin	Prof., Dr. Sci., St. Petersburg, Russia
L. B. Sheremetov	Assoc. Prof., Dr. Sci., Mexico, Mexico

Editor: A. S. Viktorova

Interpreter: Ya. N. Berezina

Art editor: N. A. Dormidontova

Editorial office address

SPC RAS, 39 litera A, 14-th line V.O., St. Petersburg, 199178, Russia

e-mail: ia@sprcras.ru, web: <http://ia.sprcras.ru>

The journal is indexed in Scopus

The journal is published under the scientific-methodological supervision of Department for Nanotechnologies and Information Technologies of the Russian Academy of Sciences

© St. Petersburg Federal Research Center of the Russian Academy of Sciences, 2025

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

Том 24 № 4, 2025

Научный, научно-образовательный журнал с базовой специализацией
в области информатики, автоматизации, робототехники, прикладной математики
и междисциплинарных исследований.

Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
(СПб ФИЦ РАН)

Главный редактор

А. Л. Ронжин, д-р техн. наук, проф., Санкт-Петербург, РФ

Редакционный совет

А. А. Ашимов	академик Национальной академии наук Республики Казахстан, д-р техн. наук, проф., Алматы, Казахстан
И. А. Каляев	академик РАН, д-р техн. наук, проф., Таганрог, РФ
Ю. А. Меркурьев	академик Латвийской академии наук, д-р, проф., Рига, Латвия
А. И. Рудской	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
В. Сгурев	академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария
Б. Я. Советов	академик РАО, д-р техн. наук, проф., Санкт-Петербург, РФ
В. А. Соифер	академик РАН, д-р техн. наук, проф., Самара, РФ

Редакционная коллегия

И. С. Азаров	д-р техн. наук, проф., Минск, Беларусь
О. Ю. Гусихин	д-р наук, Диаборн, США
В. Делич	д-р техн. наук, проф., Нови-Сад, Сербия
А. Б. Долгий	д-р наук, проф. Сент-Этьен, Франция
М. Железны	д-р наук, доцент, Пльзень, Чешская республика
Д. А. Иванов	д-р экон. наук, проф., Берлин, Германия
Х. Кайя	д-р наук, доцент, Утрехт, Нидерланды
А. А. Карпов	д-р техн. наук, доцент, Санкт-Петербург, РФ
С. В. Кулешов	д-р техн. наук, Санкт-Петербург, РФ
К. П. Марков	д-р наук, доцент, Аизу, Япония
Р. В. Мещеряков	д-р техн. наук, проф., Москва, РФ
Н. А. Молдовян	д-р техн. наук, проф., Санкт-Петербург, РФ
В. В. Никулин	д-р наук, проф., Нью-Йорк, США
В. Ю. Осипов	д-р техн. наук, проф., зам. главного редактора, Санкт-Петербург, РФ
В. Х. Пшихопов	д-р техн. наук, проф., Таганрог, РФ
Х. К. Саваж	д-р техн. наук, доцент, Мехико, Мексика
Х. Самани	д-р наук, доцент, Плимут, Соединенное Королевство
М. Сечуиски	д-р техн. наук, доцент, Нови-Сад, Сербия
А. В. Смирнов	д-р техн. наук, проф., Санкт-Петербург, РФ
Б. В. Соколов	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. В. Уткин	д-р техн. наук, проф., Санкт-Петербург, РФ
М. Н. Фаворская	д-р техн. наук, проф., Красноярск, РФ
А. Д. Хомоненко	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. Б. Шереметов	д-р техн. наук, Мехико, Мексика

Выпускающий редактор: А. С. Викторова

Переводчик: Я. Н. Березина

Художественный редактор: Н. А. Дормидонтова

Адрес редакции

14-я линия В.О., д. 39, лит. А, г. Санкт-Петербург, 199178, Россия

e-mail: ia@spcras.ru, сайт: <http://ia.spcras.ru>

Журнал индексируется в международной базе данных Scopus

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук»

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий и информационных технологий Российской академии наук

© Федеральное государственное бюджетное учреждение науки

«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», 2025
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных в составе печатного периодического издания - журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ» статей по текущим экономическим, политическим, социальным и религиозным вопросам с обязательным указанием имени автора статьи и печатного периодического издания журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»

CONTENTS

Robotics, Automation and Control Systems

A. Zaitsev, K. Yakovlev
TIME RANGE CONSTRAINTS FOR MOTION PLANNING FOR MANIPULATORS 1007

I. Arkhipov, V. Erofeeva, O. Granichin, V. Kiselev, A. Chernov
ACCELERATED LOCAL VOTING PROTOCOL FOR SINGLE REMOTE CONTROLLER
ROBOT SWARM SYSTEM 1029

V. Voronoy, D. Verhovod
HUMAN GAIT MATHEMATICAL MODEL DEVELOPMENT AND CONTROL SYSTEM
SYNTHESIS FOR ACTUATING MECHANISMS OF THE MECHATRONIC
REHABILITATION COMPLEX 1059

Artificial Intelligence, Knowledge and Data Engineering

A. Andreev, M. Kuchma, S. Malkovsky, A. Filei
RAINCAST: A HYBRID PRECIPITATION NOWCASTING ALGORITHM USING
THE HIMAWARI-8/9 SATELLITE MEASUREMENTS 1085

A. Losev, I. Popov, A. Reznikova
IMPROVING THE REPRESENTATIVENESS OF THE TRAINING DATASET
BY MEANS OF SPATIAL BALANCING 1114

A. Erpalov, V. Sinitsin, A. Shestakov
EWT-CGAN DATA AUGMENTATION FOR MEASUREMENT SYSTEMS 1157

Digital Information Telecommunication Technologies

S. Dvornikov, A. Maslova, D. Vasilieva
PROBABILISTIC ESTIMATION OF NOISE IMMUNITY OF COMPOSITE VIDEO
TRANSMISSION RADIO LINES IN CONDITIONS OF MUTUAL INTERFERENCE 1182

A. Maevsky, V. Ryzhov, T. Fedorova, I. Pechaiko, N. Burov
STOCHASTIC DYNAMIC MODEL OF AN UNDERWATER WIRELESS SENSOR
NETWORK BASED ON A MODIFIED LOUVAIN CLUSTERING ALGORITHM 1206

СОДЕРЖАНИЕ

Робототехника, автоматизация и системы управления

А.С. Зайцев, К.С. Яковлев
ИНТЕРВАЛЬНЫЕ ВРЕМЕННЫЕ ОГРАНИЧЕНИЯ ДЛЯ ПЛАНИРОВАНИЯ
СОГЛАСОВАННЫХ ДВИЖЕНИЙ МНОГОЗВЕННЫХ МАНИПУЛЯТОРОВ 1007

И.С. Архипов, В.А. Ерофеева, О.Н. Граничин, В.А. Киселев, А.О. Чернов
УСКОРЕННЫЙ ПРОТОКОЛ ЛОКАЛЬНОГО ГОЛОСОВАНИЯ ДЛЯ ГРУППЫ
РОБОТОВ С ОДНИМ ПУЛЬТОМ ДИСТАНЦИОННОГО УПРАВЛЕНИЯ 1029

В.В. Вороной, Д.П. Верховод
РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПОХОДКИ ЧЕЛОВЕКА И СИНТЕЗ
СИСТЕМЫ УПРАВЛЕНИЯ ИСПОЛНИТЕЛЬНЫМИ МЕХАНИЗМАМИ
МЕХАТРОННОГО РЕАБИЛИТАЦИОННОГО КОМПЛЕКСА 1059

Искусственный интеллект, инженерия данных и знаний

А.И. Андреев, М.О. Кучма, С.И. Мальковский, А.А. Филей
RAINCAST: ГИБРИДНЫЙ АЛГОРИТМ НАУКАСТИНГА ИНТЕНСИВНОСТИ
ОСАДКОВ ПО ДАННЫМ СПУТНИКОВЫХ НАБЛЮДЕНИЙ КА НИМАWARI-8/9 1085

А.Г. Лосев, И.Е. Попов, А.С. Резникова
ПОВЫШЕНИЕ РЕПРЕЗЕНТАТИВНОСТИ ОБУЧАЮЩЕГО НАБОРА ДАННЫХ
ЗА СЧЕТ ПРОСТРАНСТВЕННОЙ БАЛАНСИРОВКИ 1114

А.В. Ерпалов, В.В. Синицин, А.Л. Шестаков
EWT-CGAN АУГМЕНТАЦИЯ ДАННЫХ ИЗМЕРИТЕЛЬНЫХ СИСТЕМ 1157

Цифровые информационно-телекоммуникационные технологии

С.В. Дворников, А.А. Маслова, Д.В. Васильева
ВЕРОЯТНОСТНАЯ ОЦЕНКА ПОМЕХОЗАЩИЩЕННОСТИ СОСТАВНЫХ
РАДИОЛИНИЙ ПЕРЕДАЧИ ВИДЕО В УСЛОВИЯХ ВЗАИМНЫХ ПОМЕХ 1182

А.М. Маевский, В.А. Рыжов, Т.А. Федорова, И.А. Печайко, Н.М. Буров
СТОХАСТИЧЕСКАЯ ДИНАМИЧЕСКАЯ МОДЕЛЬ ПОДВОДНОЙ БЕСПРОВОДНОЙ
СЕНСОРНОЙ СЕТИ, ОСНОВАННАЯ НА МОДИФИЦИРОВАННОМ ЛУВЕНСКОМ
АЛГОРИТМЕ КЛАСТЕРИЗАЦИИ 1206

А.С. ЗАЙЦЕВ, К.С. ЯКОВЛЕВ
**ИНТЕРВАЛЬНЫЕ ВРЕМЕННЫЕ ОГРАНИЧЕНИЯ ДЛЯ
ПЛАНИРОВАНИЯ СОГЛАСОВАННЫХ ДВИЖЕНИЙ
МНОГОЗВЕННЫХ МАНИПУЛЯТОРОВ**

Зайцев А.С., Яковлев К.С. Интервальные временные ограничения для планирования согласованных движений многозвенных манипуляторов.

Аннотация. В работе рассматривается задача планирования согласованных движений для системы многозвенных роботов-манипуляторов. Одним из перспективных современных подходов к её решению является конфликтно-ориентированное планирование, позволяющее отказаться от построения комбинированного пространства поиска высокой размерности за счет перехода к последовательному решению совокупности задач более низкой размерности. Этот переход осуществляется путем задания пространственно-временных ограничений при наличии конфликта в индивидуальных планах манипуляторов и последующего перепланирования с учетом наложенных ограничений. К сожалению, известные к настоящему моменту способы наложения ограничений оперируют отдельными моментами времени, что снижает их эффективность на практике. В данной работе представляется новый алгоритм конфликтно-ориентированного планирования, основанный не на точечных, а на интервальных временных ограничениях – GECBS-T. С теоретической точки зрения, предлагаемый алгоритм гарантирует ограниченную суб-оптимальность конструируемых решений, т.е. при любом наперед заданном пользователем ограничении $w > 1$, стоимость решения GECBS-T не будет превышать стоимость оптимального решения более, чем в w раз. На практике же, предлагаемый алгоритм существенно превосходит аналогичные алгоритмы по скорости решения задач планирования, что подтверждается проведенными экспериментами в робототехническом симуляторе *tuJoso*, включающими 2-4 робота-манипулятора *Kuka* с 7 степенями свободы каждый.

Ключевые слова: планирование движения, планирование траекторий, многоагентное планирование, манипулятор, конфликтно-ориентированное планирование, поиск на основе конфликтов, поиск по графу, интервальные ограничения, временные ограничения.

1. Введение. В последние десятилетия робототехника стала одной из наиболее динамично развивающихся областей науки и техники. Особое место в этой сфере занимают многозвенные манипуляторы – механические устройства, состоящие из звеньев и шарниров с рабочим органом на последнем из них.

Манипуляторы активно применяются в промышленности, складской автоматизации, сервисной и медицинской робототехнике и др. При этом в последнее время всё чаще возникает необходимость в адаптивном поведении манипуляторов, а не выполнении одной или нескольких заранее построенных программ движений и действий. Также зачастую возникают задачи предполагающие их выполнение не одним, а системой из нескольких манипуляторов. Таким образом, весьма актуальной становится задача разработки эффективных методов

планирования согласованных траекторий для системы из нескольких робототехнических манипуляторов, действующих в общем рабочем пространстве.

Задача адаптивного планирования движений для отдельного манипулятора уже является вычислительно сложной из-за, во-первых, необходимости решения задачи обратной кинематики и частого вызова достаточно ресурсоемкой процедуры проверки на коллизии и, во-вторых, из-за большого числа степеней свободы робототехнической системы. Если же роботов-манипуляторов несколько, то сложность задачи возрастает ещё больше, т.к. для многоагентного планирования характерен экспоненциальный рост пространства комбинаций действий относительного числа агентов.

Одним из способов борьбы с проклятием размерности при многоагентном планировании является конфликтно-ориентированный поиск. При таком типе поиска комбинированного пространства действий не строится, наоборот – планирование осуществляется лишь для отдельных агентов. При обнаружении конфликтов на вовлеченных агентов накладываются пространственно-временные ограничения, определенные специальным образом так, чтобы их соблюдение устраняло обнаруженный конфликт. Затем происходит индивидуальное ерхнеуровневого дерева поиска возникает планирование с учетом наложенных ограничений и процесс повторяется до тех пор, когда не будут устранены все конфликты и не будет получен набор согласованных индивидуальных планов, одновременное выполнение которых происходит без столкновений. Этот набор планов и является решением исходной задачи.

Изначально конфликтно-ориентированное планирование было предложено для мобильных роботов [1], однако в последние несколько лет появились алгоритмы, адаптирующие этот подход к системам много-звенных робототехнических манипуляторов [2]. К сожалению, подобные алгоритмы используют способы задания конфликтов, характеризующиеся привязкой к конкретному моменту времени. Из-за этого эффективность подхода снижается, т.к. устранение конфликта в момент времени t зачастую приводит к возникновению аналогичного конфликта в момент времени $t + 1$. Именно на решение этой проблемы и направлена данная работа.

Для повышения эффективности планирования согласованных траекторий для системы много-звенных манипуляторов в работе предлагается новый алгоритм поиска, GECBS-T, который использует принцип конфликтно-ориентированного планирования, но вводит в рассмотрение интервальные временные ограничения, которые нацелены

на избегание цепочек конфликтов, возникающих при стандартном подходе. Благодаря использованию дополнительной техники фокусировки поиска, предлагаемый алгоритм GECBS-T позволяет гарантировать отыскание решений, стоимость которых не превосходит стоимость оптимальных более, чем на определенную пользователем величину. Таким образом, с одной стороны алгоритм позволяет контролировать качество отыскиваемых траекторий, с другой – превосходит предшествующие алгоритмы по скорости решения задач планирования. Последнее обстоятельство подтверждается проведенными многочисленными экспериментами с использованием робототехнического симулятора *muJoCo* и реальных моделей симуляторов *Kuka* с 7 степенями свободы.

Программный код разработанного метода и системы тестирования открыт и свободно доступен¹.

2. Обзор работ. Задача планирования движений группы роботов-манипуляторов (здесь и далее именуемая как M-RAMP, от. англ. *multi-robot/agent motion planning*) может быть рассмотрена как более сложная версия задачи многоагентного поиска путей (MAPF, *multi-agent pathfinding*) [3, 4]. В классической формулировке этой задачи на плоскости с препятствиями расположены мобильные роботы, способные перемещаться по незанятой препятствиями территории. У каждого робота есть начальное и конечное положение и нужно построить для каждого робота путь между ними так, чтобы ни в какой момент времени никакие роботы не столкнулись друг с другом или с препятствиями.

Разница задач M-RAMP и MAPF заключается в том, что в первом случае степень свободы одного робота находится в диапазоне от 6 до 8, тогда как у робота на плоскости она равна 2, из-за чего планирование движений для одного робота-манипулятора намного более сложная задача, чем планирование траектории агента на плоскости. Так, если в задаче MAPF количество агентов может достигать нескольких сотен, то в задаче M-RAMP их число не более 20. Таким образом, решения задачи MAPF, основанные на том, что планирование траектории одного агента происходит быстро, нужно адаптировать к задаче M-RAMP.

Для решения задачи MAPF применяются разные методы: метод приоритизированного планирования [5–7] и метод конфликтно-ориентированного планирования [8], которые могут быть перенесены с роботов на плоскости на роботов-манипуляторов. Так как главное отличие задач в том, что в задаче MAPF путь для одного робота строится относительно быстро, а в M-RAMP достаточно медленно, то для переноса метода конфликтно-ориентированного планирования авторы

¹https://github.com/machine-solution/motion_planning_for_manipulators

алгоритмов стараются оптимизировать их за счёт уменьшения количества вызовов планирования траекторий для одного робота, что показывает алгоритм Generalized ECBS [2], являющийся на данный момент лучшим алгоритмом на основе поиска для решения данной задачи. Это основной метод решения данной задачи на основе поиска [9]. Приоритизированное планирование можно перенести и без модификаций [10], тем не менее такой подход также существенно зависит от выбранных приоритетов и на практике является менее эффективным, чем остальные.

Для задачи M-RAMP также есть и специфичный подход на основе семплирования. Такой метод представляет группу манипуляторов как единую систему и применяет к ней алгоритм RRT [11]. Также для этого алгоритма было разработано множество модификаций [12, 13]. Такой алгоритм не так сильно подвержен замедлению, связанному с ростом размерности системы, за счёт чего является довольно быстрым. Такой подход не гарантирует оптимальности или ограниченной субоптимальности решения и даже не гарантирует нахождения решения, если оно существует, что является его недостатком по сравнению с алгоритмами на основе поиска. Также минусом является и то, что такие алгоритмы на основе семплирования [14] выдают непредсказуемые случайные результаты.

Также к задаче планирования манипуляторов применяется подход на основе машинного обучения [15–17]. Эти методы позволяют проще адаптироваться к новым задачам и легко адаптируются к условиям рабочего пространства, но само обучение является трудоёмким процессом, который требует большое количество времени и данных. Также такие алгоритмы довольно тяжело интерпретировать и они не могут давать каких-либо гарантий относительно решения.

Таким образом, одним из самых перспективных направлений изучения являются методы планирования на основе поиска, в частности конфликтно-ориентированном поиске, так как этот алгоритм имеет теоретические гарантии, выдаёт предсказуемый результат, а лучший алгоритм из этого класса Generalized ECBS показывает сравнимые показатели скорости работы с алгоритмами на основе семплирования, поэтому данный алгоритм является базовым для этой работы.

3. Постановка задачи. Дано рабочее пространство, разбитое на свободное пространство и препятствия $W = W_{free} \cup W_{obs}$, $W \subset \mathbb{R}^3$. Рассмотрим n многосвязных манипуляторов $R_1, \dots, R_n$. Конфигурацией манипулятора называется набор углов поворота k сочленений манипулятора $q = (\alpha_1, \dots, \alpha_k)$, а $B_i(q)$ обозначает множество точек рабочего пространства, которое занимает i -й манипулятор в

конфигурации q . Конфигурации каждого манипулятора являются подмножеством $\mathbb{R}^k$ и образуют конфигурационные пространства $C^i = C_{free}^i \cup C_{obs}^i$, где C_{free}^i – это множество конфигураций q , таких, что $B_i(q) \cap W_{obs} = \emptyset$, а C_{obs}^i наоборот. Так, на рисунке 1 слева изображено рабочее пространство (W-space) и двухзвенный манипулятор, а справа пространство конфигураций (C-space) этого манипулятора. Агенты R_i и R_j считаются находящимися в конфликте в конфигурациях q_i и q_j , если выполняется условие $B_i(q_i) \cap B_j(q_j) \neq \emptyset$.

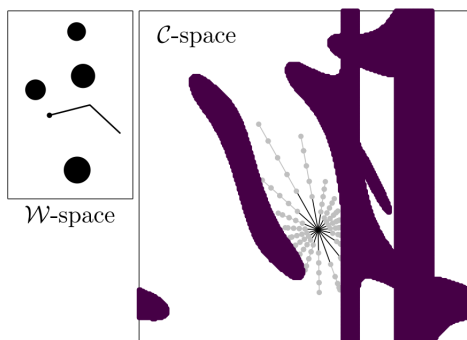


Рис. 1. Пример рабочего и конфигурационного пространств (изображение из [13])

Даны начальные конфигурации $q_{start}^i \in C_{free}^i$ и целевые конфигурации $q_{goal}^i \in C_{free}^i$. Также дано число $w > 1$, отвечающее за ограниченную субоптимальность. Решением задачи M-RAMP в дискретном времени является набор путей для отдельных агентов $\Pi = \{\pi_1, \pi_2, \dots, \pi_n\}$. Каждый путь $\pi_i = \{q_i^0, q_i^1, \dots, q_i^T\}$ удовлетворяет следующим условиям: $q_i^0 = q_i^{start}$, $q_i^T = q_i^{goal}$, все конфигурации $q_i^t \in C_{free}^i$, из конфигурации q_i^t в конфигурацию q_i^{t+1} можно добраться за одно действие, которое является поворотом в одном сочленении на угол минимального поворота δ , и между любыми парами агентов R_i , R_j нет конфликтов на всех временных шагах при следовании путям π_i и π_j . Также часто требуется не только найти допустимые пути для всех агентов, но и минимизировать стоимость этих путей. Обозначим за временную стоимость одного пути π_i такое наименьшее целое число $cost(\pi_i)$, что $\forall t \geq cost(\pi_i) : q_i^t = q_i^{goal}$, что соответствует времени, за которое конкретный агент добирается до целевой конфигурации и больше не уходит из неё. Тогда стоимостью решения задачи M-RAMP является следующее значение:

$$cost(\Pi) = \sum_{i=1}^n cost(\pi_i).$$

Набор путей считается корректным решением, если $cost(\Pi) \leq w \cdot cost(\Pi^*)$, где Π^* – это набор путей оптимальной стоимости. Такое решение называется ограниченно субоптимальным.

4. Метод. Предлагаемый в работе алгоритм реализует принципы конфликтно-ориентированного планирования, адаптирует эти принципы к рассматриваемой постановке и отличается от аналогов способом определения ограничений. Рассмотрим эти компоненты алгоритма более подробно далее.

4.1. Конфликтно-ориентированное планирование.

Конфликтно-ориентированное планирование было предложено в работе [18], где описан один из базовых алгоритмов реализующих этот принцип планирования – CBS (от англ. Conflict-Based Search). Основная идея конфликтно-ориентированного планирования заключается в двухуровневой структуре поиска. На верхнем уровне строится дерево разрешения конфликтов или же дерево ограничений (Constraint Tree), где каждая вершина содержит набор ограничений на движения агентов, а также индивидуальные пути для каждого агента, построенные с учётом этих ограничений. На нижнем уровне для каждого агента осуществляется поиск индивидуального пути с помощью стандартных алгоритмов (например, A^* или $weighted A^*$ [19]), учитывая наложенные ограничения.

Алгоритм CBS работает следующим образом:

- На начальном этапе для каждого агента строится оптимальный путь без учёта других агентов.
- Если обнаруживается конфликт (например, два агента занимают одну и ту же вершину в одно и то же время), то создаются две новые вершины дерева СТ, в каждой из которых одному из агентов добавляется ограничение, запрещающее конфликтное действие.
- Для каждой новой вершины СТ пересчитывается путь только для того агента, для которого добавлено новое ограничение.
- Процесс повторяется, пока не будет найдено решение, в котором отсутствуют конфликты между агентами.

На рисунке 2 представлен пример части верхнеуровневого дерева поиска. Пример взят для планирования путей для роботов на плоскости, так как он нагляднее и проще для понимания. На рисунке показано как конфликт в корневой вершине приводит к появлению двух новых вершин с дополнительными ограничениями.

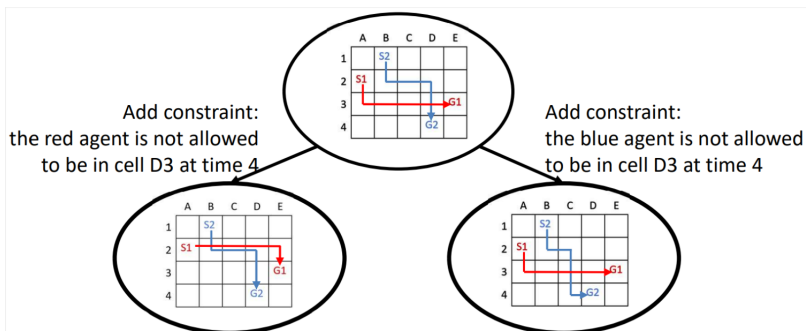


Рис. 2. Верхнеуровневое дерево поиска алгоритма CBS. В корневой вершине произошёл конфликт, из-за чего были созданы две новые вершины. В левой вершине был перепланирован красный путь с учётом ограничения, а в правой синий. Изображение взято из презентации "Model AI Assignment: Introduction to Multi-Agent Path Finding Wolfgang Hoenig, Jiaoyang Li, Sven Koenig

Данный подход позволяет эффективно разрешать конфликты, минимизируя количество перепланирований путей и обеспечивая оптимальность решения по суммарной длине путей агентов.

Алгоритм CBS является оптимальным и его эффективность опирается на предположение о том, что функция FindPath находит путь для одного робота быстро, что было верным для роботов, передвигающихся по плоскости, но не соответствует действительности для роботов-манипуляторов из-за большой размерности пространства конфигураций C^i .

По этой причине для роботов-манипуляторов был придуман алгоритм ECBS (Enhanced CBS) [20], являющийся ограниченно-субоптимальным, но позволяющим значительно быстрее находить решение. Модификация заключается в том, что помимо верхнеуровневой очереди OPEN алгоритм получает ещё и очередь FOCAL:

$$\text{FOCAL} := \{N \in \text{OPEN} \mid N.\text{cost} \leq w \cdot \min_{N' \in \text{OPEN}} \text{lb}(N')\},$$

где $\text{lb}(N) := \sum_{i=1}^n \text{lb}(\pi_i)$, и $\text{lb}(\pi_i)$ является нижней оценкой стоимости пути $N.\Pi[i]$. ECBS на каждой итерации главного цикла выбирает вершину N из FOCAL с наименьшим количеством конфликтов в ней:

$$N = \arg \min_{N' \in \text{FOCAL}} |N' \cdot \Omega|,$$

где $N \cdot \Omega$ – это множество конфликтов, обнаруженных среди набора путей N .П. Такой подход позволяет находить решения с общей стоимостью, не превышающей $w \cdot C^*$, где C^* – оптимальная суммарная стоимость. Как правило, такие решения находятся значительно быстрее, чем при использовании алгоритма CBS.

4.2. Конфликтно-ориентированное планирование для манипуляторов. Для обеспечения полноты в CBS и его вариантах крайне важно, чтобы среди используемых ограничений были полные ограничения

Определение 1 (Полные ограничения). Ограничение является полным, если при использовании для разрешения конфликта между агентами R_i и R_j , путём наложения ограничений c_i и c_j соответственно, не существует двух путей без конфликтов для R_i и R_j , где R_i нарушает c_i , а R_j нарушает c_j .

Использование только лишь полных ограничений зачастую является неэффективным, так как такие ограничения слишком слабо ограничивают движения агентов. Поэтому для планирования движений роботов-манипуляторов в работе [2] были предложены более рестриктивные, но не полные, ограничения, повышающие эффективность алгоритма конфликтно-ориентированного планирования на практике. Ниже описаны все известные на данный момент ограничения с указанием их свойств.

Ограничение на вершины (Vertex Constraints). Ограничения на вершины запрещают агенту R_i занимать определённую конфигурацию q_i^ω в момент времени t . Несложно убедиться, что это ограничение является полным, а вот следующие ограничения уже не будут являться полными

Ограничение на избегание (Avoidance Constraints). Ограничения на избегание запрещают агенту R_i сталкиваться с объёмом, занимаемым другим агентом $R_j(q_j^\omega)$ в момент времени t .

Ограничение на приоритеты (Priority Constraints). Ограничения на приоритеты запрещают агенту R_i сталкиваться с агентом R_j , пока тот движется по своему текущему пути π_j . Это ограничение используется в алгоритмах Prioritized Planning (PP) и PBS, однако оно также не является полным.

Сферическое ограничение (Sphere Constraints). Сферические ограничения запрещают агенту R_i сталкиваться с сферическим препятствием $S^2(p, r)$, имеющим радиус r и центр в точке столкновения

p , в момент времени t . При $r > 0$ это ограничение является неполным, однако при уменьшении радиуса до нуля оно становится полным.

На рисунке 3 изображения разные ограничения, которые могут быть наложены на агентов на основе одного конфликта. Преимуществом неполных ограничений является то, что они отсекают более обширную часть пространства поиска, что приводит к значительному ускорению, но при этом решение может не быть найдено из-за отсечения. Обозначим за k количество ограничений. Ради сохранения полноты алгоритм Generalized ECBS на верхнем уровне планирования рассматривает варианты с накладыванием всех k ограничений, а не только вершинного, в результате чего у каждой вершины верхнеуровневого дерева поиска возникает $2k$ потомков, а не 2, как это было в ECBS.

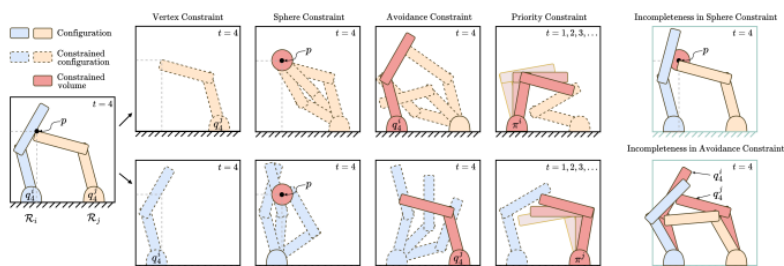


Рис. 3. Ограничения, введенные авторами алгоритма Generalized ECBS (изображение из [2])

4.3. Алгоритм GECBS-T. Идея ограничений алгоритма ECBS заключается в отсекании пространства поиска в вершинах, образованных неполными ограничениями. Но большинство из этих ограничений устанавливается на определённый момент времени, за который манипулятор может успеть выполнить только одно элементарное действие. Ограничения на избегание, приоритет и сферическое ограничение расширяют вершинное ограничение в пространстве конфигураций, то есть запрещают в конкретный момент времени находиться не в одной конфигурации, а в некотором подмножестве близких конфигураций. На практике же требуется и обратное, запрещение находиться в близкой конфигурации в течение некоторого продолжительного времени, так как для того, чтобы развести маршруты двух манипуляторов в пространстве может потребоваться большое количество шагов.

Второе соображение заключается в том, что ограничение, накладываемое на один момент времени, зачастую приводит к столкновению в следующий момент времени, при этом алгоритму

приходится накладывать ещё одно ограничение и перезапустить алгоритм планирования в вершине-потомке, что является дорогостоящей операцией. Так, например, происходит на рисунке 4 слева, при наложении ограничения на конфликтную конфигурацию в конкретный момент времени манипулятор оказывается в той же конфигурации в следующий момент времени, что также часто приводит к конфликту в реальных заданиях. В работе предлагается подход, который уменьшит количество запусков низкоуровневого планирования за счёт постановки ограничения сразу на продолжительное время.

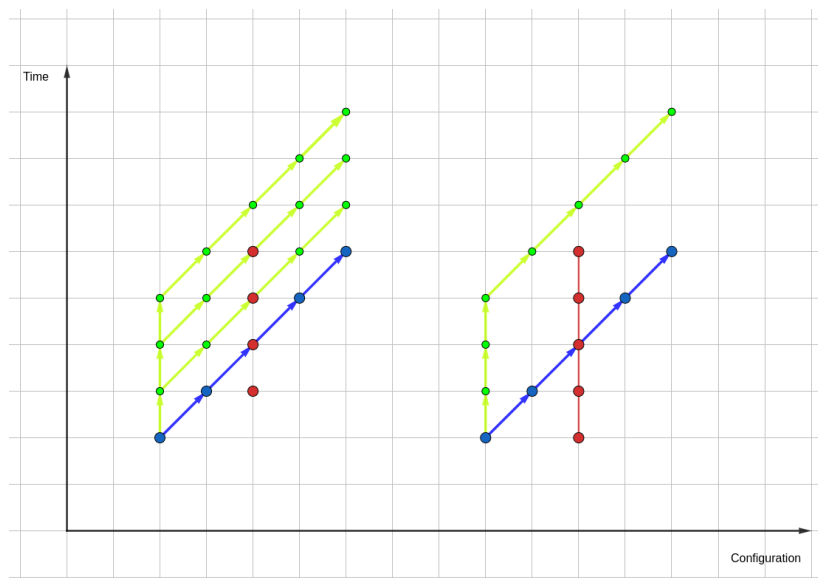


Рис. 4. По вертикали обозначена координата времени, по горизонтали обозначены координаты пространства конфигураций (одним измерением для упрощения). Синим обозначен путь манипулятора в пространстве-времени, красным обозначены конфликтные конфигурации. Синий путь – первый среди найденных планировщиков, в котором есть конфликт. Зеленым выделены пути, найденные планировщиком после синего с учётом конфликтов (в порядке наложения). Слева ограничения на конфликтные конфигурации были наложены по-отдельности, справа все конфигурации покрыты одним ограничением

Разработанный в данной работе алгоритм GECBS-T использует FOCAL очередь для ускорения поиска, а также неполные ограничения. Также алгоритм принимает на вход параметр T – длительность промежутка ограничений в тактах времени и вместо установки ограничения в момент

времени t устанавливать его с времени $t - T$ до $t + T$ включительно. Так, например, на рисунке 4 справа за счёт увеличенного по времени ограничения экономятся промежуточные вершины верхнеуровневого дерева CBS, что приводит к меньшему количеству низкоуровневых запусков планировщика. Такая формулировка ограничения не замедляет низкоуровневого планировщика и, даже наоборот, немного ускоряет за счёт того, что ограничение хранится в сжатом виде, а не как T экземпляров. Псевдокод данного алгоритма представлен ниже.

Алгоритм 1. GECBS-T

Require: agents, start, goal, T

```

1: root.constraints  $\leftarrow \emptyset$ 
2: for all agent  $i$  do
3:   root.path[ $i$ ]  $\leftarrow$  FindPath( $i$ , start[ $i$ ], goal[ $i$ ], root.constraints)
4: end for
5: root.cost  $\leftarrow$  sum of costs of all root.path[ $i$ ]
6: root.conflicts  $\leftarrow$  number of conflicts for all root.path[ $i$ ]
7: OPEN  $\leftarrow$  priority queue ordered by (cost)
8: FOCAL  $\leftarrow$  subset of OPEN with cost  $\leq w \cdot \min \text{cost}(\text{OPEN})$  ordered by
   (conflicts.size())
9: Insert root into OPEN and FOCAL
10: while FOCAL is not empty do
11:   node  $\leftarrow$  FOCAL.pop() ▷ Node with minimum conflicts
12:   if node.paths have no conflicts then
13:     return node.paths
14:   end if
15:   conflict  $\leftarrow$  earliest conflict in node.conflicts
16:   for all constraint type  $c \in$  all constraint types do
17:     for all agent  $i$  involved in  $c$  do
18:       child  $\leftarrow$  copy of node
19:       append new Constraint( $c, T, i, \text{conflict}$ ) to child.constraints
20:       child.path[ $i$ ]  $\leftarrow$  FindPath( $i, \text{start}[i], \text{goal}[i], \text{child.constraints}$ )
21:       if child.path[ $i$ ] exists then
22:         child.cost  $\leftarrow$  sum of costs of all child.path[ $j$ ]
23:         child.h  $\leftarrow$  heuristic estimate for child
24:         Insert child into OPEN
25:       end if
26:     end for
27:   end for
28:   Update FOCAL with all nodes in OPEN where cost  $\leq w \cdot \min(\text{OPEN})$ 
29: end while
30: return failure

```

В данном псевдокоде функция `FindPath` осуществляет поиск индивидуального пути для агента с учётом текущих ограничений (например, с помощью алгоритма A^*). А функция `Constraint` создаёт ограничение для агента, исходя из конфликта, параметра T и типа ограничения. Процесс продолжается до тех пор, пока не будет найдено решение без конфликтов или не будет исчерпан список возможных вариантов.

Приведем теперь утверждение, характеризующее основные теоретические гарантии, предоставляемые предлагаемым алгоритмом.

Утверждение. Алгоритм GECBS- T является полным, то есть если решение задачи планирования существует, то оно будет найдено алгоритмом. Более того, решения, возвращаемые GECBS- T являются ограниченно субоптимальными, то есть для любой заданной наперёд константы w (являющейся параметром алгоритма) верно, что стоимость найденного решения $cost$ не превосходит $w \cdot cost^*$, где $cost^*$ – стоимость оптимального решения.

Доказательство. Во-первых покажем, что в верхнеуровневом дереве ECBS всегда будет нераскрытая вершина, ограничения которой соответствуют одному из оптимальных решений, если решение задачи существует. Докажем это утверждение по индукции. Изначально это верно, так как в корневой вершине нет ограничений, а так как решение существует, то существует и оптимальное. При раскрытии каждой вершины либо раскрывается вершина, ограничения которой удовлетворяют оптимальному решению, либо нет. Если ограничения вершины противоречат оптимальному решению, то так как в дереве есть хотя бы одна вершина без противоречий ограничениям оптимальному решению – эти вершины разные и искомая вершина остаётся нераскрытой после данного шага. Если же вершина, ограничения которой удовлетворяют оптимальному решению, раскрывается, то за счёт наличия полного ограничения у неё есть два потомка, объединение решений которых содержит все решения без конкретного разрешённого конфликта по свойству полных ограничений, а значит одна из этих вершин не запрещает своими ограничениями оптимальное решение. Перейдём к доказательству ограниченной субоптимальности алгоритма с константой w . Так как в очереди OPEN всегда есть вершина, ограничения которой не противоречат оптимальному решению со стоимостью $cost^*$, то верно неравенство $\min_{N' \in OPEN} lb(N') \leq cost^*$, так как $lb(N) \leq cost^*$, где N – та самая вершина. Это означает, что в FOCAL стоимость всех вершин будет не больше чем $w \cdot cost^*$ по определению очереди FOCAL, следовательно каждая раскрытая вершина будет иметь решение, не

превосходящее $w \cdot cost^*$, значит и итоговое решение обладает этим же свойством.

5. Экспериментальное исследование. Предлагаемый метод был программно реализован на языке C++² и использован для проведения экспериментов в робототехническом симуляторе mujoco от Google Deep Mind. Все эксперименты проводились на компьютере с процессором Intel Xeon Gold 6338 с двумя ядрами и 16 ГБ оперативной памяти.

5.1. Подготовка тестовых заданий. Для проведения экспериментов использовалась программная модель современного робота-манипулятора iiwa-14 от kuka (рисунок 5), доступная в симуляторе mujoco.

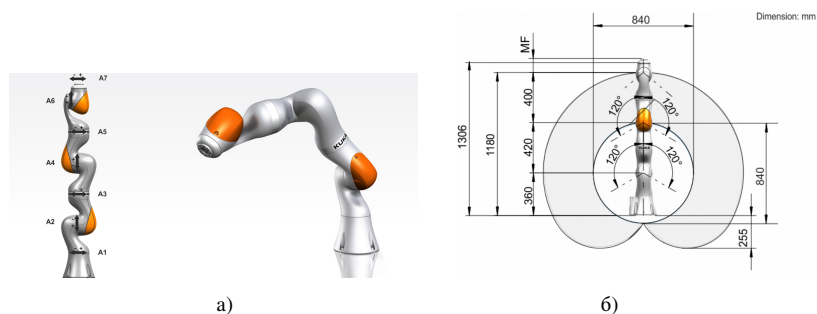


Рис. 5. Робот iiwa-14 от kuka с указанием размеров

Так как датасетов с заданиями для этой задачи очень мало и они являются специфичными для разных роботов и симуляторов, задания в этой работе были сгенерированы самостоятельно. Для подбора заданий было решено генерировать случайным образом начальные и конечные конфигурации для каждого манипулятора, при этом основания манипуляторов расположены в вершинах квадрата со стороной 1 метр.

При случайной генерации для решения многих заданий достаточно простого независимого планирования всех манипуляторов, так как при этом у найденных путей не возникало конфликтов. Такие задания также не попали в датасет, так как эти задания решаются каждым алгоритмом одинаково быстро, а значит они заведомо не позволяют сравнить два алгоритма. Также в результате случайной генерации заданий они получались достаточно простыми и почти независимыми, поэтому был найден способ генерировать в среднем более сложные задания для манипуляторов. Если рабочий орган каждого манипулятора находится

²https://github.com/machine-solution/motion_planning_for_manipulators

внутри куба $1 \times 1 \times 1$ метр, нижние вершины которого совпадают с креплениями четырёх манипуляторов, то независимо построенные траектории роботов-манипуляторов существенно мешают друг другу, что приводит к более сложным заданиям. Итоговый датасет собирался только из тех заданий, которые удовлетворяют данному условию.

Алгоритм GECBS-T сравнивался с алгоритмом Generalized ECBS для разных значений параметра T , из следующего набора: 5, 25, 50, 100, 250, 350, 425, 500 и 750.

Было подготовлено 5 сцен с 2, 3 и 4 манипуляторами, в трёх из которых нет препятствий, а в двух есть препятствие в центре сцены (рисунок 6).

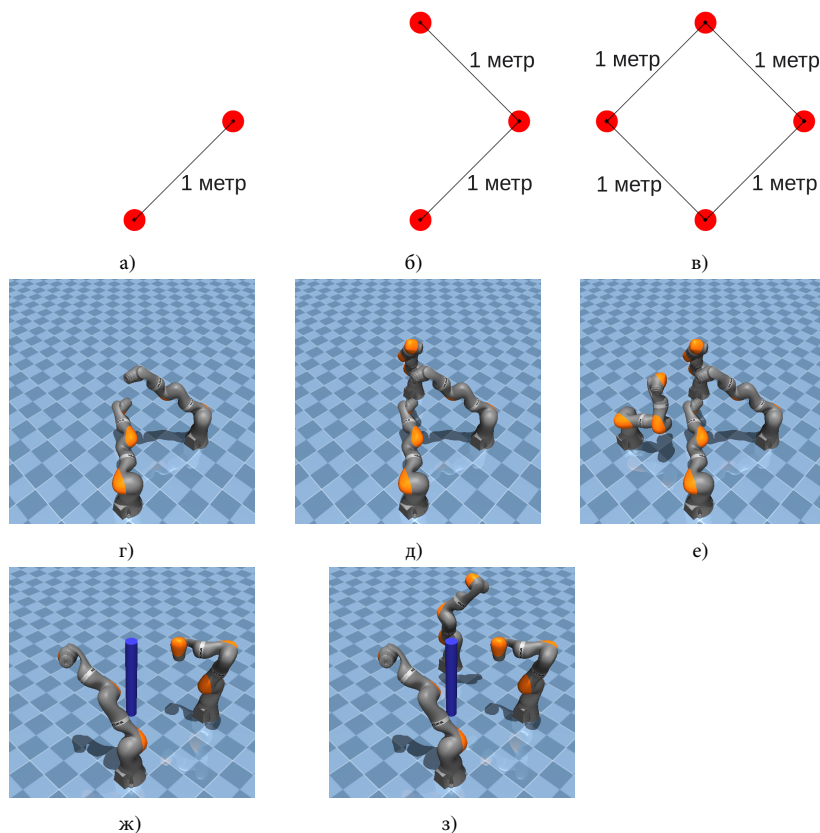


Рис. 6. Пример задания для: г, ж) двух, д, з) трёх и е) четырёх манипуляторов и а, б, в) расположение их оснований

Для каждой сцены генерировалось 100 случайных заданий как описано в предыдущей главе. Для каждого задания отводился бюджет времени 60 секунд. Если алгоритм не успевал найти ответ за это время, его выполнение останавливалось.

Для экспериментов был выбран минимальный угол поворота $\delta = \frac{\pi}{128}$ и любое элементарное действие является поворотом одного из сочленений на угол δ .

5.2. Результаты экспериментов. Алгоритмы GECBS-T и Generalized ECBS являются полными, то есть находят решение задания, если оно есть, но только в предположении бесконечного времени и вычислительных ресурсов. Если имеется ограничение по времени решения, как в проводимых экспериментах, то не для каждого алгоритм успевает найти ответ. Таким образом, возникает такой показатель вычислительной эффективности алгоритма как **доля решённых заданий** за отведенный временной лимит (60 секунд в рассматриваемом случае). Значение этого показателя приведено в таблице 1.

Таблица 1. Доля решённых заданий в зависимости от количества манипуляторов и значения параметра T

	2 ман-ра	3 ман-ра	4 ман-ра	2 ман-ра, преп.	3 ман-ра, преп.
Generalized ECBS	50%	42%	26%	45%	29%
GECBS-T, $T = 5$	55%	42%	28%	47%	30%
GECBS-T, $T = 25$	58%	44%	36%	47%	30%
GECBS-T, $T = 50$	57%	42%	31%	46%	33%
GECBS-T, $T = 100$	59%	49%	33%	49%	32%
GECBS-T, $T = 250$	69%	53%	41%	54%	33%
GECBS-T, $T = 350$	67%	57%	43%	54%	39%
GECBS-T, $T = 425$	66%	56%	31%	54%	39%
GECBS-T, $T = 500$	61%	50%	34%	54%	38%
GECBS-T, $T = 750$	55%	52%	29%	54%	35%

Первая строка в таблице 1 – это доля решенных заданий для алгоритма GECBS-T, остальные строки показывают долю решенных заданий для предлагаемого в работе алгоритма GECBS-T, для различных значений параметра T . Каждый столбец соответствует отдельной сцене с различным числом манипуляторов. Как видно из результатов, предлагаемый в работе алгоритм заметно превосходит предшественника по способности отыскивать решения за ограниченное время. Доля решённых заданий у алгоритма GECBS-T выше на 15% – 65% по сравнению с Generalized ECBS, в зависимости от количества манипуляторов, типа сцены (с препятствием или без) и значения параметра T . Для алгоритма GECBS-T наибольшее доля решенных заданий достигается для значений параметра $T = 250$, и $T = 350$. При

этом до некоторого значения параметра T эта доля растёт, после чего уменьшается за счёт того, что такая длина интервала уже является избыточной.

Второй метрикой для сравнения алгоритмов является скорость работы. На рисунках 7, 8 представлены графики зависимости количества решённых заданий (ось Y) от бюджета времени (ось X) для сцен с тремя манипуляторами.

Количество решённых заданий в зависимости от бюджета времени для 3 манипуляторов

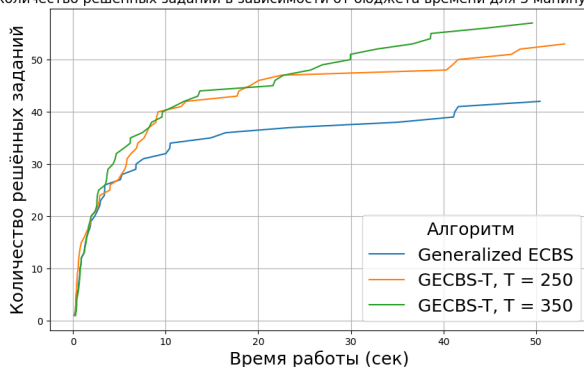


Рис. 7. Количество решённых заданий в зависимости от бюджета времени для трёх манипуляторов для сцены без препятствий

Решённые задания в зависимости от бюджета времени, 3 манипулятора

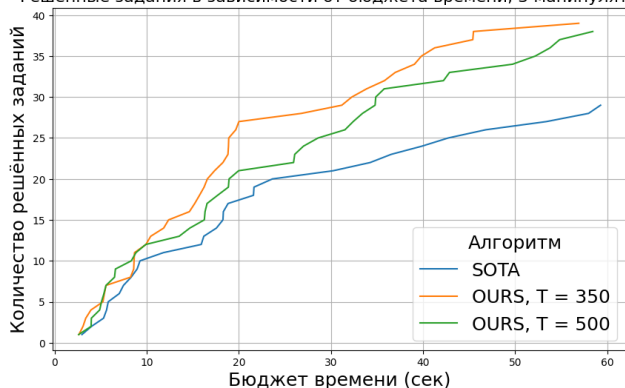


Рис. 8. Количество решённых заданий в зависимости от бюджета времени для трёх манипуляторов для сцены с препятствиями

На графиках отображены алгоритмы Generalized ECBS и GECBS-T с параметрами $T = 250, 350$. По этим результатам видно, что для каждой сцены есть какое-то количество заданий, которые решаются одинаково быстро всеми версиями алгоритма, так как они очень простые, зато для более сложных заданий, которые всё ещё решаются алгоритмами за минуту, которые составляют долю 20% – 35%, представленный в работе алгоритм решает в 2-5 раз быстрее. Именно за счёт большей скорости решения сложных заданий алгоритм GECBS-T показывает большую долю решённых заданий.

Перейдем теперь к сравнению качества решений, полученных алгоритмом Generalized ECBS и GECBS-T (для параметров $T = 250, T = 350$ и $T = 500$). Для каждого задания, которое решено алгоритмами, было рассчитано значение $\frac{cost(\Pi)}{cost(\Pi^*)} - 1$, где Π – решение, выданное алгоритмом, а Π^* – решение, полученное из Generalized ECBS. Посчитанное значение показывает превышение стоимости найденного пути по сравнению с оригинальным алгоритмом. Все эти значения, усреднённые по всем заданиям, показаны в таблице 2.

Таблица 2. Среднее превышение стоимости пути по сравнению с алгоритмом Generalized ECBS в зависимости от параметра T

	2 ман-ра	3 ман-ра	4 ман-ра	2 ман-ра, преп.	3 ман-ра, преп.
Generalized ECBS	0%	0%	0%	0%	0%
GECBS-T, $T = 250$	4.65%	0.63%	0.72%	0.04%	0.04%
GECBS-T, $T = 350$	3.30%	0.82%	0.76%	0.03%	0.01%
GECBS-T, $T = 350$	3.01%	0.71%	0.69%	0.03%	0.02%

Видно, что GECBS-T получает в среднем решения с большей стоимостью, но незначительно. Это показывает, что он практически ни в чём не уступает оригинальному алгоритму, но решает задания гораздо быстрее.

5.3. Выводы. Экспериментальные исследования показывают, что разработанный алгоритм GECBS-T превосходит Generalized ECBS по нескольким ключевым метрикам. Проведённые эксперименты выявили следующие преимущества.

1. Скорость решения заданий. GECBS-T на сложных заданиях показывает увеличение скорости в 2-5 раз по сравнению с Generalized ECBS без потери в скорости на простых заданиях, что делает его значительно более быстрым.

2. Выше доля решённых заданий. По сравнению с Generalized ECBS, GECBS-T решает больше задач в отведённый временной бюджет. Доля решённых заданий зависит от параметра алгоритма T , для которого

эксперименты показывают, что оптимальное значение находится между 250 и 350 для $\delta = \frac{\pi}{128}$, где δ – это минимальный угол поворота. Такой эффект достигается за большей скорости работы на сложных заданиях

3. Сравнимое качество решения. Алгоритм GECBS-T выдаёт решения с превышением стоимости менее чем на 5% относительно решений Generalized ECBS в худшей сцене и с превышением стоимости менее чем на 1% в остальных сценах, что означает практически такое же фактическое качество решения. При этом он сохраняет все теоретические гарантии на стоимость пути.

Исходя из этих данных можно сделать вывод, что GECBS-T из данной работы показывает себя намного быстрее и эффективнее, чем базовый алгоритм Generalized ECBS. Эти результаты достигаются благодаря меньшему количеству обращений к низкоуровневому планировщику за счёт более грубых ограничений. По этой причине GECBS-T незначительно уступает Generalized ECBS по качеству решений, но при этом сохраняет теоретические гарантии, оставаясь полным и ограниченно субоптимальным. Таким образом, предложенный алгоритм получает значительно более эффективны решения задач M-RAMP с гарантией ограниченной субоптимальности.

6. Заключение. В данной работе предложен алгоритм планирования согласованных траекторий системы из нескольких многозвенных манипуляторов GECBS-T, основанный на методе конфликтно-ориентированного планирования и доказана его корректность. Алгоритм был реализован и проведены экспериментальные исследования, в которых алгоритм GECBS-T сравнивался с алгоритмом Generalized ECBS. Эксперименты показали значительно большую эффективность представленного в статье алгоритма. В дальнейшем планируется исследование динамического определения длительности интервала ограничений в алгоритме в зависимости от уже построенных путей, улучшение системы приоритетов для разных ограничений в зависимости от показанной ими эффективности в прошлом и проведение экспериментов на реальных роботах.

Литература

1. Sharon G., Stern R., Felner A., Sturtevant N.R. Conflict-based search for optimal multi-agent pathfinding. Artificial intelligence. 2015. vol. 219. pp. 40–66.
2. Shaoul Y., Veerapaneni R., Likhachev M., Li J. Unconstraining multi-robot manipulation: Enabling arbitrary constraints in ecbs with bounded sub-optimality. Proceedings of the International Symposium on Combinatorial Search. 2024. vol. 17. pp. 109–117.
3. Ma H., Koenig S. Ai buzzwords explained: multi-agent path finding (mapf). AI Matters. 2017. vol. 3(3). pp. 15–19.

4. Stern R., Sturtevant N., Felner A., Koenig S., Ma H., Walker T., Li J., Atzmon D., Cohen L., Kumar T.K., et al. Multi-agent pathfinding: Definitions, variants, and benchmarks. *Proceedings of the International Symposium on Combinatorial Search*. 2019. vol. 10. pp. 151–158.
5. Ma H., Harabor D., Stuckey P.J., Li J., Koenig S. Searching with consistent prioritization for multi-agent path finding. In *Proceedings of the AAAI conference on artificial intelligence*. 2019. vol. 33. pp. 7643–7650.
6. Chan S.-H., Stern R., Felner A., Koenig S. Greedy priority-based search for suboptimal multi-agent path finding. *Proceedings of the International Symposium on Combinatorial Search*. 2023. vol. 16. pp. 11–19.
7. Zhang S., Li J., Huang T., Koenig S., Dilkina B. Learning a priority ordering for prioritized planning in multi-agent path finding. *Proceedings of the International Symposium on Combinatorial Search*. 2022. vol. 15. pp. 208–216.
8. Sharon G., Stern R., Felner A., Sturtevant N.R. Conflict-based search for optimal multi-agent pathfinding. *Artificial intelligence*. 2015. vol. 219. pp. 40–66.
9. Shaoul Y., Mishani I., Likhachev M., Li J. Accelerating search-based planning for multi-robot manipulation by leveraging online-generated experiences. *Proceedings of the International Conference on Automated Planning and Scheduling*. 2024. vol. 34. pp. 523–531.
10. Chen J., Li J., Huang Y., Garrett C., Sun D., Fan C., Hofmann A., Mueller C., Koenig S., Williams B.C. Cooperative task and motion planning for multi-arm assembly systems. *arXiv preprint arXiv:2203.02475*. 2022.
11. LaValle S. Rapidly-exploring random trees: A new tool for path planning. *Research Report 9811*. 1998.
12. Kuffner J.J., LaValle S.M. Rrt-connect: An efficient approach to single-query path planning. *Proceedings 2000 ICRA. Millennium conference. IEEE international conference on robotics and automation. Symposia proceedings (Cat. No. 00CH37065). IEEE*, 2000. vol. 2. pp. 995–1001.
13. Lacevic B., Osmankovic D.. Improved c-space exploration and path planning for robotic manipulators using distance information. *2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE*, 2020. pp. 1176–1182.
14. Karaman S., Frazzoli E. Sampling-based algorithms for optimal motion planning. *The international journal of robotics research*. 2011. vol. 30(7). pp. 846–894.
15. Zhang D., Zhu G., Zhang Q. Multi-robot motion planning: A learning-based artificial potential field solution. *2nd Conference on Fully Actuated System Theory and Applications (CFASTA). IEEE*, 2023. pp. 233–238.
16. Shaoul Y., Mishani I., Vats S., Li J., Likhachev M. Multi-robot motion planning with diffusion models. *arXiv preprint arXiv:2410.03072*. 2024.
17. Aljalbout E., Karl M., van der Smagt P. Clas: Coordinating multi-robot manipulation with central latent action spaces. *Learning for Dynamics and Control Conference. PMLR*, 2023. pp. 1152–1166.
18. Sharon G., Stern R., Felner A., Sturtevant N.R. Conflict-based search for optimal multi-agent pathfinding. *Artificial intelligence*. 2015. vol. 219. pp. 40–66.
19. Ebendt R., Drechsler R. Weighted a search-unifying view and application. *Artificial Intelligence*. 2009. vol. 173(14). pp. 1310–1342.
20. Li J., Ruml W., Koenig S. Eecbs: A bounded-suboptimal search for multi-agent path finding. In *Proceedings of the AAAI conference on artificial intelligence*. 2021. vol. 35(14). pp. 12353–12362.

Зайцев Андрей Сергеевич — студент, факультет математики и компьютерных наук, Санкт-Петербургский государственный университет. Область научных интересов: компьютерные науки, наука о данных. Число научных публикаций — 3. Dusha.Zaitsev@yandex.ru; Университетская набережная, 7–9, 199034, Санкт-Петербург, Россия; р.т.: +7(812)328–2000.

Яковлев Константин Сергеевич — канд. физ.-мат. наук, доцент, ведущий научный сотрудник, факультет математики и компьютерных наук, Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН). Область научных интересов: искусственный интеллект, интеллектуальные динамические системы, интеллектуальные системы управления, интеллектуальная робототехника, интеллектуальное планирование, планирование траектории, эвристический поиск, многоагентные системы, когнитивные агенты. Число научных публикаций — 75. yakovlev@isa.ru; улица Вавилова, 44/2, 119333, Москва, Россия; р.т.: +7(926)270-8979.

A.S. ZAITSEV, K.S. YAKOVLEV
**TIME RANGE CONSTRAINTS FOR MOTION PLANNING FOR
MANIPULATORS**

Zaitsev A.S., Yakovlev K.S. **Time Range Constraints for Motion Planning for Manipulators.**

Abstract. This paper addresses the problem of coordinated motion planning for multi-link robotic manipulator systems. One of the promising modern approaches to solving this problem is conflict-based planning, which avoids constructing a high-dimensional joint search space by sequentially solving a series of lower-dimensional problems. This is achieved by introducing spatio-temporal constraints whenever conflicts arise in individual manipulator plans, followed by replanning with these constraints in place. Unfortunately, existing methods that use constraints operate with individual time points, which reduces their practical efficiency. In this work, we present a novel conflict-based planning algorithm that utilizes interval-based temporal constraints rather than point-based ones – GECBS-T. Theoretically, the proposed algorithm guarantees bounded sub-optimality of the generated solutions; that is, for any user-defined bound $w > 1$, the cost of the GECBS-T solution will not exceed w times the cost of the optimal solution. In practice, the proposed algorithm significantly outperforms analogous algorithms in terms of planning speed, as confirmed by experiments conducted in the Mujoco robotics simulator involving 2–4 KUKA robotic manipulators, each with 7 degrees of freedom.

Keywords: motion planning, manipulators, M-RAMP, CBS, search-based, conflict-based, graph search, constraints, multi-agent, time range.

References

1. Sharon G., Stern R., Felner A., Sturtevant N.R. Conflict-based search for optimal multi-agent pathfinding. *Artificial intelligence*. 2015. vol. 219. pp. 40–66.
2. Shaoul Y., Veerapaneni R., Likhachev M., Li J. Unconstraining multi-robot manipulation: Enabling arbitrary constraints in ecbs with bounded sub-optimality. *Proceedings of the International Symposium on Combinatorial Search*. 2024. vol. 17. pp. 109–117.
3. Ma H., Koenig S. Ai buzzwords explained: multi-agent path finding (mapf). *AI Matters*. 2017. vol. 3(3). pp. 15–19.
4. Stern R., Sturtevant N., Felner A., Koenig S., Ma H., Walker T., Li J., Atzmon D., Cohen L., Kumar T.K., et al. Multi-agent pathfinding: Definitions, variants, and benchmarks. *Proceedings of the International Symposium on Combinatorial Search*. 2019. vol. 10. pp. 151–158.
5. Ma H., Harabor D., Stuckey P.J., Li J., Koenig S. Searching with consistent prioritization for multi-agent path finding. In *Proceedings of the AAAI conference on artificial intelligence*. 2019. vol. 33. pp. 7643–7650.
6. Chan S.-H., Stern R., Felner A., Koenig S. Greedy priority-based search for suboptimal multi-agent path finding. *Proceedings of the International Symposium on Combinatorial Search*. 2023. vol. 16. pp. 11–19.
7. Zhang S., Li J., Huang T., Koenig S., Dilkina B. Learning a priority ordering for prioritized planning in multi-agent path finding. *Proceedings of the International Symposium on Combinatorial Search*. 2022. vol. 15. pp. 208–216.
8. Sharon G., Stern R., Felner A., Sturtevant N.R. Conflict-based search for optimal multi-agent pathfinding. *Artificial intelligence*. 2015. vol. 219. pp. 40–66.

9. Shaoul Y., Mishani I., Likhachev M., Li J. Accelerating search-based planning for multi-robot manipulation by leveraging online-generated experiences. *Proceedings of the International Conference on Automated Planning and Scheduling*. 2024. vol. 34. pp. 523–531.
10. Chen J., Li J., Huang Y., Garrett C., Sun D., Fan C., Hofmann A., Mueller C., Koenig S., Williams B.C. Cooperative task and motion planning for multi-arm assembly systems. *arXiv preprint arXiv:2203.02475*. 2022.
11. LaValle S. Rapidly-exploring random trees: A new tool for path planning. *Research Report 9811*. 1998.
12. Kuffner J.J., LaValle S.M. Rrt-connect: An efficient approach to single-query path planning. *Proceedings 2000 ICRA. Millennium conference. IEEE international conference on robotics and automation. Symposia proceedings (Cat. No. 00CH37065)*. IEEE, 2000. vol. 2. pp. 995–1001.
13. Lacevic B., Osmankovic D. Improved c-space exploration and path planning for robotic manipulators using distance information. *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020. pp. 1176–1182.
14. Karaman S., Frazzoli E. Sampling-based algorithms for optimal motion planning. *The international journal of robotics research*. 2011. vol. 30(7). pp. 846–894.
15. Zhang D., Zhu G., Zhang Q. Multi-robot motion planning: A learning-based artificial potential field solution. *2nd Conference on Fully Actuated System Theory and Applications (CFASTA)*. IEEE, 2023. pp. 233–238.
16. Shaoul Y., Mishani I., Vats S., Li J., Likhachev M. Multi-robot motion planning with diffusion models. *arXiv preprint arXiv:2410.03072*. 2024.
17. Aljalbout E., Karl M., van der Smagt P. Clas: Coordinating multi-robot manipulation with central latent action spaces. *Learning for Dynamics and Control Conference*. PMLR, 2023. pp. 1152–1166.
18. Sharon G., Stern R., Felner A., Sturtevant N.R. Conflict-based search for optimal multi-agent pathfinding. *Artificial intelligence*. 2015. vol. 219. pp. 40–66.
19. Ebdend R., Drechsler R. Weighted a search—unifying view and application. *Artificial Intelligence*. 2009. vol. 173(14). pp. 1310–1342.
20. Li J., Ruml W., Koenig S. Eecbs: A bounded-suboptimal search for multi-agent path finding. In *Proceedings of the AAAI conference on artificial intelligence*. 2021. vol. 35(14). pp. 12353–12362.

Zaitsev Andrey – Student, Faculty of mathematics and computer science, St. Petersburg State University. Research interests: computer science, data science. The number of publications — 3. Dusha.Zaitsev@yandex.ru; 7–9, University Emb., 199034, St. Petersburg, Russia; office phone: +7(812)328–2000.

Yakovlev Konstantin – Ph.D., Associate Professor, Leading researcher, Faculty of mathematics and computer science, Federal Research Center for Computer Science and Control of the Russian Academy of Sciences (FRCSC RAS). Research interests: artificial intelligence, intelligent dynamic systems, intelligent control systems, intelligent robotics, intelligent planning, trajectory planning, heuristic search, multi-agent systems, cognitive agents. The number of publications — 75. yakovlev@isa.ru; 44/2, Vavilova St., 119333, Moscow, Russia; office phone: +7(926)270-8979.

И.С. АРХИПОВ, В.А. ЕРОФЕЕВА, О.Н. ГРАНИЧИН, В.А. КИСЕЛЕВ,
А.О. ЧЕРНОВ

УСКОРЕННЫЙ ПРОТОКОЛ ЛОКАЛЬНОГО ГОЛОСОВАНИЯ ДЛЯ ГРУППЫ РОБОТОВ С ОДНИМ ПУЛЬТОМ ДИСТАНЦИОННОГО УПРАВЛЕНИЯ

Архипов И.С., Ерофеева В.А., Граничин О.Н., Киселев В.А., Чернов А.О. Ускоренный протокол локального голосования для группы роботов с одним пультом дистанционного управления.

Аннотация. Управление группой роботов с помощью одного пульта дистанционного управления является актуальной и сложной задачей, особенно в условиях нестабильной связи, когда часть роботов может временно терять управляющий сигнал, что требует разработки надёжных децентрализованных механизмов для поддержания строя. В данной работе предложена и апробирована полуцентрализованная система управления, позволяющая оператору координировать группу роботов как единое целое. Система объединяет централизованные команды от базовой станции с децентрализованной коррекцией положения на основе протокола ESP-NOW. Для сравнения эффективности в задаче поддержания жёсткой формации были применены протокол локального голосования (LVP) и его ускоренная версия (ALVP). Их эффективность оценивалась в среде моделирования на группе из четырёх дронов в ходе экспериментов, включавших резкие манёвры (повороты на 50° и 75°) и имитацию значительных потерь пакетов данных (50% и 80%). Результаты показали, что ALVP продемонстрировал значительные преимущества по сравнению со стандартным LVP: более высокую скорость восстановления формации, меньшую среднюю ошибку позиционирования и большую устойчивость. В частности, при выполнении серии из 20 полётов с поворотом на 50° , ALVP успешно сохранил строй в 17 случаях против 3 у LVP, а также показал лучшую робастность в условиях потерь пакетов. Таким образом, предложенный полуцентрализованный подход с использованием протокола ALVP является эффективным и надёжным решением для управления строем группы роботов. Дальнейшие исследования будут направлены на проведение натурных экспериментов и интеграцию механизмов обхода препятствий.

Ключевые слова: кооперативное управление, адаптивное управление, робототехника.

1. Введение. В последние годы исследователи проявляют все больший интерес к использованию групп автономных роботов. Эти группы, включающие беспилотные летательные аппараты (а иногда наземные или морские системы), применяются в широком спектре задач, таких как сельское хозяйство, спасательные операции, мониторинг дикой природы, климатология, картографирование, инспекция инфраструктуры и доставка грузов. [1,2]. С практической точки зрения, одной из важнейших задач в этой области является поддержание строгого порядка формации группы роботов [3].

Существует два основных подхода к формированию структуры группы: централизованный и децентрализованный [1]. Централизованный

способ предполагает управление каждым роботом через базовую станцию, которая собирает данные от всех членов группы, определяет их местоположение и рассчитывает управляющие воздействия. Однако этот подход сталкивается с ограничениями: по мере увеличения числа роботов вычислительная и коммуникационная нагрузка на базовую станцию значительно возрастает, что снижает общую эффективность системы [3]. Кроме того, базовая станция должна поддерживать постоянную связь со всеми роботами, что приводит к неизбежным задержкам при обработке поступающей информации. Разрывы связи для такой системы критичны.

Децентрализованный подход позволяет каждому роботу самостоятельно определять необходимые для поддержания формации управляющие воздействия, обмениваясь информацией с соседними роботами. Для поддержки таких взаимодействий часто используются консенсусные алгоритмы [4], позволяющие роботам корректировать свои движения в зависимости от состояния находящихся поблизости членов группы и обеспечивающие согласованное поведение всей группы [4, 5].

В этом исследовании представлен полуцентрализованный способ управления группой роботов. Оператор при помощи пульта или базовая станция при помощи автоматического алгоритма транслирует общее для всех дронов управляющее воздействие (централизованная часть управления). Роботы, получившие эти команды, исполняют их. Роботы, связь с которыми оказалась разорвана, поддерживают формацию за счет локального взаимодействия с роботами, получившими управление от базовой станции (децентрализованная часть управления). Способность роботов группы самостоятельно регулировать свое положение и поддерживать формацию также снижает нагрузку на базовую станцию, связанную с необходимостью обрабатывать информацию от каждого отдельного робота. Предлагаемый подход облегчает управление группой как единым целым, упрощая контроль и координацию.

В этой статье рассматривается применение протокола локального голосования (Local Voting Protocol, LVP) [6] и ускоренного протокола локального голосования (Accelerated LVP, ALVP) [7] для поддержания жесткой формации группы роботов. Ранее протокол локального голосования был успешно применен к задаче планирования в реальной беспроводной сети [8]. Позже было предложено применение LVP для задачи о группе роботов, а именно для задачи синхронизации курса движения колесных роботов [9].

Протокол локального голосования – это масштабируемый распределенный протокол достижения консенсуса. Его основная идея заключается в постепенном выравнивании некоторых параметров агентов

системы между собой. Распределенный характер этого протокола также позволяет преодолеть коммуникационные ограничения. Известно, что он работает в сетях с переменной топологией и в условиях шумов, где обмен информацией между сетевыми агентами может быть потерян или задержан [10]. Также, во время работы системы некоторые роботы могут потерять связь с базовой станцией. LVP частично компенсирует такую потерю данных: роботы, имеющие связь с базовой станцией продолжают движение по заданной траектории, а роботы, связь потерявшие, ориентируются на движение соседей и за счет этого продолжают движение по траектории.

Протокол ускоренного локального голосования – это новый децентрализованный протокол достижения консенсуса, разработанный для нестационарных сетевых сред, где состояние узлов (агентов) динамически меняется. Степень несбалансированности состояний узлов сети характеризуется потенциалом Лапласа [11]. ALVP основывается на принципах консенсуса, поскольку каждый узел итеративно обновляет свою стратегию выравнивания состояний, используя локально-обмениваемые зашумленные измерения соседних узлов и свои измерения. Используя метод ускорения по Нестерову [12], и модифицируя классический потенциал Лапласа с помощью регуляризующего члена для обеспечения сильной выпуклости и гладкости, алгоритм обеспечивает более быструю сходимость к сбалансированному состоянию всех узлов, т.е. к минимуму потенциала Лапласа. Его конструкция, в которой используются постоянные настройки размера шага и вспомогательные переменные, позволяет надежно отслеживать состояния даже в условиях внезапных изменений. В задаче балансировки нагрузки вычислительной сети ALVP превосходит традиционные протоколы, такие как стандартный протокол локального голосования и round-robin планирование, с точки зрения сокращения общесистемного времени обработки [7]. Кроме ALVP для ускорения сходимости используется группа Fast Laplacian-методов, например, [13]. Такие методы используют свойства спектра матрицы Лапласа графа для ускорения сходимости. Также применяются различные модификации Push-Sum-подхода, такие как [14, 15]. Эти методы могут работать в условиях однонаправленности связей, но эффективность таких методов может сильно зависеть от структуры и стабильности графа связей [16], чего не наблюдается у LVP и ALVP.

Оставшаяся часть этой статьи построена следующим образом: в разделе 2 даётся формальная постановка задачи управления группой роботов в условиях меняющихся связей и внешних возмущений. Раздел 3 описывает архитектуру системы, включая компоненты базовой станции,

бортовое оборудование (ArduPilot, ESP32) и операционную систему реального времени MartOS. В разделе 4 рассматривается протокол локального голосования (LVP) для поддержания формации, его алгоритмы и свойства сходимости в условиях помех. Раздел 5 посвящён анализу ускоренного протокола (ALVP) и оценке его скорости сходимости. В разделе 6 приводятся результаты экспериментального сравнения LVP и ALVP в среде моделирования при выполнении резких манёвров. Наконец, в разделе 7 кратко отмечаются направления будущих исследований, такие как обход препятствий и проведение натурных испытаний.

2. Постановка задачи. Рассмотрим группу, состоящую из N роботов. Динамическая топология связей в группе моделируется как изменяющийся во времени неориентированный граф $\mathcal{G}_t = (\mathcal{N}, \mathcal{E}_t)$, где множество вершин $\mathcal{N} = \{1, 2, \dots, N\}$ обозначает роботов, а набор ребер $\mathcal{E}_t \subseteq \mathcal{N} \times \mathcal{N}$ обозначает наличие связи между парами роботов в момент времени t . Для случая беспроводной коммуникации, рассматриваемого в этой работе, соседями будут считаться роботы группы, попадающие в радиус вещания робота i .

Предположение 1: Граф $\mathcal{G}_t$ является связным в течение определенного периода или в среднем: для любой пары вершин $(i, j) \in \mathcal{N}$ существует по крайней мере один путь, соединяющий их в течение некоторого интервала времени.

Такое предположение обычно используется в исследованиях, посвященных задачам оптимизации на динамических графах [17].

Рассмотрим следующую проблему управления группой роботов. Пусть $X_t \in \mathbb{R}^{Nd}$ – глобальный вектор состояния системы, представляющий собой объединение векторов размерности d , обозначающих состояния отдельных роботов. Обозначим целевую траекторию группы как $X_t^* \in \mathbb{R}^{Nd}$, определяющую движение компонент системы роботов в группе в виде определённой формации вдоль некоторой траектории.

Предположение 2: Для всех роботов существует единая стратегия управления U_t , направляющая их по желаемой траектории.

Эта стратегия управления направлена на решение следующей оптимизационной задачи: $\int_0^t \|X_t - X_t^*\| \xrightarrow[t \rightarrow \infty]{} \min$. В этой работе не будет выводиться такая управляющая стратегия. Вместо этого будем считать, что оптимизирующая управляющая стратегия U_t генерируется с использованием таких методов, как Model Predictive Control (MPC) [18] или Linear Quadratic Regulator (LQR) [19–21], или задаются человеком-оператором при помощи пульта дистанционного управления.

Таким образом, требуется разработать систему, которая позволит доводить управляющее воздействие U_t до всех членов группы роботов в условиях возможной потери связи между базовой станцией и некоторыми из роботов, при этом поддерживая заданную формацию.

3. Архитектура системы. Система состоит из группы роботов, соединенных между собой в коммуникационную сеть, и базовой станции, сигналы от которой способны получать все роботы группы. Общая структура системы представлена на рисунке 1.

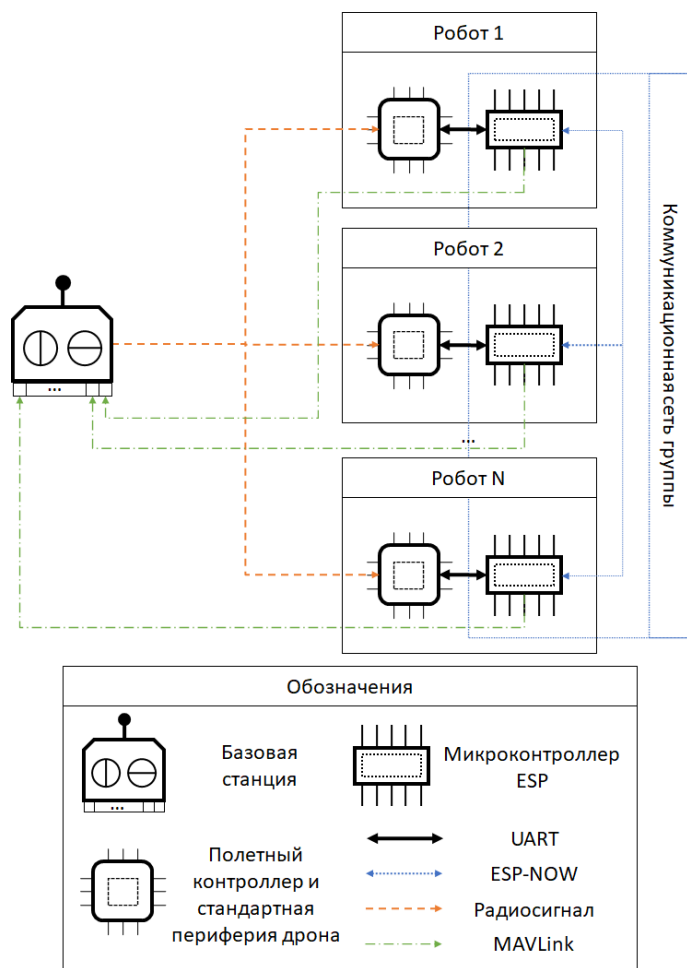


Рис. 1. Архитектура системы

Базовая станция является источником управления группой роботов. Управление может задаваться либо оператором при помощи пульта дистанционного управления, который отправляет команды для регулировки крена, тангажа, рыскания и тяги робота, либо автоматически при помощи алгоритма. Все роботы группы получают от базовой станции одинаковые управляющие сигналы, которые впоследствии преобразуются их бортовыми системами в команды управления двигателями.

Другим важным элементом базовой станции является компьютер, оснащенный специализированным программным обеспечением для мониторинга и управления группой. Каждый робот передает телеметрические данные, включая его текущее местоположение, скорость и ориентацию, используя протокол MAVLink [22]. Эти данные отображаются в программном интерфейсе (в представленном прототипе используется Mission Planner [23]), что позволяет оператору наблюдать за группой и управлять ей с повышенной точностью.

Регулировка положения в группе осуществляется программным обеспечением автопилота, работающим на контроллере полета каждого робота. В предлагаемой системе используется ArduPilot [24], программная платформа с открытым исходным кодом, которая может быть настроена для реализации конкретных алгоритмов управления, адаптированных к потребностям этого приложения.

Чтобы точно определить необходимые корректировки для движения, каждый робот полагается на данные о местоположении своих ближайших соседей. Для достижения этой цели каждый робот группы оснащен микроконтроллером ESP [25], который интегрирован с контроллером полета через универсальный асинхронный интерфейс приема-передающего устройства (UART).

Программное обеспечение автопилота постоянно предоставляет микроконтроллеру ESP обновленные данные о местоположении, которые в этой системе включают широту, долготу, высоту и ориентацию робота (рыскание). Затем микроконтроллер ESP передает эту информацию на устройства ESP ближайших роботов, используя Wi-Fi соединение.

Для обеспечения обмена данными без использования маршрутизатора используется протокол ESP-NOW [26]. Он отправляет важную информацию, заключенную в служебных сообщениях. Эти сообщения передаются в широковещательном режиме, что устраняет необходимость знать MAC-адреса роботов группы, создавать Wi-Fi-сеть или подключаться к уже существующей. Каждый находящийся в радиусе вещания микроконтроллер ESP способен получить эти сообщения и переслать содержащиеся в нем данные диспетчеру полета. Затем

диспетчер полета записывает местоположение соседа или обновляет его, если это местоположение было сохранено ранее.

Одной из ключевых функций автопилота робота является периодическая корректировка его движения. Это достигается путем использования сохраненных позиций соседних роботов для вычисления необходимых значений крена, тангажа, рыскания и тяги для достижения желаемого местоположения внутри группы. Затем эти рассчитанные значения объединяются с управляющими входами, полученными с пульта дистанционного управления.

Как передача телеметрии, так и внутригрупповой обмен информацией должны осуществляться параллельно. Для этой цели микроконтроллер ESP запущен на MartOS [27]. MartOS – это операционная система реального времени, специально разработанная для мультиагентного управления совместными роботизированными системами и распределенными сенсорными сетями. Она оснащена специализированным диспетчером задач без вытеснения, основанным на round-robin алгоритме, интегрированным управлением памятью и точных функциях таймера, которые облегчают синхронизацию срочных задач между несколькими агентами, тем самым оптимизируя совместное управление. Разработанная в основном на языке программирования Rust с поддержкой C в качестве альтернативы, MartOS использует особенности Rust для повышения безопасности и производительности с целью обеспечения надежной работы даже на устройствах с ограниченными ресурсами, таких как микроконтроллеры ESP32. Ключевым аспектом её архитектуры является встроенный сетевой интерфейс, который обеспечивает эффективную связь между агентами в режиме реального времени.

4. Протокол синхронизации. Управление согласно протоколу локального голосования для i -го робота в момент времени t может быть обозначено как $\bar{u}_t^i$. В простейшем случае $\bar{u}_t^i = ([\bar{u}_x]_{t+1}^i, [\bar{u}_y]_{t+1}^i, [\bar{u}_z]_{t+1}^i, [\bar{u}_h]_{t+1}^i)$. Также можно рассмотреть управляющий вектор с большим количеством измерений, где 4 его элементами являются $[\bar{u}_x]_{t+1}^i, [\bar{u}_y]_{t+1}^i, [\bar{u}_z]_{t+1}^i, [\bar{u}_h]_{t+1}^i$, а остальные члены равны 0.

4.1. Выравнивание курса и высоты. Для обеспечения единообразия курса (рыскания) и высоты полета роботов группы используются методы, основанные на консенсусе. Протокол локального голосования, который облегчает согласование конкретных параметров, может быть применен непосредственно для решения этой проблемы.

Пусть z_t^i обозначает высоту полета i -го робота в момент времени t , а $\mathcal{N}_t^i$ обозначает набор соседей для i -го робота в момент времени t . Используя LVP, каждый робот вычисляет свою локальную настройку управления $[\bar{u}_z]_{t+1}^i$ следующим образом:

$$[\bar{u}_z]_{t+1}^i = \gamma_z \sum_{j \in \mathcal{N}_t^i} (z_t^j - z_t^i), \quad (1)$$

где γ_z – это размер шага для протокола.

Направление движения робота определяется как угол по часовой стрелке между его текущим направлением и истинным севером. Пусть направление движения i -го робота в момент времени t обозначается как h_t^i . Для точных расчетов все углы должны оставаться в пределах $[-180; 180]$. Элемент управления корректировкой для выравнивания курса, $[\bar{u}_h]_{t+1}^i$, вычисляется как:

$$[\bar{u}_h]_{t+1}^i = \gamma_h \sum_{j \in \mathcal{N}_t^i} (h_t^j - h_t^i), \quad (2)$$

где γ_h представляет собой размер шага для выравнивания курса.

4.2. Выравнивание положения по горизонтали. Для достижения координации в горизонтальном расположении роботов группы применяется стратегия распределенного консенсуса. Пусть (x_t^i, y_t^i) представляет горизонтальное положение i -го робота в момент времени t . При использовании данных глобальных навигационных спутниковых систем (ГНСС) эти координаты могут потребовать преобразования из значений широты и долготы. Предполагается, что для предлагаемого метода необходимо ГНСС или любое другое позиционирование.

Подход, основанный на консенсусе, позволяет каждому роботу уточнять свою позицию, взаимодействуя с соседними роботами. Этот процесс включает в себя выбор роботом i лучшего положения, обозначаемого как $(\hat{x}_t^j, \hat{y}_t^j)$, из нескольких предложенных соседом j положений. Рисунок 2 иллюстрирует эту идею.

Корректировки для горизонтальных координат, $[\bar{u}_x]_{t+1}^i$ и $[\bar{u}_y]_{t+1}^i$, определяются с использованием следующих уравнений:

$$\begin{cases} [\bar{u}_x]_{t+1}^i = \gamma_{xy} \sum_{j \in \mathcal{N}_t^i} (\hat{x}_t^j - x_t^i) \\ [\bar{u}_y]_{t+1}^i = \gamma_{xy} \sum_{j \in \mathcal{N}_t^i} (\hat{y}_t^j - y_t^i) \end{cases}, \quad (3)$$

где γ_{xy} – это размер шага протокола, равномерно применяемый к обоим горизонтальным измерениям управления.

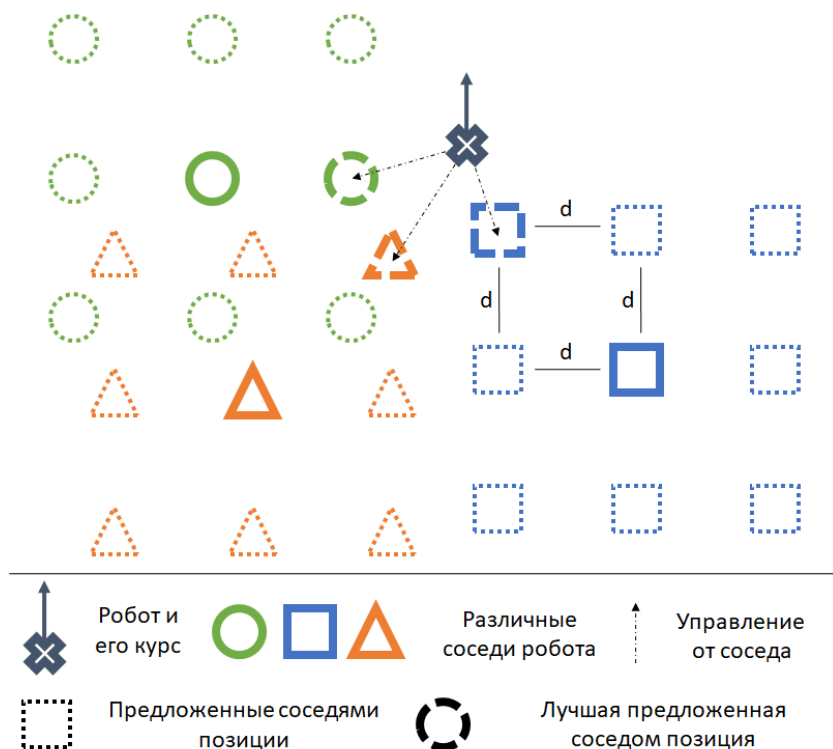


Рис. 2. Расчёт позиции по данным от соседа

В представленном примере протокол локального голосования используется для поддержания группой сеточной структуры с расстоянием d между соседями. Рисунок 3 демонстрирует некоторые формации, которые могут образоваться благодаря механизму консенсуса, представленному на рисунке 2. При помощи протокола локального голосования возможно поддерживать и другие конфигурации формации: для этого необходимо лишь изменить предлагаемые соседями положения.

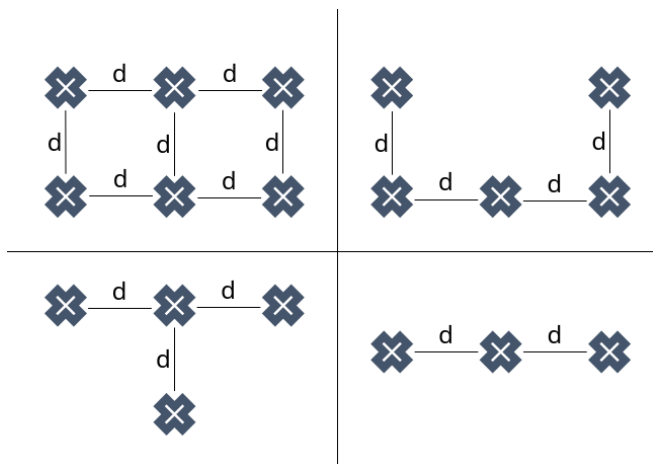


Рис. 3. Допустимые формации группы

4.3. Управление роботами. Объединенное глобальное и локальное управление u_t^i i -го робота в момент времени t может быть описано как:

$$u_t^i = U_t^i + \bar{u}_t^i. \quad (4)$$

Внешнее управление U_t^i в большинстве случаев соответствует стратегии управления U_t , описанной в разделе 2. Предполагается, что в некоторые моменты времени t связь между наземной станцией и роботом i может быть потеряна, и в этом случае U_t^i отличается от U_t . В этом случае U_t^i может остаться таким же, каким он был на предыдущем шаге (то есть $U_t^i = U_{t-1}^i$) или измениться на нулевой вектор. Более подробно решение проблемы управления группой описано в статье [28].

Динамика робота i моделируется уравнением состояния в пространстве с дискретным временем:

$$X_{t+1}^i = A_t^i X_t^i + B_t^i u_t^i + w_{t+1}^i, \quad (5)$$

где A_t^i – матрица перехода состояния, B_t^i – матрица управления состоянием, а w_t^i – возмущение.

Ожидается, что в идеальных условиях – при отсутствии шума и непрерывной связи – группа будет следовать по заданной траектории, определенной в предположении 2. Однако в реальных сценариях измерения часто сопровождаются шумом, задержками и случайной потерей связи между роботами или с наземной станцией. Свойства конвергенции LVP [10], установленные в предыдущих исследованиях, гарантируют, что даже в этих сложных условиях группа сможет достичь консенсуса и поддерживать желаемую формацию в течение долгого времени.

На практике важно, чтобы процесс достижения консенсуса происходил как можно быстрее. Для этой цели предлагается использовать протокол ускоренного локального голосования [7].

5. Ускоренный протокол локального голосования. Ускоренный протокол локального голосования [7] улучшает децентрализованную балансировку нагрузки в мультиагентных сетях в нестационарных условиях. Алгоритм основан на протоколе локального голосования, который основан на оптимизации на основе консенсуса, но использует методы ускорения, основанные на методах Нестерова, для повышения скорости сходимости [12]. ALVP работает за счет динамического перераспределения рабочей нагрузки между узлами, используя локальную связь и измерения для эффективной балансировки системы. Он включает в себя корректировку функции потенциала Лапласа, добавление регуляризационных условий для обеспечения сильной выпуклости и оптимизации распределения задач даже в изменяющихся во времени средах.

Одним из ключевых нововведений ALVP является его способность работать в нестационарных условиях, когда поступающие задачи постоянно нарушают баланс системы. В отличие от LVP, который эффективен в основном в стационарных сценариях, ALVP адаптируется к динамическим изменениям в распределении рабочей нагрузки и отслеживает эти изменения, чтобы создавать решения, приближенные к оптимальной стратегии балансировки нагрузки. Алгоритм достигает этого за счет использования механизма с переменным размером шага, который ускоряет сходимость, сохраняя при этом устойчивость к зашумленным измерениям и сбоям связи. ALVP устойчиво работает в условиях unknown-but-bounded шумов, задержках и потерях информации, что показано в [7].

Скорость выравнивания курса летательного аппарата определяется требуемым числом итераций алгоритма. Для ALVP это число может быть

определено с помощью теоремы 1, доказанной в [29] для другого случая применения протокола ALVP.

Теорема 1. Оценка ошибки ALVP [29].

Рассмотрим последовательность оценок $\{\hat{\mathbf{y}}_t\}_{t=0}^{\infty}$, сгенерированных ALVP, и предположим, что существует глобальный оптимум $\bar{\mathbf{y}}_t$ функции

$$R_t(\hat{\mathbf{y}}) = \frac{(1 + \lambda)}{2} \hat{\mathbf{y}}^T \left(\mathbf{I}_m \otimes \mathcal{L}(B(t)) \right) \hat{\mathbf{y}}. \quad (6)$$

Предположим, что для каждой итерации t выполняется следующее неравенство [7]:

$$\mathbb{E}_t [R_t(\hat{\mathbf{y}}_t) - R_t(\bar{\mathbf{y}}_t)] \leq \lambda_t (\phi_0(\bar{\mathbf{y}}_0) - R_t(\bar{\mathbf{y}}_t) + \Phi) + D_t, \quad (7)$$

где:

- $\phi_0(\bar{\mathbf{y}}_0)$ – начальная суррогатная функция,
- Φ – конечная константа,
- λ_t характеризует скорость затухания ошибок,
- $D_t \geq 0$ – это остаточная ошибка, удовлетворяющая $D_t < D_{\infty} < \infty$ и $D_t \rightarrow D_{\infty}$ при $t \rightarrow \infty$.

Тогда выполняются следующие утверждения:

1. **Ошибка после K итераций.** Для любого $K \geq 1$:

$$\begin{aligned} & \mathbb{E}_K [R_K(\hat{\mathbf{y}}_K) - R_K(\bar{\mathbf{y}}_K)] \\ & \leq \lambda_K (\phi_0(\bar{\mathbf{y}}_0) - R_K(\bar{\mathbf{y}}_K) + \Phi) + D_K. \end{aligned} \quad (8)$$

2. **Предел асимптотической погрешности.** В пределе:

$$\limsup_{K \rightarrow \infty} \mathbb{E}_K [R_K(\hat{\mathbf{y}}_K) - R_K(\bar{\mathbf{y}}_K)] \leq D_{\infty}. \quad (9)$$

3. **Минимальное количество итераций для заданной точности.**

Если требуется целевая точность $\epsilon > D_{\infty}$, то минимальное число итераций $K_{\min}$ удовлетворяет:

$$K_{\min} = \left\lceil \frac{\ln \left(\frac{\epsilon - D_K}{C} \right)}{\ln(1/\lambda)} \right\rceil, \quad (10)$$

где $C = \phi_0(\bar{\mathbf{y}}_0) - R_K(\bar{\mathbf{y}}_K) + \Phi$.

Доказательство теоремы представлено в [29]. Доказательство соответствия постановки задачи условиям теоремы приведено в приложении к статье.

6. Эксперименты. Предлагается следующий план экспериментов:

Цель экспериментов:

- Сравнить скорость восстановления идеальной формации квадрата при применении стандартного протокола локального голосования и его ускоренной версии ALVP.

- Проверить устойчивость каждого алгоритма к различным углам поворотов группы и длительности манёвра.

- Исследовать эффективность алгоритмов при восстановлении формации из произвольного начального состояния с ошибкой позиционирования.

- Оценить воспроизводимость результатов при многократных запусках.

Окружение:

- Специализированный симулятор [30], предназначенный для воспроизведения динамики групп дронов, что позволяет всесторонне протестировать их поведение в различных условиях.

- 4 однородных дрона в симуляционной среде с бортовыми контроллерами ArduPilot и ESP32 (протокол ESP-NOW).

- Наземная станция с пультом дистанционного управления и интерфейсом Mission Planner в симуляционной среде.

Метрика оценки:

- Расстояние (в метрах), которое должен пройти каждый дрон, чтобы вернуться в исходную, идеально заданную формацию. Иными словами, ошибка измеряется как суммарный путь, необходимый для полного восстановления правильной формации после нарушения.

Сценарии экспериментов:

- Поворот 50° . Длительность манёвра: 15 секунд.

- Поворот 75° . Длительность манёвра: 7 секунд.

- Поворот 50° . Длительность манёвра: 15 секунд. Потеря 50% пакетов.

- Поворот 50° . Длительность манёвра: 15 секунд. Потеря 80% пакетов.

– Восстановление формации из произвольного начального состояния. Дроны запускаются с нарушенной формацией (случайные отклонения от идеальных позиций) и стремятся достичь правильного построения квадрата без внешних возмущающих маневров.

Настройки протоколов:

– Для ALVP: $L = 3, \mu - \eta = 10, h = 0.3, \gamma_0 = 10, \alpha_x = 0.3$.
Параметры взяты, исходя из рекомендаций, описанных в [7].

– Для LVP: $\alpha = 0.3$. Параметр взят исходя из рекомендаций, описанных в [10].

Процедура проведения:

– Для сценария восстановления формации: запуск симуляции с дронами в произвольных начальных позициях, отклоняющихся от идеальной формации квадрата на случайные расстояния. Фиксация процесса сходимости к целевой формации без внешних возмущений.

– Для сценариев с поворотами: запуск симуляции полёта 4-х дронов в формации в виде квадрата, подача команды поворота под заданным углом на указанную длительность.

– Фиксация данных телеметрии каждого дрона (положение, расстояние до восстановления формации).

– Повторение сценария восстановления формации по 20 запусков для LVP и 20 запусков для ALVP.

– Повторение каждого из сценариев с поворотами по 20 запусков для LVP и 20 запусков для ALVP.

– Для сценариев с потерей сообщений (50% и 80%) – по 10 запусков для LVP и по 10 запусков для ALVP.

Траекторию полета дронов для первого сценария можно увидеть на рисунке 4. Графики ошибок полёта для первого сценария показаны на рисунке 5. Для оценки качества алгоритмов на протяжении всего полёта дополнительно было посчитано среднее и отклонение метрики. Для первого сценария среднее и отклонение метрики представлены в Таблице 1.

Траекторию полета дронов для второго сценария можно увидеть на рисунке 6. Графики ошибок полёта для второго сценария показаны на рисунке 7. Для оценки качества алгоритмов на протяжении всего полёта дополнительно было посчитано среднее и отклонение метрики. Для второго сценария среднее и отклонение метрики представлены в Таблице 2.

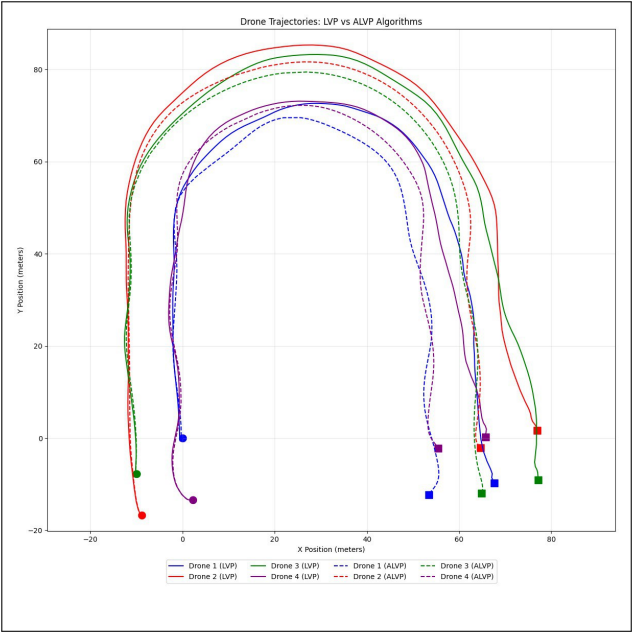


Рис. 4. Траектория полёта на первом сценарии

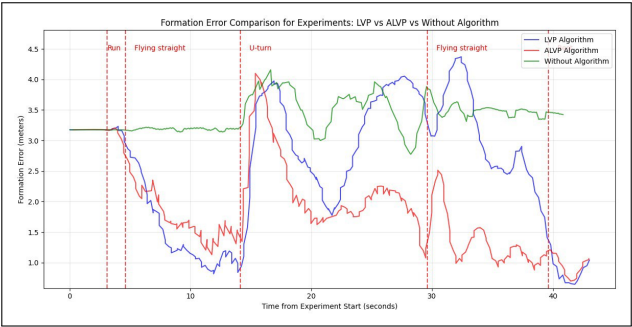


Рис. 5. Ошибка алгоритмов на первом сценарии

Таблица 1. Среднее и отклонение метрики для первого сценария

Алгоритм	Среднее	Отклонение
Без коррекции полёта	3.36 м	0.24 м
LVP	1.93 м	1.24 м
ALVP	1.31 м	0.79 м

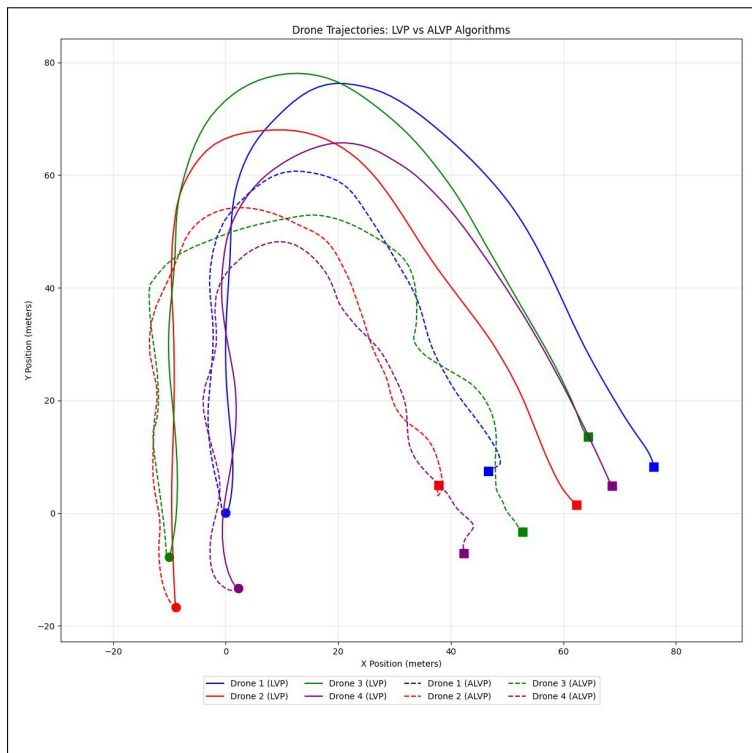


Рис. 6. Траектория полёта на втором сценарии

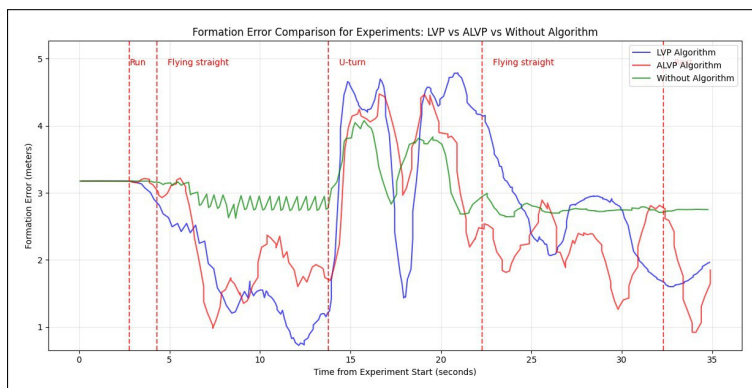


Рис. 7. Ошибка алгоритмов на втором сценарии

Таблица 2. Среднее и отклонение метрики для второго сценария

Алгоритм	Среднее	Отклонение
Без коррекции полёта	3.03 м	0.34 м
LVP	2.76 м	1.06 м
ALVP	2.62 м	0.86 м

Результаты экспериментов показывают, что во время движения по прямой система естественным образом стремится к идеальному построению. Однако при выполнении поворотов построение временно ухудшается, поскольку алгоритмы работают над его восстановлением. Примечательно, что система успешно восстанавливает свое построение после завершения поворота.

Результаты также демонстрируют, что оба алгоритма превосходят случай без использования коррекции положения дрона и протокол ускоренного локального голосования превосходит стандартный протокол локального голосования.

Для оценки воспроизводимости результата были проведены дополнительные эксперименты. Для угла поворота в 50 градусов в течение 15 секунд было произведено 20 полётов с алгоритмом LVP и 20 полётов с алгоритмом ALVP. С алгоритмом LVP только в 3-х полётах дроны оказались на финише в формации квадрата, в остальных случаях формация превращалась либо в зигзаг, либо в букву "Т". С алгоритмом ALVP в 17-ти полётах дроны успешно удержали формацию квадрата до финиша.

На рисунке 8 отображен разброс ошибок для 3-х экспериментов. На рисунке 9 отображены ошибки этих 3-х экспериментов. На рисунке 10 отображен разброс ошибок для 9-ти экспериментов. На рисунке 11 отображены ошибки этих 9-ти экспериментов.

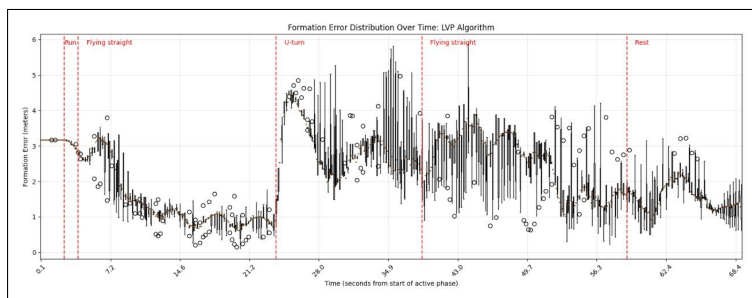


Рис. 8. Распределение ошибки для алгоритма LVP

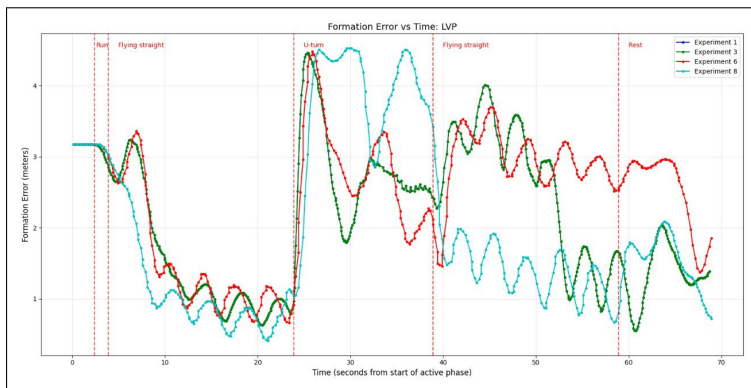


Рис. 9. Ошибки запусков с алгоритмом LVP

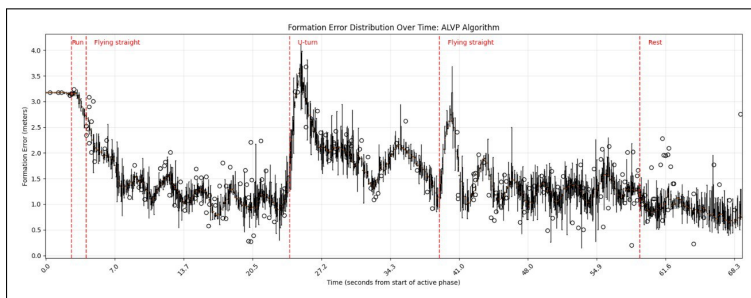


Рис. 10. Распределение ошибки для алгоритма ALVP

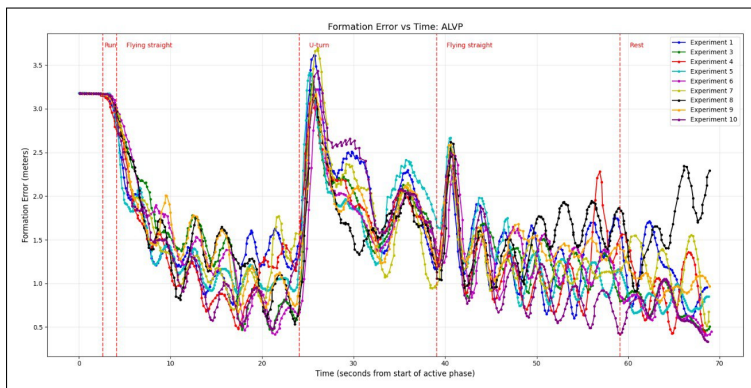


Рис. 11. Ошибки запусков с алгоритмом ALVP

В сценариях 3 и 4 оценивалась устойчивость алгоритмов к массовым потерям внутри группы. Для каждого уровня потерь (50% и 80%) и каждого алгоритма (LVP/ALVP) выполнено по 10 запусков; в каждом случае 5 запусков завершались восстановлением идеальной формации квадрата. Результаты экспериментов по сценарию 3 для LVP представлены на рисунке 12 и рисунке 13. Результаты экспериментов по сценарию 3 для ALVP представлены на рисунке 14 и рисунке 15. Результаты экспериментов по сценарию 4 для LVP представлены на рисунке 16 и рисунке 17. Результаты экспериментов по сценарию 4 для ALVP представлены на рисунке 18 и рисунке 19.

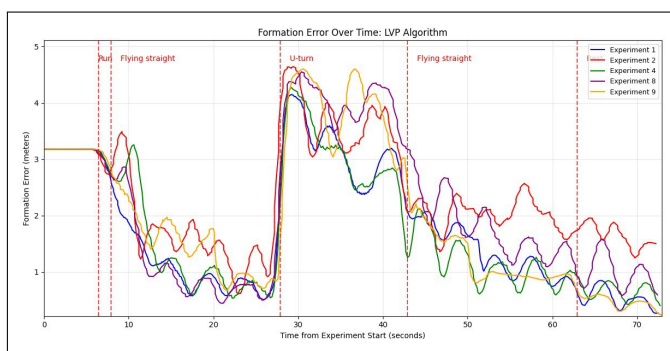


Рис. 12. Ошибки запусков с алгоритмом LVP с 50% потерями сообщений

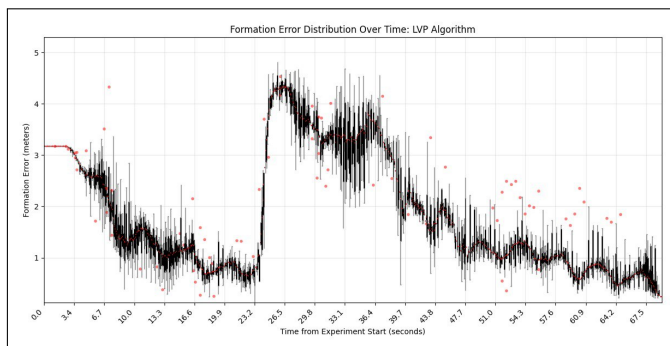


Рис. 13. Распределение ошибки для алгоритма LVP с 50% потерями сообщений

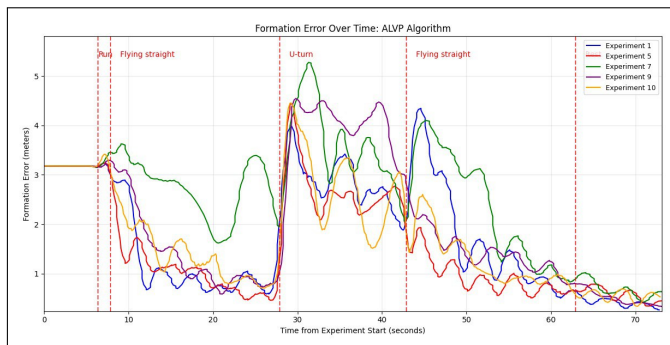


Рис. 14. Ошибки запусков с алгоритмом ALVP с 50% потерями сообщений

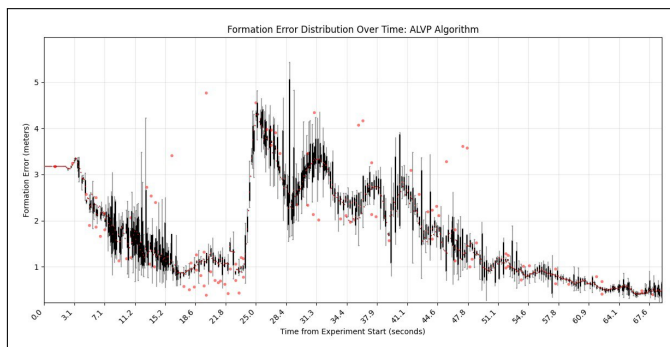


Рис. 15. Распределение ошибки для алгоритма ALVP с 50% потерями сообщений

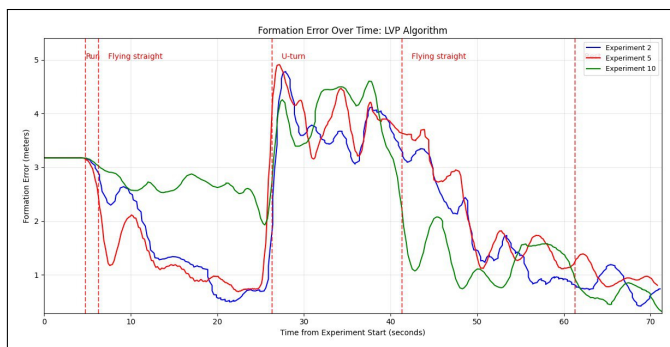


Рис. 16. Ошибки запусков с алгоритмом LVP с 80% потерями сообщений

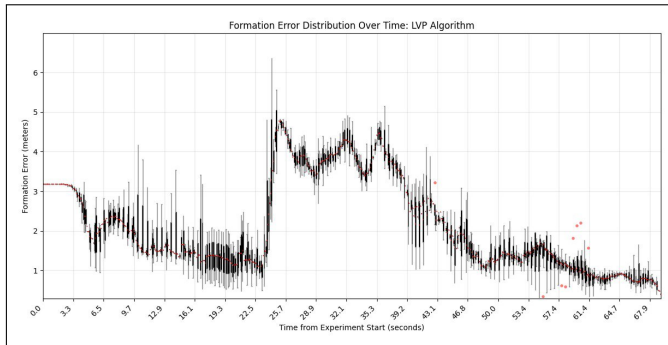


Рис. 17. Распределение ошибки для алгоритма LVP с 80% потерями сообщений

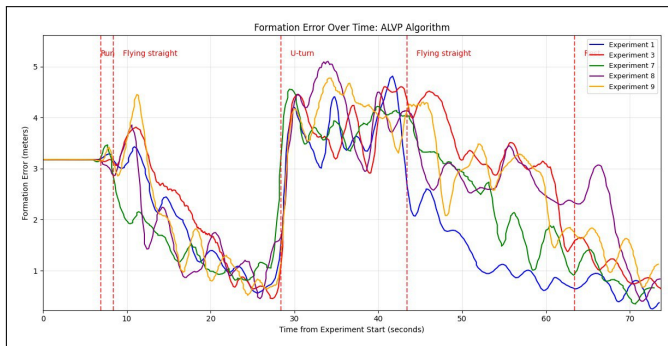


Рис. 18. Ошибки запусков с алгоритмом ALVP с 80% потерями сообщений

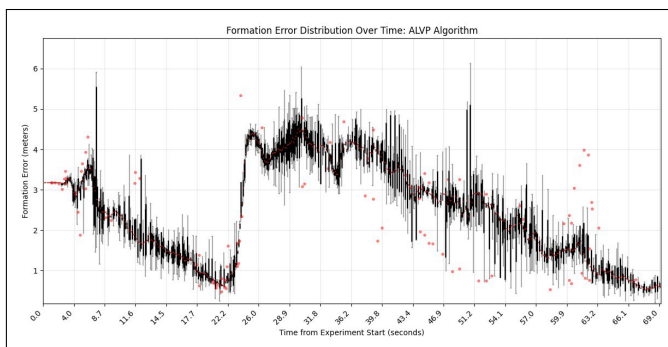


Рис. 19. Распределение ошибки для алгоритма ALVP с 80% потерями сообщений

В рамках пятого сценария экспериментов проводились запуски исключительно с использованием алгоритма ALVP с целью сравнить полученные значения ошибки формации с теоретическими предсказаниями согласно теореме 1. Данный сценарий предполагал восстановление формации дронов из произвольного начального состояния с нарушенным позиционированием без применения дополнительных маневров группы. Результаты экспериментов по данному сценарию, представленные на рисунке 20, демонстрируют хорошее соответствие между экспериментальными данными и теоретической моделью сходимости. Несмотря на наличие периодических колебаний в процессе восстановления формации, обусловленных особенностями работы алгоритма управления и динамикой системы, общая тенденция убывания ошибки полностью соответствует ожидаемому поведению согласно формуле $\varepsilon = C \cdot \lambda^K + D_K$.

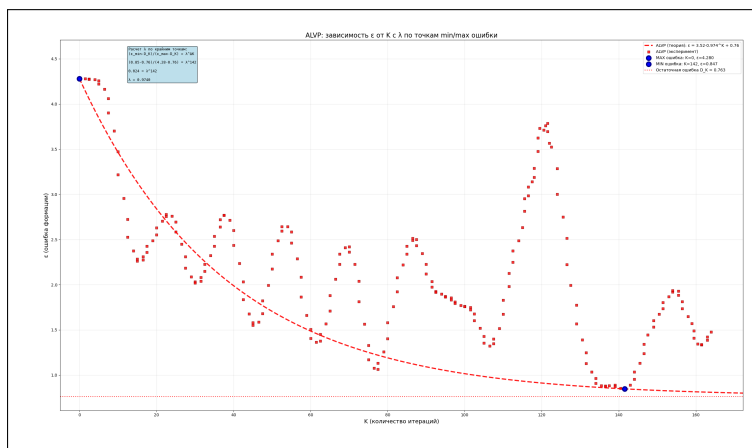


Рис. 20. Теоретический минимум и реальная ошибка ALVP на пятом сценарии

Проведённые множественные эксперименты показывают, что алгоритм ALVP работает стабильнее стандартного LVP.

Для оценки эффективности алгоритмов был использован специализированный симулятор [30]. Этот инструмент предназначен для воспроизведения динамики групп дронов, что позволяет всесторонне протестировать их поведение в различных условиях.

7. Заключение. В этой статье предлагается применение протокола локального голосования и его ускоренного варианта для поддержания формации групп роботов. Для тестирования алгоритмов

были разработаны архитектура системы и симулятор. Результаты экспериментов продемонстрировали эффективность предложенного подхода, показав, что ALVP значительно улучшает восстановление формации группы после резких маневров. ALVP продемонстрировал более высокую скорость сближения и устойчивость к динамическим изменениям окружающей среды по сравнению с LVP. В дальнейшем планируется провести натурные испытания предложенного подхода для повышения практической значимости результатов.

Возможным направлением исследований является добавление средств, позволяющих избежать препятствий. В [31] было предложено приложение LVP для обхода препятствий для группы колесных роботов, движущихся к источнику света. К предлагаемой динамике системы может быть добавлено дополнительное управляющее воздействие, позволяющее избегать препятствий всей группе при их обнаружении хотя бы одним из роботов.

Приложение: Доказательство выполнения условий теоремы 1 в системе управления роем роботов. Условия теоремы 1 ([29]): Для последовательности оценок $\{\bar{y}_t\}_{t=0}^{\infty}$, генерируемой протоколом ALVP, требуется:

1. Существование глобального оптимума $\bar{y}_t^*$ для функции:

$$R_t(\bar{y}) = \frac{(1 + \lambda)}{2} \bar{y}^T (I_m \otimes L(B(t))) \bar{y}.$$

2. Выполнение неравенства для каждой итерации t :

$$\mathbb{E}_t [R_t(\bar{y}_t) - R_t(\bar{y}_t^*)] \leq \lambda_t (\phi_0(\bar{y}_0) - R_t(\bar{y}_t^*) + \Phi) + D_t,$$

где λ_t характеризует скорость затухания ошибок, $D_t \geq 0$ – остаточная ошибка, причём $D_t < D_\infty < \infty$ и $D_t \rightarrow D_\infty$ при $t \rightarrow \infty$.

Доказательство для системы управления роем:

1. Применимость ALVP: В системе управления роем (раздел 5 статьи) ALVP используется для синхронизации состояний роботов: курса (h_i), высоты (z_i) и горизонтальных координат (x_i, y_i). Динамика i -го робота описывается уравнением:

$$X_{t+1}^i = A_t^i X_t^i + B_t^i w_t^i + w_{t+1}^i,$$

где w_t^i – возмущения (потери связи, шумы измерений, внешние воздействия). Глобальный вектор состояния:

$$\bar{y}(t) = [s_1(t)^T, \dots, s_n(t)^T]^T, \quad s_i(t) = [x_i(t), y_i(t), z_i(t), h_i(t)]^T.$$

Целевая функция имеет вид:

$$R_t(\bar{y}) = \frac{(1 + \lambda)}{2} \bar{y}^T (I_4 \otimes L(B(t))) \bar{y},$$

где I_4 – единичная матрица, $L(B(t))$ – матрица Лапласа графа связи.

2. Существование оптимума: Функция $R_t(\bar{y})$ является сильно выпуклой и гладкой (раздел 5). Матрица $I_m \otimes L(B(t))$ положительно полуопределена. При условии сильной связности графа (Предположение 1), $\lambda_{\min}(L(B(t))) > 0$, что обеспечивает существование и единственность глобального оптимума $\bar{y}_t^*$ – состояния консенсуса.

3. Выполнение неравенства (7): В [29] доказано, что ALVP удовлетворяет неравенству:

$$\mathbb{E}_t [R_t(\bar{y}_t) - R_t(\bar{y}_t^*)] \leq \lambda_t (\phi_0(\bar{y}_0) - R_t(\bar{y}_t^*) + \Phi) + D_t$$

для динамических сетей с неизвестными, но ограниченными возмущениями (раздел 5.2). В системе управления роем:

1. Возмущения w_t^i моделируют потери связи (до 80%), шумы измерений и внешние воздействия (разделы 4.3, 6).

2. Эти возмущения ограничены (Предположение 2 в [29]), что соответствует экспериментальным условиям (раздел 6).

Параметры λ_t и D_t :

1. $\lambda_t = \mathcal{O}(e^{-ct})$ (экспоненциальное затухание) доказано в [29] для сетей со стохастическими возмущениями.

2. D_t включает ошибку от возмущений w_t ; при $t \rightarrow \infty$ сходится к D_∞ , связанной с дисперсией шумов.

Заключение: Все условия теоремы 1 [29] выполняются в постановке задачи управления роем роботов:

1. ALVP генерирует последовательность $\{\bar{y}_t\}$ для динамики роботов.

2. Глобальный оптимум $\bar{y}_t^*$ существует и единственен при сильной связности графа.

3. Неравенство (7) выполняется благодаря ограниченности возмущений w_t и свойствам ALVP, доказанным в [29].

Таким образом, оценка ошибки ALVP (уравнения (8)–(10) теоремы 1) применима к системе управления роем роботов.

Литература

1. Bu Y., Yan Y., Yang Y. Advancement challenges in UAV swarm formation control: A comprehensive review // *Drones*. 2024. vol. 8. no. 7. DOI: 10.3390/drones8070320.
2. Amala Arokia Nathan R.J., Kurmi I., Bimber O. Drone swarm strategy for the detection and tracking of occluded targets in complex environments // *Communications Engineering*. 2023. vol. 2. DOI: 10.1038/s44172-023-00104-0.
3. Do H.T., Hua H.T., Nguyen M.T., Nguyen C.V., Nguyen H.T.T., Nguyen H.T., Nguyen N.T.T. Formation control algorithms for multiple-UAVs: a comprehensive survey // *EAI Endorsed Trans. Ind. Networks Intell. Syst.* 2021. vol. 8. no. 27.
4. Ren W. Consensus based formation control strategies for multi-vehicle systems // *Proceedings of the American Control Conference*. IEEE, 2006. pp. 6.
5. Amelin K., Amelina N., Granichin O., Granichina O., Andrievsky B. Randomized algorithm for UAVs group flight optimization // *IFAC Proceedings Volumes*. 2013. vol. 46. no. 11. pp. 205–208.
6. Amelina N., Fradkov A., Amelin K. Approximate consensus in multi-agent stochastic systems with switched topology and noise // *IEEE International Conference on Control Applications (CCA)*. 2012. pp. 445–450.
7. Erofeeva V., Granichin O., Volodina E. Accelerated decentralized load balancing in multi-agent networks // *IEEE Access*. 2024. vol. 12.
8. Vergados D.J., Amelina N., Jiang Y., Kravetska K., Granichin O. Toward optimal distributed node scheduling in a multihop wireless network through local voting // *IEEE Transactions on Wireless Communications*. 2017. vol. 17. no. 1. pp. 400–414.
9. Амелин К.С., Амелина Н.О., Граничин О.Н., Сергеев С.Ф. Децентрализованное групповое управление роем автономных роботов без маршрутизации данных // *Робототехника и техническая кибернетика*. 2021. Т. 9. № 1. С. 42.
10. Amelina N., Fradkov A., Jiang Y., Vergados D.J. Approximate consensus in stochastic networks with application to load balancing // *IEEE Transactions on Information Theory*. 2015. vol. 61. no. 4. pp. 1739–1752.
11. Olfati-Saber R., Murray R.M. Consensus problems in networks of agents with switching topology and time-delays // *IEEE Transactions on Automatic Control*. 2004. vol. 49. no. 9. pp. 1520–1533.
12. Nesterov Y. *Lectures on convex optimization*. Berlin: Springer, 2018. vol. 137. 589 p.
13. Kovalev D., Borodich E., Gasnikov A., Feoktistov D. Lower bounds and optimal algorithms for non-smooth convex decentralized optimization over time-varying networks. *arXiv preprint arXiv:2405.18031*. 2024.
14. Chen X., Huang L., Ding K., Dey S., Shi L. Privacy-preserving push-sum average consensus via state decomposition // *IEEE Transactions on Automatic Control*. 2023. vol. 68. no. 12. pp. 7974–7981.

15. Zhou Y., Cheng Y., Xu L., Chen E. Adaptive weighting push-SUM for decentralized optimization with statistical diversity // IEEE Transactions on Control of Network Systems. 2025. DOI: 10.1109/TCNS.2025.3566329.
16. Kenyeres M., Kenyeres J., Skorpil V. The analysis of the push-sum protocol in various distributed systems // European Scientific Journal. 2016. vol. 12. no. 12. DOI: 10.19044/esj.2016.v12n12p64.
17. Nedic A., Olshevsky A. Distributed optimization over time-varying directed graphs // IEEE Transactions on Automatic Control. 2014. vol. 60. no. 3. pp. 601–615. DOI: 10.1109/TAC.2014.2364096.
18. Mayne D.Q., Rawlings J.B., Diehl M.M. Model predictive control theory and design // Nob Hill Pub, Llc. 1999.
19. Kalman R.E. Contributions to the theory of optimal control // Bol. soc. mat. mexicana. 1960. vol. 5. no. 2. pp. 102–119.
20. Cohen A., Hasidim A., Koren T., Lazic N., Mansour Y., Talwar K. Online linear quadratic control // International Conference on Machine Learning. PMLR, 2018. vol. 80. pp. 1029–1038.
21. Goel G., Wierman A. An online algorithm for smoothed regression and LQR control // The 22nd International Conference on Artificial Intelligence and Statistics. PMLR, 2019. vol. 89. pp. 2504–2513.
22. The Dronecode Foundation. MAVLink. URL: <https://mavlink.io/en> (дата обращения: 25.06.2025).
23. M. Osborne and ArduPilot Dev Team. Mission planner. URL: <https://ardupilot.org/planner> (дата обращения: 25.06.2025).
24. ArduPilot Dev Team. (2024) ArduPilot. URL: <https://ardupilot.org> (дата обращения: 25.06.2025).
25. Espressif Systems. ESP32 Wi-Fi and Bluetooth SoC. URL: <https://www.espressif.com/en/products/socs/esp32> (дата обращения: 25.06.2025).
26. Espressif Systems. ESP-NOW wireless communication protocol. URL: <https://www.espressif.com/en/solutions/low-power-solutions/esp-now> (дата обращения: 25.06.2025).
27. Arkhipov I. MartOS. URL: <https://github.com/IvanArkhipov1999/Martos> (дата обращения: 25.06.2025).
28. Erofeeva V., Granichin O., Uzhva D. Meso-scale coalitional control in large-scale networks // Automatica. 2025. vol. 177.
29. Erofeeva, V., Granichin, O., Pankov, V., Volkovich, Z. Communication-efficient decentralized clustering for dynamical multi-agent systems // PlosOne. 2025.
30. Kiselev V. Swarm simulator. URL: <https://github.com/CroccoRush/swarm-simulator> (дата обращения: 25.06.2025).
31. Amelin K., Granichin O., Sergeenko A., Volkovich Z.V. Emergent intelligence via self-organization in a group of robotic devices // Mathematics. 2021. vol. 9. no. 12. pp. 1314. DOI: 10.3390/math9121314.

Архипов Иван Сергеевич — аспирант, ведущий разработчик программного обеспечения, Санкт-Петербургский государственный университет. Область научных интересов: математическое обеспечение и администрирование информационных систем. Число научных публикаций — 0. arkhipov.iv99@mail.ru; Университетская набережная, 7/9, 199034, Санкт-Петербург, Россия; р.т.: +7(812)328-2000.

Ерофеева Виктория Александровна — канд. техн. наук, ведущий специалист, центр ИИ и науки о данных, Санкт-Петербургский государственный университет. Область научных

интересов: стохастическая оптимизация, рандомизированные алгоритмы, адаптивное управление, идентификация систем, мультиагентные технологии, машинное обучение. Число научных публикаций — 20. vicki.ultramarine@gmail.com; Университетская набережная, 7/9, 199034, Санкт-Петербург, Россия; р.т.: +7(812)328-2000.

Граничин Олег Николаевич — д-р физ.-мат. наук, профессор кафедры, кафедра системного программирования математико-механического факультета, Санкт-Петербургский государственный университет; Институт проблем машиностроения Российской академии наук. Область научных интересов: многоагентное адаптивное управление, сжатие данных, кластеризация (интеллектуальный анализ данных), рандомизированные алгоритмы, анализ и проектирование сложных систем с неопределенностью, приложения в области информационных технологий. Число научных публикаций — 300. oleg_granichin@mail.ru; Университетская набережная, 7/9, 199034, Санкт-Петербург, Россия; р.т.: +7(812)328-2000.

Киселев Владимир Александрович — студент, Санкт-Петербургский государственный университет; разработчик программного обеспечения, ООО «Технологии Видеоанализа» и «Aezakmi Group». Область научных интересов: стохастическая оптимизация, рандомизированные алгоритмы, адаптивное управление, идентификация систем, мультиагентные технологии, машинное обучение. Число научных публикаций — 0. kiselyovvld@mail.ru; Университетская набережная, 7/9, 199034, Санкт-Петербург, Россия; р.т.: +7(812)328-2000.

Чернов Андрей Олегович — инженер-программист, центр ИИ и науки о данных, Санкт-Петербургский государственный университет; младший научный сотрудник, Институт проблем машиностроения Российской академии наук. Область научных интересов: стохастическая оптимизация, рандомизированные алгоритмы, адаптивное управление, идентификация систем, мультиагентные технологии, машинное обучение. Число научных публикаций — 10. a.o.chernov@mail.ru; Университетская набережная, 7/9, 199034, Санкт-Петербург, Россия; р.т.: +7(812)328-2000.

Поддержка исследований. Работа выполнена при финансовой поддержке СПбГУ, шифр проекта 121061000159-6.

I. ARKHIPOV, V. EROFEEVA, O. GRANICHIN, V. KISELEV, A. CHERNOV
**ACCELERATED LOCAL VOTING PROTOCOL FOR SINGLE
REMOTE CONTROLLER ROBOT SWARM SYSTEM**

Arkhipov I., Erofeeva V., Granichin O., Kiselev V., Chernov A. **Accelerated Local Voting Protocol for Single Remote Controller Robot Swarm System.**

Abstract. Controlling a robot swarm with a single remote controller is a challenging task, especially under unstable communication conditions where agents can temporarily lose the control signal, necessitating robust decentralized mechanisms for formation maintenance. This paper presents and tests a semi-centralized control system that enables an operator to coordinate the entire swarm as a unified entity. The system integrates centralized commands from a base station with decentralized position correction via the ESP-NOW protocol. To compare performance in maintaining a rigid formation, the Local Voting Protocol (LVP) and its Accelerated version (ALVP) were applied. Their effectiveness was evaluated in a simulation environment with a group of four drones through experiments involving sharp maneuvers (50° and 75° turns) and significant data packet loss simulations (50% and 80%). The results demonstrate that the Accelerated Local Voting Protocol (ALVP) offers significant advantages over the standard LVP, including faster formation recovery, lower mean positioning error, and greater stability. Specifically, in a series of 20 flight tests with a 50° turn, ALVP successfully maintained the formation in 17 cases, compared to only 3 for LVP, and also showed superior robustness under packet loss conditions. Therefore, the proposed semi-centralized approach using the ALVP protocol is an effective and robust solution for swarm formation control. Future work will focus on conducting physical experiments and integrating obstacle avoidance mechanisms.

Keywords: cooperative control, adaptive control, robotics.

References

1. Bu Y., Yan Y., Yang Y. Advancement challenges in UAV swarm formation control: A comprehensive review. *Drones*. 2024. vol. 8. no. 7. DOI: 10.3390/drones8070320.
2. Amala Arokia Nathan R.J., Kurmi I., Bimber O. Drone swarm strategy for the detection and tracking of occluded targets in complex environments. *Communications Engineering*. 2023. vol. 2. DOI: 10.1038/s44172-023-00104-0.
3. Do H.T., Hua H.T., Nguyen M.T., Nguyen C.V., Nguyen H.T.T., Nguyen H.T., Nguyen N.T.T. Formation control algorithms for multiple-UAVs: a comprehensive survey. *EAI Endorsed Trans. Ind. Networks Intell. Syst.* 2021. vol. 8. no. 27.
4. Ren W. Consensus based formation control strategies for multi-vehicle systems. *Proceedings of the American Control Conference. IEEE*, 2006. pp. 6.
5. Amelin K., Amelina N., Granichin O., Granichina O., Andrievsky B. Randomized algorithm for UAVs group flight optimization. *IFAC Proceedings Volumes*. 2013. vol. 46. no. 11. pp. 205–208.
6. Amelina N., Fradkov A., Amelin K. Approximate consensus in multi-agent stochastic systems with switched topology and noise. *IEEE International Conference on Control Applications (CCA)*. 2012. pp. 445–450.
7. Erofeeva V., Granichin O., Volodina E. Accelerated decentralized load balancing in multi-agent networks. *IEEE Access*. 2024. vol. 12.

8. Vergados D.J., Amelina N., Jiang Y., Kravetska K., Granichin O. Toward optimal distributed node scheduling in a multihop wireless network through local voting. *IEEE Transactions on Wireless Communications*. 2017. vol. 17. no. 1. pp. 400–414.
9. Amelin K.S., Amelina N.O., Granichin O.N., Sergeev S.F. [Decentralized group control of autonomous robots swarm without data Routing]. *Robototekhnika i tekhnicheskaja kibernetika – Robotics and Technical Cybernetics*. 2021. vol. 9. no. 1. (In Russ.).
10. Amelina N., Fradkov A., Jiang Y., Vergados D.J. Approximate consensus in stochastic networks with application to load balancing. *IEEE Transactions on Information Theory*. 2015. vol. 61. no. 4. pp. 1739–1752.
11. Olfati-Saber R., Murray R.M. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*. 2004. vol. 49. no. 9. pp. 1520–1533.
12. Nesterov Y. *Lectures on convex optimization*. Berlin: Springer, 2018. vol. 137. 589 p.
13. Kovalev D., Borodich E., Gasnikov A., Feoktistov D.. Lower bounds and optimal algorithms for non-smooth convex decentralized optimization over time-varying networks. *arXiv preprint arXiv:2405.18031*. 2024.
14. Chen X., Huang L., Ding K., Dey S., Shi L. Privacy-preserving push-sum average consensus via state decomposition. *IEEE Transactions on Automatic Control*. 2023. vol. 68. no. 12. pp. 7974–7981.
15. Zhou Y., Cheng Y., Xu L., Chen E. Adaptive weighting push-SUM for decentralized optimization with statistical diversity. *IEEE Transactions on Control of Network Systems*. 2025. DOI: 10.1109/TCNS.2025.3566329.
16. Kenyeres M., Kenyeres J., Skorpil V. The analysis of the push-sum protocol in various distributed systems. *European Scientific Journal*. 2016. vol. 12. no. 12. DOI: 10.19044/esj.2016.v12n12p64.
17. Nedic A., Olshevsky A. Distributed optimization over time-varying directed graphs. *IEEE Transactions on Automatic Control*. 2014. vol. 60. no. 3. pp. 601–615. DOI: 10.1109/TAC.2014.2364096.
18. Mayne D.Q., Rawlings J.B., Diehl M.M. *Model predictive control theory and design*. Nob Hill Pub, Llc. 1999.
19. Kalman R.E. Contributions to the theory of optimal control. *Bol. soc. mat. mexicana*. 1960. vol. 5. no. 2. pp. 102–119.
20. Cohen A., Hasidim A., Koren T., Lazic N., Mansour Y., Talwar K. Online linear quadratic control. *International Conference on Machine Learning*. PMLR, 2018. vol. 80. pp. 1029–1038.
21. Goel G., Wierman A. An online algorithm for smoothed regression and LQR control. *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR, 2019. vol. 89. pp. 2504–2513.
22. The Dronecode Foundation. MAVLink. Available at: <https://mavlink.io/en> (accessed 25.06.2025).
23. M. Osborne and ArduPilot Dev Team. Mission planner. Available at: <https://ardupilot.org/planner> (accessed 25.06.2025).
24. ArduPilot Dev Team. (2024) ArduPilot. Available at: <https://ardupilot.org> accessed 25.06.2025).
25. Espressif Systems. ESP32 Wi-Fi and Bluetooth SoC. Available at: <https://www.espressif.com/en/products/socs/esp32> (accessed 25.06.2025).
26. Espressif Systems. ESP-NOW wireless communication protocol. Available at: <https://www.espressif.com/en/solutions/low-power-solutions/esp-now> (accessed 25.06.2025).

27. Arkhipov I. MartOS. Available at: <https://github.com/IvanArkhipov1999/Martos> (accessed 25.06.2025).
28. Erofeeva V., Granichin O., Uzhva D. Meso-scale coalitional control in large-scale networks. *Automatica*. 2025. vol. 177.
29. Erofeeva, V., Granichin, O., Pankov, V., Volkovich, Z. Communication-efficient decentralized clustering for dynamical multi-agent systems. *PlosOne*. 2025.
30. Kiselev V. Swarm simulator. Available at: <https://github.com/CroccoRush/swarm-simulator> (accessed 25.06.2025).
31. Amelin K., Granichin O., Sergeenko A., Volkovich Z.V. Emergent intelligence via self-organization in a group of robotic devices. *Mathematics*. 2021. vol. 9. no. 12. pp. 1314. DOI: 10.3390/math9121314.

Arkhipov Ivan — Ph.D. student, leading software developer, St. Petersburg State University. Research interests: mathematical support and administration of information systems. The number of publications — 0. arkhipov.iv99@mail.ru; 7/9, University Emb., 199034, St. Petersburg, Russia; office phone: +7(812)328-2000.

Erofeeva Victoria — is a lead specialist of AI and Data Science Center of St. Petersburg State University. She earned her Specialist degree from Tambov State Technical University in 2013, followed by an MSc from St. Petersburg State University in 2015, and a PhD in Mathematical Cybernetics from the same institution in 2018. From 2016 to 2024, she was associated with the Institute of Problems of Mechanical Engineering of the Russian Academy of Sciences, conducting research in the Control of Complex Systems Laboratory. Her expertise encompasses stochastic optimization, randomized algorithms, adaptive control, system identification, multi-agent technologies, machine learning, and related areas.

Granichin Oleg — Ph.D., Dr.Sci., Professor of the department, System programming department, Faculty of mathematics and mechanics, St. Petersburg State University; Institute of Problems in Mechanical Engineering of the Russian Academy of Sciences. Research interests: multi-agent adaptive control, compressive sensing, clustering (data mining), randomized algorithms, analysis and design of complex systems with uncertainty, applications within information technology. The number of publications — 300. oleg_granichin@mail.ru; 7/9, University Emb., 199034, St. Petersburg, Russia; office phone: +7(812)328-200.

Kiselev Vladimir — Student, St. Petersburg State University; Software developer, «Technologies of Video Analysis» and «Aezakmi Group». Research interests: stochastic optimization, randomized algorithms, adaptive control, system identification, multi-agent technologies, machine learning. The number of publications — 0. kiselyovvld@mail.ru; 7/9, University Emb., 199034, St. Petersburg, Russia; office phone: +7(812)328-2000

Chernov Andrey — Software engineer, AI and data science center, St. Petersburg State University; Junior researcher, Institute of Problems in Mechanical Engineering of the Russian Academy of Sciences. Research interests: stochastic optimization, randomized algorithms, adaptive control, system identification, multi-agent technologies, machine learning. The number of publications — 10. a.o.chernov@mail.ru; 7/9, University Emb., 199034, St. Petersburg, Russia; office phone: +7(812)328-2000.

Acknowledgements. The research was carried out with the financial support of St. Petersburg State University, project No. 121061000159-6.

В.В. Вороной, Д.П. ВЕРХОВОД
**РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПОХОДКИ
ЧЕЛОВЕКА И СИНТЕЗ СИСТЕМЫ УПРАВЛЕНИЯ
ИСПОЛНИТЕЛЬНЫМИ МЕХАНИЗМАМИ МЕХАТРОННОГО
РЕАБИЛИТАЦИОННОГО КОМПЛЕКСА**

Вороной В.В., Верховод Д.П. Разработка математической модели походки человека и синтез системы управления исполнительными механизмами мехатронного реабилитационного комплекса.

Аннотация. В статье рассматривается разработка математической модели походки человека для синтеза системы управления мехатронного реабилитационного комплекса. Актуальность исследования обусловлена необходимостью создания эффективных реабилитационных технологий для пациентов с нарушениями двигательной функции. Существующие реабилитационные комплексы можно разделить на экзоскелеты и устройства с механическим соединением (end effector), при этом экзоскелеты демонстрируют более высокую реабилитационную эффективность, имитируя естественную походку. Исследование включает анализ различных моделей движения, таких как кинематические, динамические, биомеханические, нейронные сети и алгоритмы управления. Научная новизна работы заключается в создании модели, позволяющей учитывать индивидуальные антропометрические параметры пациента, включая массу и длины сегментов конечностей, а также моделировать поворот ступней. В рамках исследования предложен метод разбиения цикла шага на четыре фазы, каждая из которых описывается отдельной системой математических уравнений, что обеспечивает высокую точность воспроизведения различных этапов движения. Для валидации модели использовалась маркерная система захвата движений, что позволило получить данные о траекториях движения. Результаты показали, что модель эффективно генерирует траектории сагиттальных углов возвышения бедра, голени и стопы, что способствует улучшению управления реабилитационным устройством. В заключение, работа подчеркивает важность математического моделирования для создания адаптивных систем управления, которые могут существенно улучшить процесс реабилитации. Дальнейшие исследования будут направлены на доработку модели и её интеграцию с методами машинного обучения для повышения точности и надежности реабилитационных программ.

Ключевые слова: математическая модель, мехатронный реабилитационный комплекс, экзоскелет, кинематическая модель, параметрическая модель.

1. Введение. Роботизированный реабилитационный комплекс предназначен для проведения функциональной терапии нижних конечностей и вертикализации. Аппарат представляет собой массивный фиксированный стол, условно разделяемый на две половины: плоский стол, на который крепится торс пациента, и экзоскелет, на который крепятся ноги. Предполагается, что экзоскелет будет имитировать реальную походку, тем самым помогая пациенту вновь обрести навыки ходьбы. Подобные методы активно изучаются и применяются в рамках реабилитационных практик.

Текущие исследования в области биомеханики показывают, что она может предложить огромный неиспользованный потенциал в биомеханическом моделировании и симуляции. Математическое моделирование походки человека исходит из биомеханической системы, и соответствующая математическая модель используется для описания характеристик опорно-двигательного аппарата человека. Этот анализ предоставляет пользователям неограниченный доступ ко всем параметрам модели. В сочетании с анализом движений человека пользователи будут стремиться выявить конкретные отклонения, указав параметры модели, которые приводят к лежащей в основе патологии. В процессе принятия решений пользователи хотят оценить наиболее вероятный результат после конкретного терапевтического вмешательства и отрегулировать эти параметры таким образом, чтобы кодировать физиологические изменения, возникающие в результате конкретного вмешательства. Эффективность этого подхода зависит от достоверности моделирования двигательных систем и физиологических процессов человека, а также от того, как протекают сами процессы, порождающие движение.

Помимо поиска оптимального конструкторского и инженерного решения мехатронного комплекса, включающего электроприводы и актуаторы, воздействующие на нижние конечности человека так, что возможно их перемещение по траекториям, имитирующим ходьбу, необходимо также выбрать оптимальную систему управления исполнительными механизмами. Для чего нужно решить две основные задачи. Первое – это выбор математической модели походки человека, позволяющей наиболее точно описать траекторию движения ноги, а второе – выбор системы автоматического регулирования, позволяющей осуществлять точное и плавное движение исполнительных механизмов в совокупности, повторяющих заданную моделью траекторию движения.

Также целью работы является разработка такой системы, которая могла бы описать движение объекта в пространстве координат с плавными и непрерывными переходными процессами, а также обеспечивающую возможность моделирования движения с учетом произвольной временной «сетки». Что подразумевает не только учёт плавного перехода между заданными координатами, но и учёт вариативности скорости изменения координат в различные моменты времени.

2. Обзор основных видов реабилитационных комплексов.

Реабилитационные комплексы для людей, перенесших инсульт и пациентов с нарушением двигательной активности, можно разделить

на несколько категорий, учитывая специфические потребности этих групп пациентов. Первая группа – это активные реабилитационные комплексы (экзоскелеты, тренажеры), которые способствуют активному участию пациента в процессе восстановления двигательных функций [1, 2]. Вторая группа – пассивные комплексы (системы пассивной мобилизации, массажные аппараты), предназначены в основном для пациентов с ограниченной подвижностью и минимальным участием в активных упражнениях [3]. К третьей группе относят смешанные реабилитационные комплексы (роботизированные тренажеры, комплексы виртуальной реальности [4, 5]), которые комбинируют активные и пассивные методы, позволяя адаптировать реабилитацию под индивидуальные потребности пациента. И последняя группа – это информационные и аналитические системы [6] (системы захвата движений, программное обеспечение для анализа), которые акцентируют внимание на мониторинге и анализе данных о процессе реабилитации и позволяют адаптировать реабилитационные планы в соответствии с динамикой состояния пациента.

На данный момент существующие активные реабилитационные комплексы по принципу построения можно разделить на два основных вида. В первом случае концевые приводы (*end effector*) комплекса механически соединяются со ступнями человека [6], а во втором случае комплекс работает по принципу экзоскелета и жестко закреплен на бедрах и голени человека, обеспечивая перемещение конечностей [7, 8], помогая пациентам выполнять шаги и восстанавливать координацию.

Выбор конструктивных и программных решений, в том числе для применяемых приводов, определяется требованиями к их кинематическим и динамическим параметрам, т.е. реализуемым траекториям движения конечностей, располагаемым скоростям и ускорениям. Данные по этим параметрам неоднократно изучены и будут представлены ниже.

Существующие конструкции реабилитационных комплексов для людей, перенесших инсульт и пациентов с нарушением двигательной активности, различаются по уровню сложности и типам упражнений, форме интерфейса, а также габаритами, требованиями к электропитанию, массой и возможностью транспортировки и использования в различных условиях, что немаловажно при эксплуатации комплекса в медицинских учреждениях. Основные сравнительные характеристики представлены в таблице 1.

Таблица 1. Конструктивные характеристики реабилитационных комплексов

Критерий	Экзоскелет	End effector
Габариты комплекса	высота 1.5-1.8 м ширина 0.5-0.7 м	высота 1.5-2 м длина/ширина платформы 2-3 м/1.5-2 м
Масса	30-40 кг	200-500 кг
Потребление электроэнергии	3-5 кВт	6-8 кВт
Удерживание ноги	обеспечивается	Обеспечивается при использовании ортезов
Мобильность комплекса	да	нет
Сложность реализации	высокая	средняя

Экзоскелетная модель демонстрирует высокую реабилитационную эффективность благодаря меньшим массо-габаритным параметрам, низкому энергопотреблению и мобильности. Кроме того, возможность имитации естественной походки на основе математической модели позволяет пациентам, перенесшим травмы или инсульт, быстрее восстанавливать утраченные навыки ходьбы. Подобные методы активно изучаются и применяются на практике [10, 11]. Однако, данные комплексы обладают более высокой сложностью разработки, как программной части, так и механической, что может повлечь за собой и увеличение вероятности инженерных ошибок и снижение надежности.

3. Разработка математической модели походки человека.

В настоящее время существует множество вариантов моделирования походки как робота, так и человека. Большинство существующих моделей разработаны для различной мультипликации и не подходят для применения в реабилитационных комплексах, т.к. траектории движения не настолько гладкие, а при практической реализации в реабилитационных комплексах важны траектории, описываемые математическими уравнениями хотя бы до второй производной. Кроме того, данные модели ориентированы на динамических персонажей, а в нашем случае пациенты с заболеваниями опорно-двигательного аппарата обладают малой подвижностью и высокая плавность движений очень важна в реабилитационных целях. Также существующие модели не предоставляются в открытом доступе, а дорогостоящие лицензии на их применение ограничены для применения в медицинских целях.

В реабилитационных роботизированных комплексах для восстановления двигательной активности после инсульта обычно используются несколько типов моделей движения как робота, так и человека:

1. Модели на основе кинематики [12]. Эти модели описывают движение суставов и конечностей человека. Они могут использоваться для анализа и синтеза движений, а также для создания алгоритмов управления роботизированными системами.

2. Модели на основе динамики [13]. Динамические модели учитывают силы и моменты, действующие на тело человека, и могут использоваться для симуляции походки и разработки контроллеров для реабилитационных устройств.

3. Модели на основе биомеханики [14]. Эти модели исследуют механические аспекты движения человека, включая распределение нагрузки на суставы и мышцы. Они помогают понять, как различные факторы влияют на походку и могут быть использованы для адаптации реабилитационных программ.

4. Модели на основе нейронных сетей [15, 16]. Современные подходы могут включать использование нейронных сетей для анализа и воспроизведения паттернов движения. Эти модели могут адаптироваться к индивидуальным потребностям пациента.

5. Модели, основанные на алгоритмах управления [17]. Эти модели используются для управления роботизированными экзоскелетами или другими устройствами, обеспечивая плавное и естественное движение, имитирующее походку человека.

Каждый из этих типов моделей может быть использован в комбинации с другими для создания эффективных реабилитационных программ и устройств.

В рамках разработки реабилитационного комплекса была выбрана модель, описанная в статьях [18, 19], поскольку она позволяет задать множество антропологических параметров, а также моделирует поворот ступней. В выбранной модели ноги человека представлены в виде соединённых между собой отрезков (рисунок 1). Голова, руки и туловище, для удобства, представляются в виде единого целого. Предполагается, что человек не машет руками при ходьбе слишком сильно [20, 21]. Такая модель обеспечивает баланс между вычислительной эффективностью и биомеханической достоверностью. Упрощение верхней части тела допустимо, так как реабилитация сосредоточена на восстановлении паттернов ходьбы, где ключевую роль играют движения ног и таза [22]. Кроме того, подобный подход широко

применяется в роботизированной реабилитации, например, в экзоскелетах EksoNR и ReWalk, где акцент также делается на нижние конечности.

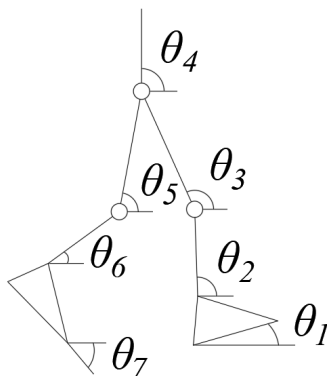


Рис. 1. Визуальное представление модели ног

Цикл шага состоит из четырёх фаз:

1) От касания пола правой пяткой (*RHS*, от англ. *right-heel-strike*) до отрыва левого носка (*LTO*, от англ. *left-toe-off*). Правая пятка и левый носок опираются на пол.

2) От *LTO* до касания правым носком (*RFF*, от англ. *right-foot-flat*). Только правая нога опирается на пол.

3) *RFF*. Правая нога не двигается, оставшиеся 6 сегментов опираются на правую ногу.

4) От отрыва правой пятки (*RHO*, от англ. *right-heel-off*) до касания левой пятки. Только правый носок опирается на пол.

Разные фазы шага описываются двумя различными математическими моделями. Первая модель описывает фазу 1, в которой 6 переменных, описывающих ноги, зависимы друг от друга. Система уравнений, описывающая подобную ситуацию, была бы чрезвычайно сложна, поэтому было принято решение считать, что правое бедро соединяется с тазом посредством жёсткой пружины с демпфером. Фазы 2, 3 и 4 существенно отличаются от фазы 1, но между собой отличаются лишь немного. Положение ног описывается с помощью углов относительно уровня пола.

Чтобы сформулировать уравнение, опишем Лагранжиан L :

$$L = T - V, \quad (1)$$

где T и V – кинетическая и потенциальная энергии, соответственно. Энергии описываются как функции независимых переменных θ_i , указанных на рисунке 1. Полученные уравнения движения описываются как:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}_i} \right) - \frac{\partial L}{\partial \theta_i} = Q_i, \quad i = 1, \dots, 7, \quad (2)$$

где Q_i – виртуальная работа, включающая в себя эффекты перемещения ног и демпфера в районе таза во время первой фазы шага. Уравнение, описывающее первую фазу шага:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} & a_{45} & a_{46} & a_{47} \\ 0 & 0 & 0 & a_{54} & a_{55} & a_{56} & a_{57} \\ 0 & 0 & 0 & a_{64} & a_{65} & a_{66} & a_{67} \\ 0 & 0 & 0 & a_{74} & a_{75} & a_{76} & a_{77} \end{bmatrix} \begin{bmatrix} \ddots \\ \theta_1 \\ \ddots \\ \theta_2 \\ \ddots \\ \theta_3 \\ \ddots \\ \theta_4 \\ \ddots \\ \theta_5 \\ \ddots \\ \theta_6 \\ \ddots \\ \theta_7 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix}, \quad (3)$$

где $a_{ij} = c_{ij} \cos(\theta_i - \theta_j)$. Значения c_{ij} симметричны, то есть $c_{ij} = c_{ji}$, а значит a_{ij} тоже симметричны [23].

Стоит отметить, что первые три уравнения системы (3) кажутся несвязанными друг с другом из-за того, что правая нога отделена от остального тела пружиной. Связь через пружину отражена в b_i , включающей в себя воздействие сил, связанных с ней.

Для фаз 2 и 4, где одна нога движется, опираясь на другую ногу, уравнение движения выглядит так.

Значения a_{ij} и b_i в данном случае отличны от аналогичных в уравнении первой фазы. Фаза 3 определяется таким же уравнением, но из-за того, что опорная нога стоит на полу, первый столбец и первая строка матрицы a в ней отсутствуют.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ a_{41} & a_{42} & a_{43} & a_{44} & 0 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & 0 & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & 0 & a_{65} & a_{66} & a_{67} \\ a_{71} & a_{72} & a_{73} & 0 & a_{75} & a_{76} & a_{77} \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{\theta}_3 \\ \ddot{\theta}_4 \\ \ddot{\theta}_5 \\ \ddot{\theta}_6 \\ \ddot{\theta}_7 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix}. \quad (4)$$

Значения c_{ij} и b_i для всех фаз зависят от углов θ и антропологических параметров пациента (массы тела, длин разных частей ног и так далее).

Значения c_{ij} для первой фазы:

$$\begin{aligned} c_{11} &= m_f C^2 + I_f + (m_s + m_t) B^2, \\ c_{12} &= m_s DB + m_t (D + E), \\ c_{13} &= m_t FB, \\ c_{22} &= I_s + m_s D^2 + m_t (D + E)^2, \\ c_{23} &= m_t F(D + E), \\ c_{33} &= I_t + m_t F^2, \end{aligned} \quad (5)$$

$$\begin{aligned}
c_{44} &= I_h + m_h H^2, \\
c_{45} &= m_h H(F + G), \\
c_{46} &= m_h H(D + E), \\
c_{47} &= 2m_h HA, \\
c_{55} &= I_t + m_t F^2 + m_h (F + G)^2 \\
c_{56} &= m_t F(D + E) + m_h (D + E)(F + G), \\
c_{57} &= 2m_t FA + 2m_h A(F + G), \\
c_{66} &= I_s + m_s D^2 + (m_h + m_t)(D + E)^2, \\
c_{67} &= 2(m_h + m_t)(D + E)A + 2m_s AD, \\
c_{77} &= I_f + (0.25m_f + m_h + m_t + m_s)4A^2.
\end{aligned} \tag{5}$$

где A – расстояние от лодыжки до центра масс, B – расстояние от пятки до лодыжки, C – расстояние от пятки до центра масс, m_f – масса стопы, I_f – момент инерции стопы относительно ее центра масс, D – расстояние от лодыжки до центра масс голени, E – расстояние от центра масс голени до колена, m_s – масса голени, I_s – момент вращения голени относительно ее центра масс, F – расстояние от колена до бедра, G – расстояние от центра масс бедра до таза, m_t – масса бедра, I_t – момент инерции бедра относительно центра масс, H – расстояние от таза до центра масс верхней половины тела, m_h – масса верхней половины тела, I_h – момент инерции верхней половины тела вокруг его центра масс, f – коэффициент затухания демпфера в области таза, g – гравитационная постоянная.

Значения c_{ij} для фаз 2, 3, 4:

$$\begin{aligned}
 c_{11} &= K^2(2m_s + 2m_t + m_h + m_f) + m_f Z^2 + I_f, \\
 c_{12} &= K(m_s D + (2m_t + m_h + m_s + m_f)(D + E)), \\
 c_{13} &= K(m_t F + (m_h + m_t + m_s + m_f)(F + G)), \\
 c_{14} &= K m_h H, \\
 c_{15} &= -K(m_t G + (m_s + m_f)(F + G)), \\
 c_{16} &= -K(m_s E + m_f(D + E)), \\
 c_{17} &= -K m_f A, \\
 c_{22} &= I_s + m_s D^2 + (m_s + 2m_t + m_h + m_f)(D + E)^2, \\
 c_{23} &= (m_t F + (m_h + m_t + m_s + m_f)(F + G))(D + E), \\
 c_{24} &= m_h H(D + E), \\
 c_{25} &= -(m_t G + (m_s + m_f)(F + G))(D + E), \\
 c_{26} &= -(m_s E + m_f(D + E))(D + E), \\
 c_{27} &= -m_f A(D + E), \\
 c_{33} &= I_t + m_t F^2 + (m_h + m_t + m_s + m_f)(F + G)^2, \\
 c_{34} &= m_h H(F + G), \\
 c_{35} &= -(m_t G + (m_s + m_f)(F + G))(F + G), \\
 c_{36} &= -(m_s E + m_f(D + E))(F + G), \\
 c_{37} &= -m_f A(F + G), \\
 c_{44} &= I_h + m_h H^2, \\
 c_{45} &= c_{46} = c_{47} = 0, \\
 c_{55} &= I_s + m_s G^2 + (m_s + m_f)(F + G)^2, \\
 c_{56} &= (m_s E + m_f(D + E))(F + G), \\
 c_{57} &= m_f A(F + G), \\
 c_{66} &= I_s + m_s E^2 + m_f(D + E)^2, \\
 c_{67} &= m_f A(D + E), \\
 c_{77} &= I_f + m_f A^2.
 \end{aligned} \tag{6}$$

K принимает значения: B для фазы 2; 0 для фазы 3; $2A$ для фазы 4. Z принимает значения: C для фазы 2; 0 для фазы 3; A для фазы 4.

Значения c_{ij} , не указанные выше равны нулю.

Определим значения d_j , необходимые для описания выражений b_i :

$$d_1 = B; d_2 = D + E; d_3 = F + G; d_4 = 0; d_5 = -(F + G); \\ d_6 = -(D + E); d_7 = -2A.$$

Значения b_i задаются как:

$$b_i = - \sum_{j=1}^7 c_{ij} \dot{\theta}_i^2 \sin(\theta_i - \theta_j) + \\ + S_k \left[\sum_{j=1}^7 d_i d_j \sin(\theta_i - \theta_j) + d_i S \sin \theta_i \right] + R_i, \quad (7)$$

где S – изначальное расстояние от левого носка до правой пятки,

$$R_1 = -m_f g C \cos(\theta_1 - \beta + \gamma) - (m_s + m_t) g B \cos \theta_1 + M_{ra} + D_1, \\ R_2 = -C_{21} g \cos \theta_2 + M_{rk} - M_{ra} + D_2, \\ R_3 = -C_{31} g \cos \theta_3 + M_{rh} - M_{rk} + D_3, \\ R_4 = -m_h H g \cos \theta_4 - M_{rh} - M_{lh} + D_4, \\ R_5 = -[m_h (F + G) + m_t F] g \cos \theta_5 + M_{lh} - M_{lk} + D_5, \\ R_6 = -[(m_h + m_t)(D + E) + M_s D] g \cos \theta_6 + M_{lk} - M_{la} + D_6, \\ R_7 = -[2A(m_h + m_t + m_s) + A m_f] g \cos \theta_7 + M_{la} + D_7, \quad (8)$$

где D_i – моменты, возникшие в результате работы демпфера, g – гравитационная постоянная. Чтобы найти эти моменты, необходимо вычислить виртуальную работу демпфера. Выразив эту работу через $\delta\theta_i$, можно увидеть данные коэффициенты:

$$\delta W_D = D_1 \delta\theta_1 + D_2 \delta\theta_2 + \dots + D_7 \delta\theta_7. \quad (9)$$

Чтобы найти виртуальную работу, перейдём в Декартову систему координат. Сила демпфера пропорциональна относительной скорости верхних точек левого и правого бёдер. Относительно левого бедра, сила демпфера F_D равна:

$$F_D = -f(v_R - v_L), \quad (10)$$

где v_R и v_L – скорости верхних точек правого и левого бедра соответственно. Переписав выражения для отдельных компонентов, получим: $F_{DX} = -fv_x$, $F_{DY} = -fv_y$, где v_x и v_y – компоненты искомой относительной скорости. Заметим, что F_{DX} и F_{DY} могут быть найдены с помощью перевода угловых скоростей в декартову систему координат.

Уравнение виртуальной работы в декартовой системе координат выглядит таким образом:

$$\delta W_D = F_{DX}\delta x + F_{DY}\delta y. \quad (11)$$

Выразив δx и δy через угловые скорости и подставив представление δW_D в терминах угловых скоростей получим выражения для искоемых моментов:

$$D_i = d_i (-F_{DX}\sin\theta_i + F_{DY}\cos\theta_i). \quad (12)$$

Для фаз 2-4, значения b_i описываются следующим образом:

$$b_i = - \sum_{j=1}^7 c_{ij} \dot{\theta}_i^2 \sin(\theta_i - \theta_j) + R_i. \quad (13)$$

Здесь для фазы 2:

$$R_1 = -Bgm_f \cos(\theta_1 - \beta + \gamma) - (2m_s + 2m_t + m_h + m_f) \cos\theta_1 + M_r, \quad (14)$$

где β – угол между подошвой и стопой, измеренный на пятке, γ – угол между подошвой и центром массы, измеренный на пятке.

Для фазы 3 $R_1 = 0$, а для фазы 4:

$$R_1 = -2Agm_f \cos(\theta_1 - \beta + \gamma) - (2m_s + 2m_t + m_h + m_f) \cos\theta_1 + M_{ra}. \quad (15)$$

Для всех фаз:

$$\begin{aligned}
 R_2 &= -c_{12}\cos\theta_2 + M_{rk} - M_{ra}, \\
 R_3 &= -c_{13}g\cos\theta_3 + M_{rh} - M_{rk}, \\
 R_4 &= -c_{14}g\cos\theta_4 - M_{rh} - M_{lh}, \\
 R_5 &= -c_{15}g\cos\theta_5 + M_{lh} - M_{lk}, \\
 R_6 &= -c_{16}g\cos\theta_6 + M_{lk} - M_{la}, \\
 R_7 &= -c_{17}g\cos\theta_7 + M_l.
 \end{aligned} \tag{16}$$

Для решения сформулированной задачи необходимо знать изначальные положения ног, а также моменты вращения в суставах. Изначальные значения были подобраны вручную, а для определения моментов использовались данные из [18]. Использовался следующий метод решения:

- имея начальные значения углов θ_i , находим $a_{ij} \forall i, j$;
- имея начальные значения углов θ_i и угловых скоростей $\dot{\theta}_i$, находим $b_i \forall i$;
- имея значения a_{ij} и b_{ij} в начальный момент, находим угловые ускорения с помощью уравнений 1 и 2;
- находим угол и угловую скорость спустя время h по схеме Эйлера:

$$\theta_i(t+h) = \theta_i(t) + h(t). \tag{17}$$

Модель была реализована на языке *Python* с помощью библиотеки *numpy*. Для визуализации результата была написана утилита с помощью движка *Godot*. В качестве результатов работы были получены данные о положении ног в процессе шага (рисунок 2). Также, была разработана визуализация процесса ходьбы для выявления недостатков в выбранной модели. В дальнейшем модель будет доработана, будет проведено сравнение с моделью, основанной на машинном обучении.

4. Параметрическая модель. Для построения параметрической модели походки человека, которая позволит генерировать траектории сагиттальных углов возвышения бедра, голени и стопы каждой нижней конечности, необходим набор данных, полученных в результате съёмки походки человека с помощью системы захвата движений. Для

этого в рамках данного проекта разработана маркерная система захвата движений.

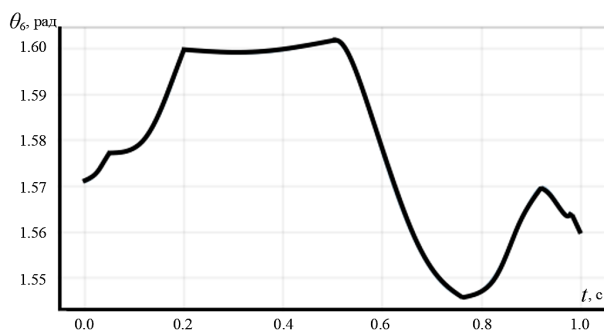


Рис. 2. Изменение θ_6 в процессе шага

Маркерная система захвата движений состоит из трех основных компонентов. Стереокамера, осуществляющая захват позиций маркеров, маркеры, наносимые на человека и хорошо фиксируемые камерами, траекторию движения которого захватывается и программное обеспечение, обеспечивающее функционал калибровки стерео-видеосистемы захвата, сохранения и визуализации стерео-видео и полу автоматизированной разметки снятых данных.

На первом этапе для каждой камеры калибруются параметры матрицы внутренних параметров модели камеры-обскуры и параметры модели искажений объектива «рыбий глаз». Модель камеры-обскуры:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R & T \end{bmatrix} p_\omega = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix}_{4 \times 4} \begin{bmatrix} x_\omega \\ y_\omega \\ z_\omega \\ 1 \end{bmatrix}, \quad (18)$$

где (u, v) – координаты точки в пространстве изображения; K – матрица внутренних параметров камеры; f_x, f_y – фокальные расстояния, выраженные в пикселях; c_x, c_y – координаты центра изображения; $\begin{bmatrix} R & T \end{bmatrix}$ – матрица внешних параметров, определяющая позицию и ориентации камеры в мировых координатах; $(x_\omega, y_\omega, z_\omega)$ – мировые координаты проецируемой точки.

Модель искажений объектива добавляется к модели камеры-обскура как трансформация вектора $p_a = [R \ T] p_\omega = [x_a \ y_a \ z_a \ 1]^T$ следующим образом:

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} x_a / z_a \\ y_a / z_a \end{bmatrix}, \quad (19)$$

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = \frac{\theta}{r} \left[1 + k_1 \theta^2 + k_2 \theta^4 + k_3 \theta^6 + k_4 \theta^8 \right] \begin{bmatrix} x_n \\ y_n \end{bmatrix},$$

где $r = \sqrt{x_n^2 + y_n^2}$ и $\theta = \tan^{-1}(r)$, k_1, k_2, k_3, k_4 , – калибруемые параметры.

Параметры $f_x, f_y, c_x, c_y, k_1, k_2, k_3, k_4$ восстанавливаются в результате процесса оптимизации по набору изображений заранее известного паттерна (шахматной доски), снятых камерой с разных ракурсов.

Взаимное расположение камер, матрицы внешней калибровки для первой и второй камеры соответственно, находится в результате процесса минимизации ошибки триангуляции точек по набору изображений шахматной доски, снятых закрепленными между собой камерами с разных ракурсов.

Перед фактической съемкой траекторий камеры устанавливаются на штатив и по нескольким специальным образом нанесенным в некоторые точки окружения опорным маркерам производится привязка позиций камер в мировых координатах помещения, а также выравнивание осей координат так, чтобы ось x соответствовала направлению движению, ось y была направлена вверх, ноль на уровне пола, ось z перпендикулярна двум предыдущим осям и направлена в сторону от камер.

Маркеры закрепляются на человеке так, чтобы была возможность восстановить позицию каждого интересующего нас сустава с некоторой избыточностью в данных. Сами маркеры должны удовлетворять нескольким требованиям: должны быть яркие, контрастные, хорошо различимые со своим локальным фоном, при этом, желательно, быть простыми/доступными, легкими в закреплении. В нашем случае маркерами являются вырезанные в форме круга диаметром 1 см красно-оранжевые наклейки, обладающие хорошим свет возвращающим покрытием.

Каждый записываемый участник проходит несколько раз по намеченной прямой траектории, каждый раз с разной, но постоянной скоростью. Одна запись соответствует одному проходу, длительностью 5-9 секунд, включает 1.5-3 полных цикла шага.

Алгоритм полуавтоматизированной разметки принимает на вход последовательность пар стереоизображений, ищет на них потенциальные позиции маркеров. Пользователь выбирает конкретную позицию каждого маркера на каждом изображении стереопары. Для уже выбранного однажды маркера алгоритм пытается, основываясь на предыдущих наблюдениях, предсказать позицию на последующих кадрах. По предсказанной позиции, алгоритмом выбирается ближайшее похожее на маркер пятно.

При перекрывании маркера отслеживание теряется. Пользователь должен заново выделить маркер, когда тот вернется в поле зрения.

Каждая пара маркеров преобразуется в 3D точку на основе триангуляции с откалиброванной пары камер. Наборы 3D точек для каждой пары изображений стерео-видео для каждого маркера составляют целевые захватываемые траектории движения.

На вход алгоритма нахождения потенциальных позиций маркеров поступает изображение формата RGB, на котором нужно найти потенциальные позиции маркеров. Первым шагом изображение переводится в gray-scale формат следующим образом. За основу берется R канал, на него накладывается штраф, уменьшающий интенсивность, зависящий от величины G и B каналов данного пикселя. Величина штрафа является параметром алгоритма и настраивается пользователем. Затем gray-scale изображение бинаризуется по порогу, где порог также является параметром, настраиваемым пользователем. Затем на бинаризованном изображении находятся связные компоненты (непрерывные белые пятна). Для каждой связной компоненты находится ее центр как взвешенная средняя позиций пикселей связной компоненты, где вес позиции – интенсивность пикселя на gray-scale изображении. Центр каждой связной компоненты – центр потенциального маркера.

5. Синтез системы управления исполнительными механизмами. Полученные данные траекторий движения используются в качестве задающего сигнала в системе подчиненного регулирования приводами исполнительных механизмов с применением ПД-регуляторов. Подробное описание реализации системы управления представлено в [17].

На этапе проектирования была выбрана модель подчиненного управления с целью управления не только координатами положения, но и скоростью и моментом на валу двигателей исполнительных механизмов. Общая структурная схема системы представлена на рисунке 3.

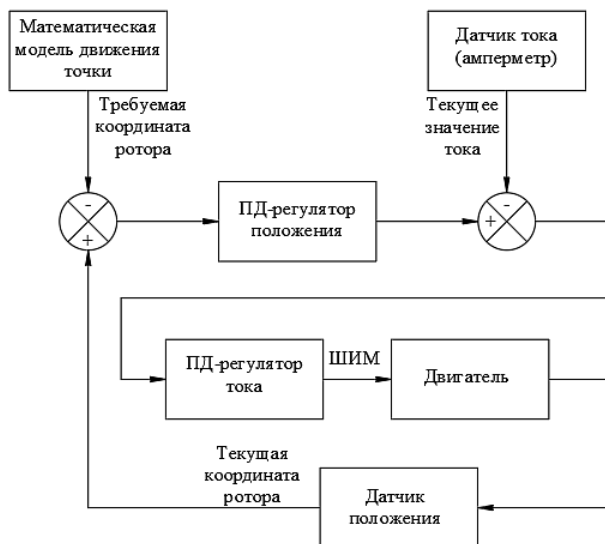


Рис. 3. Структурная схема системы управления

Общее описание функционирования системы: математическая модель походки человека вычисляет актуальную координату в данный момент времени, которая затем передается в ПД-регулятор для расчета необходимого тока. После этого выходное значение по положению от ПД-регулятора поступает на вход ПД-регулятора по току. В результате работы регулятора формируется напряжение, которое будет подано на обмотки фаз, обеспечивая необходимое управление вращением ротора.

На рисунках 4, 5 представлены результаты работы системы управления, полученные в ходе тестирования экспериментального образца мехатронного реабилитационного комплекса. Эксперимент проводился при различных скоростях шага. На рисунке 5 скорость шага соответствует средней скорости человека, равной 1.4-1.5 м/с, в то время как на рисунке 6 скорость шага снижена вдвое. Это связано с тем, что на начальном этапе реабилитации пациентов может потребоваться начинать с более низких скоростей движения и меньшей высоты подъема ног.

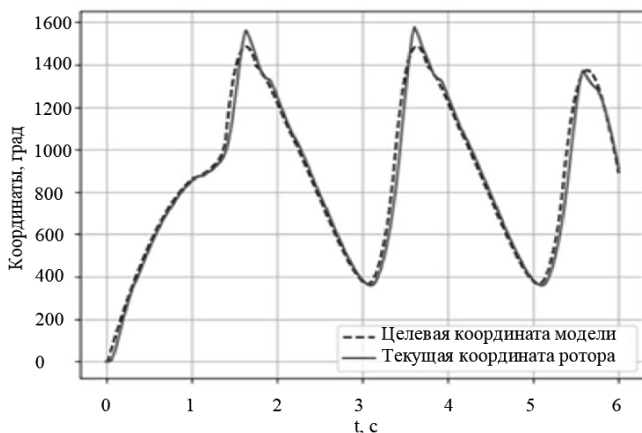


Рис. 4. Координаты при средней скорости шага человека 1.4-1.5 м/с

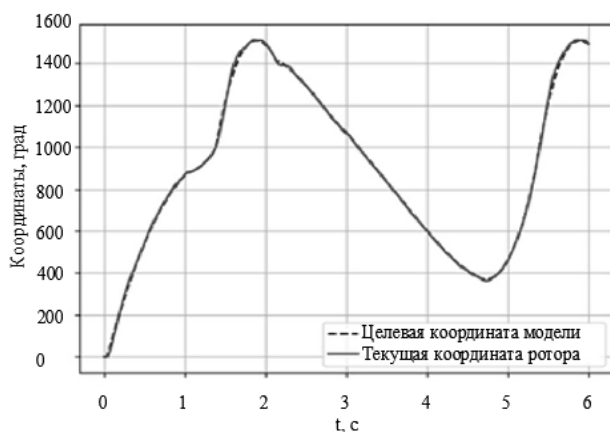


Рис. 5. Координаты при средней скорости шага человека 0.7-0.8 м/с

6. Разработка механической конструкции реабилитационного аппарата. Конструкторское решение комплекса включает в себя каркас, поворотную опору, позволяющую осуществлять процедуры как в горизонтальном, так и в вертикальном положении пациента, индивидуальные электроприводы для бедренного, коленного и голеностопного суставов ног, формирующие экзоскелетную конструкцию нижних конечностей. При применении экзоскелетной схемы комплекса привод каждой конечности должен быть способен справляться с полным весом пациента в случае, если он опирается на стопу во время ходьбы. Необходимость управления

полным весом пациента требует использования рычажного механизма, который способен в различных сегментах траектории шага (при разных положениях ноги) обеспечивать изменяющееся максимальное усилие и скорость электропривода. Такой подход позволяет уменьшить требования к максимальной мощности каждого привода экзоскелетной части реабилитационного комплекса и улучшить его габариты и массу. Трехмерная модель (цифровой двойник) реабилитационного комплекса представлена на рисунке 6.

При использовании рычагов для приведения в движение бедренного сустава для обеспечения необходимой полноты движений необходим диапазон вращения от -30 градусов до плюс 30 градусов, при этом ведущий рычаг может вращаться в диапазоне от -90 градусов до плюс 90 градусов, что при использовании например цепной передачи в этом звене могло бы обеспечить передаточное соотношение $1/3$, однако в нижней части траектории все нагрузки веса пациента воспринимает подшипниковый узел и потребность в моменте минимальна, при отклонении же на 30 градусов полный вес пациента приложенный к рычагу даст значение 300Н/м ($120\text{кг} \cdot 9.8\text{н/м}^2 \cdot \sin 30^\circ \cdot 0.5\text{м}$), причем в крайней точке скорость при траекторном движении будет равно 0 .

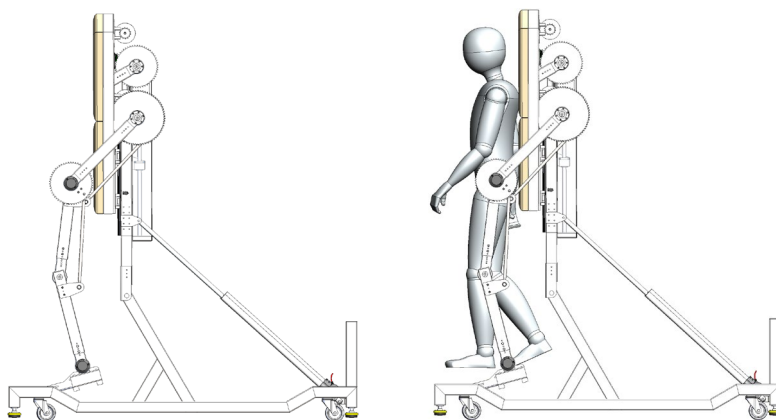


Рис. 6. Трехмерная модель реабилитационного комплекса

При использовании рычажного механизма можно так подобрать длины рычагов, чтобы при отклонении от нулевого положения коэффициент передачи уменьшался, позволяя прилагать все большее усилие для преодоления все возрастающего момента в соответствии с весом пациента, приложенным к бедренному рычагу экзоскелета.

Представим модель рычажной передачи для бедренного сустава на рисунке 7.

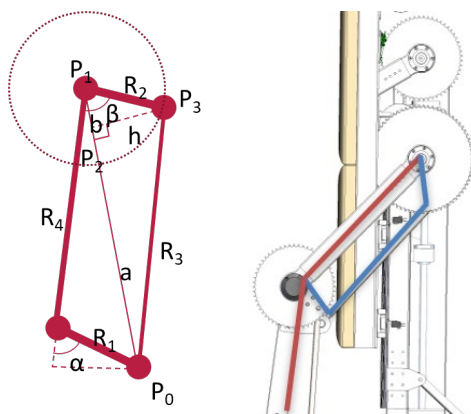


Рис. 7. Модель рычажной передачи

Кинематика рычагов описывается следующей системой уравнений:

$$\begin{aligned}
 P_0 &= (R_4 + R_1 \cos \alpha, R_1 \sin \alpha), \\
 d &= a + b = \sqrt{(R_4 + R_1 \cos \alpha)^2 + (R_1 \sin \alpha)^2}, \\
 a &= \frac{R_3^2 - R_2^2 + d^2}{2d}, P_2 = P_0 \left(1 - \frac{a}{d}\right), \\
 h &= \sqrt{R_3^2 - a^2}, x_3 = x_2 \pm \frac{hy_0}{d}, \\
 y_3 &= y_2 \pm \frac{hx_0}{d}, \beta = \tan^{-1} \frac{x_3}{y_3}.
 \end{aligned} \tag{19}$$

Предложенное конструкторское решение дает нам ряд преимуществ:

1. Реализация функции вертикализации пациента. Позволяет максимально эффективно работать с лежачими больными, начинать реабилитацию на ранних этапах.
2. Имитация различных режимов движения (движение по плоскости с замедлением и ускорением, подъем или спуск по лестнице, перешагивание препятствий).

3. Уникальная мобильность комплекса. Может быть перемещен одним человеком, легко проходит через дверные проемы.

При помощи реабилитационных методик (которые по своей сути являются программными алгоритмами воздействия исполнительных механизмов комплекса на нижние конечности пациента) комплекс поможет пациенту заново учиться делать первые шаги, ходить по лестнице, ускорять и замедлять шаг.

7. Заключение. В статье рассматриваются все этапы разработки мехатронного реабилитационного комплекса, начиная с общего анализа существующих решений, разработке математической модели походки человека и заканчивая разработкой механической составляющей комплекса. Основное внимание уделяется кинематическим и динамическим аспектам модели, что позволяет более точно воспроизводить механизмы движения.

Использование антропометрических параметров и разбивка цикла шага на четыре фазы обеспечивают детальное представление о процессе ходьбы. Валидация модели с применением маркерных систем захвата движений подтвердила высокую точность полученных данных, что является важным шагом в направлении создания эффективных реабилитационных технологий.

Результаты исследования показывают, что предложенная модель может значительно улучшить адаптивные системы управления, что, в свою очередь, может способствовать более эффективной реабилитации пациентов с нарушениями двигательной функции. В дальнейшем необходимо сосредоточиться на интеграции методов машинного обучения, что позволит повысить точность и адаптивность реабилитационных программ.

Для обеспечения точного следования траекториями для комплекса было разработано специализированное управляющее программное обеспечение, обеспечивающее не только комплексное управление всеми приводами и отображение положения комплекса на экране, но и прием сигналов от датчиков состояния пациента и обеспечение в будущем обратной связи этих показателей и интенсивности или типа реабилитационной процедуры. Например, при увеличении частоты сердечных сокращений может быть принято решение об уменьшении скорости «ходьбы», а в ряде случаев и о прекращении процедуры.

Предполагаемый эффект от внедрения данного аппарата – снижение стоимости приобретаемого оборудования медицинскими организациями, что повлечет снижение себестоимости проводимых процедур. Возможно уменьшение времени восстановления

двигательной активности пациента. Снижение трудозатрат медицинского персонала на реабилитационные процедуры больных, перенесших инсульт.

Литература

1. Vaughan-Graham J., Brooks D., Rose L., Nejat G., Pons J., Patterson K. The Use of Exoskeletons in Stroke Rehabilitation: a qualitative study of the perspectives of persons post-stroke and physiotherapists // *Journal of NeuroEngineering and Rehabilitation*. 2020. vol. 17(1). DOI: 10.1186/s12984-020-00750-x.
2. Chien W-T, Chong Y-Y, Tse M-K, Chien C-W, Cheng H-Y. Robot-assisted therapy for upper-limb rehabilitation in subacute stroke patients: A systematic review and meta-analysis // *Brain and Behavior*. 2020. vol. 10(8). DOI: 10.1002/brb3.1742.
3. Латышева В.Я., Чечетин Д.А., Ядченко Н.М., Иванцов О.А., Федоров В.В., Барбарович А.С., Филостин А.Е., Иванова Н.М. Реабилитация двигательной активности пациентов в постинсультном периоде: практическое руководство для врачей / Гомель: ГУ «РНПЦ РМиЭЧ», УО «ГТМУ», 2015. 78 с.
4. Chen G., Xu Y. Combining Virtual Reality and Robotics for Effective Stroke Rehabilitation: A Systematic Review // *Journal of NeuroEngineering and Rehabilitation*. 2021. vol. 18. pp. 1–14.
5. Краснова-Гольева В.В., Гольев М.А. Виртуальная реальность в реабилитации после инсульта // *Современная зарубежная психология*. 2015. Т. 4. № 4. С. 39–44. DOI: 10.17759/jmfp.2015040406.
6. Котов С.В., Слюнькова Е.В., Борисова В.А., Исакова Е.В. Эффективность применения интерфейсов «мозг – компьютер» и когнитивных тренингов с использованием компьютерных технологий в восстановлении когнитивных функций у пациентов после инсульта // *Журнал неврологии и психиатрии им. С.С. Корсакова*. Спецвыпуски. 2022. Т. 122(12). С. 67–75. DOI: 10.17116/jnevro202212212267.
7. REHA Technology G-EOL Brochure. URL: https://www.rehatechnology.com/wp-content/uploads/181130a_Brochure_G-EOL_En_Web_cm_in.pdf (дата обращения: 01.08.2021).
8. Hocoma Lokomat Pro. Functional Robotic Gait Therapy. URL: <https://www.hocoma.com/solutions/lokomat/> (дата обращения: 01.08.2021).
9. Motorika High-Level Gait Training Platform – ReoAmbulator. URL: <https://motorika.com/reoambulator> (дата обращения: 01.08.2021).
10. Bortole M., del Ama A., Rocon E., Moreno J.C., Brunetti F., Pons J.L. A robotic exoskeleton for overground gait rehabilitation // *IEEE International Conference on Robotics and Automation*. IEEE. 2013. pp. 3356–3361. DOI: 10.1109/ICRA.2013.6631045.
11. Beyaert C., Vasa R., Frykberg G.E. Gait post-stroke: pathophysiology and rehabilitation strategies // *Neurophysiologie Clinique/Clinical Neurophysiology*. 2015. vol. 45. № 4-5. pp. 335–355. DOI: 10.1016/j.neucli.2015.09.005.
12. Torricelli D., Cortes C., Lete N., Bertelsen A., Gonzalez-Vargas J.E., del-Ama A.J., Dimbwadyo I., Moreno J.C., Florez J., Pons J.L. A Subject-Specific Kinematic Model to Predict Human Motion in Exoskeleton-Assisted Gait // *Front Neurorobot*. 2018. vol. 12. DOI: 10.3389/fnbot.2018.00018.
13. Sun J. Dynamic Modeling of Human Gait Using a Model Predictive Control Approach // *Dissertations (1934-)*. 2015. pp. 161 p.
14. Колесникова Г.П., Формальский А.М. Об одном способе моделирования походки человека // *Инженерный журнал: наука и инновации*. 2014. Т. 1(25). DOI: 10.18698/2308-6033-2014-1-1181.

15. Королева Т.М., Липский В.Е., Липская О.В., Паркетова А.В. Использование нейросетевых технологий в биомеханике // Медицинская визуализация и компьютерная графика. 2018. № 6(3). С. 80–86.
16. Hoellinger T., Petieau M., Duvinage M., Castermans T., Seetharaman K., Cebolla A.-M., Bengoetxea A., Ivanenko Yu., Dan B., Cheron G. Biological oscillations for learning walking coordination: dynamic recurrent neural network functionally models physiological central pattern generator // *Frontiers in Computational Neuroscience*. 2013. vol. 7(70). DOI: 10.3389/fncom.2013.00070.
17. Верховод Д.П., Вороной В.В., Побединский С.Ю. Разработка системы управления и позиционирования узлов мехатронного реабилитационного комплекса // Системы анализа и обработки данных. 2023. Т. 92. № 4. С. 23–34.
18. Slim M., Rokbani N., Neji B., Terres M.A., Beyrouthy T. Inverse Kinematic Solver Based on Bat Algorithm for Robotic Arm Path Planning. *Robotics*. 2023. vol. 12(2). pp. 1–26. DOI: 10.3390/robotics12020038.
19. Вороной В.В., Верховод Д.П. Выбор математической модели походки человека для применения в реабилитационных устройствах // Сибирский журнал клинической и экспериментальной медицины. 2023. № 38(4). С. 231–235.
20. Cavagna G., Kaneko M. Mechanical work and efficiency in level walking and running // *The Journal of physiology*. 1977. vol. 268. no. 2. pp. 467–481. DOI: 10.1113/jphysiol.1977.sp011866.
21. Цыгулин А.А. Разработка реабилитационного комплекса Гефест для пациентов после перенесенного инсульта // Экстремальная роботехника. 34-я международная научно-техническая конференция. Сборник тезисов. Санкт-Петербург: ООО «Типография Фурсова», 2023. С. 188–191.
22. McGregor A.H., Warren M., Brooks J.M., Feldman D.S. The Role of Pelvic and Lower Limb Biomechanics in Post-Stroke Gait Rehabilitation: A Systematic Review // *Journal of NeuroEngineering and Rehabilitation*. 2021. vol. 18. pp. 45.
23. Chen Z., Franklin D.W. Musculotendon Parameters in Lower Limb Models: Simplifications, Uncertainties, and Muscle Force Estimation Sensitivity // *Annals of Biomedical Engineering*. 2023. vol. 51. pp. 1147–1164. DOI: 10.1007/s10439-023-03166-5.

Вороной Вадим Владимирович — канд. техн. наук, доцент кафедры, кафедра систем сбора и обработки данных, Новосибирский государственный технический университет (НГТУ). Область научных интересов: разработка систем управления технологическими процессами, математическое моделирование систем. Число научных публикаций — 29. voronovj@corp.nstu.ru; проспект Карла Маркса, 20, 630073, Новосибирск, Россия; р.т.: +7(383)346-0843.

Верховод Дмитрий Павлович — младший научный сотрудник, центр технологического превосходства, Новосибирский государственный технический университет (НГТУ). Область научных интересов: разработка реабилитационных комплексов, исследования и моделирование походки человека. Число научных публикаций — 2. verhovod@corp.nstu.ru; проспект Карла Маркса, 20, 630073, Новосибирск, Россия; р.т.: +7(383)346-0843.

Поддержка исследований. Работа выполнена в рамках программы «Приоритет 2030» Новосибирского Государственного Технического Университета.

V. VORONOV, D. VERHOVOD
**HUMAN GAIT MATHEMATICAL MODEL DEVELOPMENT
AND CONTROL SYSTEM SYNTHESIS FOR ACTUATING
MECHANISMS OF THE MECHATRONIC REHABILITATION
COMPLEX**

Voronov V., Verhovod D. Human Gait Mathematical Model Development and Control System Synthesis for Actuating Mechanisms of the Mechatronic Rehabilitation Complex.

Abstract. The article discusses the development of a mathematical model of human gait for the synthesis of a control system for a mechatronic rehabilitation complex. The relevance of the research is determined by the necessity to create effective rehabilitation technologies for patients with motor function impairments. Existing rehabilitation complexes can be divided into exoskeletons and devices with mechanical linkage (end effectors), with exoskeletons demonstrating higher rehabilitation effectiveness by mimicking natural gait. The scientific novelty of this study lies in the development of a model that takes into account the individual anthropometric parameters of the patient, including body mass and the lengths of limb segments, as well as the ability to simulate foot rotation. Within the framework of the study, a method for dividing the gait cycle into four phases is proposed, each described by a separate system of mathematical equations, which ensures high accuracy in reproducing various stages of movement. To validate the model, a marker-based motion capture system was used, which provided data on movement trajectories. The results showed that the model effectively generates trajectories of sagittal angles of hip, shank, and foot elevation, contributing to improved control of the rehabilitation device. In conclusion, the work emphasizes the importance of mathematical modeling for the development of adaptive control systems that can significantly enhance the rehabilitation process. Further research will focus on refining the model and integrating it with machine learning methods to improve the accuracy and reliability of rehabilitation programs.

Keywords: mathematical model, mechatronic rehabilitation complex, exoskeleton, kinematic model, parametric model.

References

1. Vaughan-Graham J., Brooks D., Rose L., Nejat G., Pons J., Patterson K. The Use of Exoskeletons in Stroke Rehabilitation: a qualitative study of the perspectives of persons post-stroke and physiotherapists. *Journal of NeuroEngineering and Rehabilitation*. 2020. vol. 17(1). DOI: 10.1186/s12984-020-00750-x.
2. Chien W-T, Chong Y-Y, Tse M-K, Chien C-W, Cheng H-Y. Robot-assisted therapy for upper-limb rehabilitation in subacute stroke patients: A systematic review and meta-analysis. *Brain and Behavior*. 2020. vol. 10(8). DOI: 10.1002/brb3.1742.
3. Latysheva V.Ja., Chechetin D.A., Jadchenko N.M., Ivancov O.A., Fedorov V.V., Barbarovich A.S., Filjustin A.E., Ivanova N.M. Reabilitacija dvigatel'noj aktivnosti pacientov v postinsul'nom periode: prakticheskoe rukovodstvo dlja vrachej [Rehabilitation of motor activity of patients in the post-stroke period: a practical guide for doctors]. Gomel': GU «RNPС RMiJeCh», UO «GGMU», 2015. 78 p. (In Russ.).
4. Chen G., Xu Y. Combining Virtual Reality and Robotics for Effective Stroke Rehabilitation: A Systematic Review. *Journal of NeuroEngineering and Rehabilitation*. 2021. vol. 18. pp. 1–14.

5. Krasnova-Gol'eva V.V., Gol'ev M.A. [Virtual'naja real'nost' v rehabilitacii posle insulta]. *Sovremennaja zarubezhnaja psihologija – Modern foreign psychology*. 2015. vol. 4. no. 4. pp. 39–44. DOI: 10.17759/jmfp.2015040406. (In Russ.).
6. Kotov S.V., Sljun'kova E.V., Borisova V.A., Isakova E.V. [Efficiency of using brain-computer interfaces and cognitive training using computer technologies in restoring cognitive functions in patients after stroke]. *Zhurnal nevrologii i psichiatrii im. S.S. Korsakova. Specvypuski – Journal of Neurology and Psychiatry named after S.S. Korsakov. Special issues*. 2022. vol. 122(12). pp. 67–75. DOI: 10.17116/jnevro202212212267. (In Russ.).
7. REHA Technology G-EOL Brochure. Available at: https://www.rehatechnology.com/wp-content/uploads/181130a_Brochure_G-EOL_En_Web_cm_in.pdf (accessed 01.08.2021).
8. Hocoma Lokomat Pro. Functional Robotic Gait Therapy. Available at: <https://www.hocoma.com/solutions/lokomat/> (accessed 01.08.2021).
9. Motorika High-Level Gait Training Platform – ReoAmbulator. Available at: <https://motorika.com/reoambulator> (accessed 01.08.2021).
10. Bortole M., del Ama A., Rocon E., Moreno J.C., Brunetti F., Pons J.L. A robotic exoskeleton for overground gait rehabilitation. *IEEE International Conference on Robotics and Automation. IEEE*. 2013. pp. 3356–3361. DOI: 10.1109/ICRA.2013.6631045.
11. Beyaert C., Vasa R., Frykberg G.E. Gait post-stroke: pathophysiology and rehabilitation strategies. *Neurophysiologie Clinique/Clinical Neurophysiology*. 2015. vol. 45. № 4-5. pp. 335–355. DOI: 10.1016/j.neucli.2015.09.005.
12. Torricelli D., Cortes C., Lete N., Bertelsen A., Gonzalez-Vargas J.E., del-Ama A.J., Dimbwadyo I., Moreno J.C., Florez J., Pons J.L. A Subject-Specific Kinematic Model to Predict Human Motion in Exoskeleton-Assisted Gait. *Front Neurorobot*. 2018. vol. 12. DOI: 10.3389/fnbot.2018.00018.
13. Sun J. Dynamic Modeling of Human Gait Using a Model Predictive Control Approach. *Dissertations (1934-)*. 2015. pp. 161 p.
14. Kolesnikova G.P., Formal'skij A.M. [Ob odnom sposobe modelirovanija pohodki cheloveka]. *Inzhenernyj zhurnal: nauka i innovacii – Engineering Journal: Science and Innovation*. 2014. vol. 1(25). DOI: 10.18698/2308-6033-2014-1-1181. (In Russ.).
15. Королева Т.М., Липский В.Е., Липская О.В., Паркетова А.В. Использование нейросетевых технологий в биомеханике. Медицинская визуализация и компьютерная графика. 2018. № 6(3). С. 80–86.
16. Hoellinger T., Petieau M., Duvinage M., Castermans T., Seetharaman K., Cebolla A.-M., Bengoetxea A., Ivanenko Yu., Dan B., Cheron G. Biological oscillations for learning walking coordination: dynamic recurrent neural network functionally models physiological central pattern generator. *Frontiers in Computational Neuroscience*. 2013. vol. 7(70). DOI: 10.3389/fncom.2013.00070.
17. Verhovod D.P., Voronoy V.V., Pobedinskij S.Ju. [Razrabotka sistemy upravlenija i pozicionirovanija uzlov mehatronnogo reabilitacionnogo kompleksa]. *Sistemy analiza i obrabotki dannyh – Data analysis and processing systems*. 2023. vol. 92. no. 4. pp. 23–34. (In Russ.).
18. Slim M., Rokbani N., Neji B., Terres M.A., Beyrouthy T. Inverse Kinematic Solver Based on Bat Algorithm for Robotic Arm Path Planning. *Robotics*. 2023. vol. 12(2). pp. 1–26. DOI: 10.3390/robotics12020038.
19. Voronoy V.V., Verhovod D.P. [Vybor matematicheskoy modeli pohodki cheloveka dlja primenenija v reabilitacionnyh ustroystvah]. *Sibirskij zhurnal klinicheskoy i jeksperimental'noj mediciny – Siberian Journal of Clinical and Experimental Medicine*. 2023. no. 38(4). pp. 231–235. (In Russ.).

20. Cavagna G., Kaneko M. Mechanical work and efficiency in level walking and running. *The Journal of physiology*. 1977. vol. 268. no. 2. pp. 467–481. DOI: 10.1113/jphysiol.1977.sp011866.
21. Cygulin A.A. Razrabotka reabilitacionnogo kompleksa Gefest dlja pacientov posle perenesennogo insulta [Development of the rehabilitation complex Gefest for patients after a stroke]. *Jekstremal'naja robotehnika. 34-ja mezhdunarodnaja nauchno-tehnicheskaja konferencija. Sbornik tezisov [Extreme Robotics. 34th International Scientific and Technical Conference. Abstracts]*. Sankt-Peterburg: OOO «Tipografija Fursova», 2023. pp. 188–191. (In Russ.).
22. McGregor A.H., Warren M., Brooks J.M., Feldman D.S. The Role of Pelvic and Lower Limb Biomechanics in Post-Stroke Gait Rehabilitation: A Systematic Review. *Journal of NeuroEngineering and Rehabilitation*. 2021. vol. 18. pp. 45.
23. Chen Z., Franklin D.W. Musculotendon Parameters in Lower Limb Models: Simplifications, Uncertainties, and Muscle Force Estimation Sensitivity. *Annals of Biomedical Engineering*. 2023. vol. 51. pp. 1147–1164. DOI: 10.1007/s10439-023-03166-5.

Voronoy Vadim — Ph.D., Associate professor of the department, Data collection and processing systems department, Novosibirsk State Technical University (NSTU). Research interests: development of control systems for technological processes, mathematical modeling of systems. The number of publications — 29. voronovj@corp.nstu.su; 20, Karl Marx Av., 630073, Novosibirsk, Russia; office phone: +7(383)346-0843.

Verhovod Dmitriy — Junior research fellow, Technology excellence center, Novosibirsk State Technical University (NSTU). Research interests: development of rehabilitation complexes, research and modeling of human gait. The number of publications — 2. verxovod@corp.nstu.ru; 20, Karl Marx Av., 630073, Novosibirsk, Russia; office phone: +7(383)346-0843.

Acknowledgements. The work was carried out as part of the «Priority 2030» program at Novosibirsk State Technical University.

А.И. АНДРЕЕВ, М.О. КУЧМА, С.И. МАЛЬКОВСКИЙ, А.А. ФИЛЕЙ
**RAINCAST: ГИБРИДНЫЙ АЛГОРИТМ НАУКАСТИНГА
ИНТЕНСИВНОСТИ ОСАДКОВ ПО ДАННЫМ СПУТНИКОВЫХ
НАБЛЮДЕНИЙ КА HIMAWARI-8/9**

Андреев А.И., Кучма М.О., Мальковский С.И., Филей А.А. RainCast: Гибридный алгоритм наукастинга интенсивности осадков по данным спутниковых наблюдений КА Himawari-8/9.

Аннотация. В работе предложен алгоритм RainCast (Rain Rate NowCast), предназначенный для краткосрочного прогноза интенсивности осадков на срок до двух часов. Данное направление метеорологии, в англоязычной литературе получившее название «наукастинг» (nowcasting), является одним из наиболее востребованных информационных продуктов во многих сферах человеческой деятельности. Вместе с тем, его доступность может быть сильно ограничена имеющейся наземной инфраструктурой. В настоящей работе авторами поставлена цель создания алгоритма прогноза осадков, предназначенного для одной из таких территорий на примере Азиатско-Тихоокеанского регионе с использованием спутниковых измерений космического аппарата Himawari-8/9. Предложенный алгоритм объединяет преимущества детерминированного и статистического подходов к задаче прогноза и основан на использовании двух нейросетевых моделей. Первая модель осуществляет предварительный прогноз общего направления движения осадков на мезомасштабном уровне с помощью физически-обусловленной нейронной сети NowcastNet. Вторая модель на основе архитектуры CasFormer служит для постобработки полученного прогноза методом диффузий с целью повышения его детализации. Разработанный с использованием данного метода гибридный алгоритм RainCast позволяет получать краткосрочный прогноз интенсивности осадков на срок до двух часов с частотой каждые 10 минут и пространственным разрешением 2 км/пиксель на основе спутниковых измерений в инфракрасном диапазоне. Спутниковые измерения используются для расчета интенсивности осадков по ранее предложенному авторами алгоритму. На основе карт осадков сформированы обучающие, валидационные и тестовые наборы данных, применяемые для обучения, настройки алгоритма и оценки качества прогнозов соответственно. Для предложенного алгоритма RainCast, а также применяемых в мировой практике актуальных решений, таких как NowcastNet, CasFormer, Earthformer, было проведено обучение с использованием сформированных наборов данных, после чего для указанных алгоритмов был рассчитан ряд показателей качества получаемых прогнозов. Из полученных результатов можно сделать вывод, что предложенный в настоящей работе гибридный алгоритм RainCast имеет более высокое качество прогноза в сравнении с другими вышеперечисленными решениями при равных условиях. Для валидационного набора данных на территории Дальневосточного региона при прогнозе алгоритмом RainCast с заблаговременностью 2 часа были получены следующие показатели для последнего спрогнозированного кадра: корень среднеквадратической ошибки RMSE составил 0.88 мм/ч, вероятность обнаружения POD – 0.78, коэффициент корреляции Пирсона – 0.75, индекс структурного сходства SSIM – 0.91, а пиковое отношение сигнала к шуму PSNR – 36.63. Дополнительно проведенный визуальный анализ прогнозов подтвердил, что наиболее близкие к фактическим наблюдениям результаты имеет предложенный гибридный региональный алгоритм RainCast.

Ключевые слова: наукастинг, осадки, интенсивность осадков, прогноз осадков, Himawari, RainCast, диффузионные модели.

1. Введение. Термином «наукастинг» английский метеоролог Кит Браунинг обозначил в 80-х годах прошлого века «детализированное описание текущего состояния погоды и прогноз изменений, которые ожидаются в течение первых 2 ч» [1]. Его востребованность связана с его способностью обеспечивать потребителей более детализированной информацией, чем это позволяют другие виды метеорологических прогнозов. В особенности это важно при прогнозе опасных явлений погоды, многие из которых характеризуются коротким жизненным циклом и небольшой пространственной протяженностью. Поля интенсивности осадков были исторически первым элементом наукастинга [2]. Несмотря на расширение круга прогнозируемых метеорологических параметров, осадки остаются ключевым объектом исследований и моделирования на всех пространственно-временных масштабах, включая наукастинг.

В контексте данной работы под наукастингом понимается краткосрочный (заблаговременностью до двух часов) прогноз интенсивности осадков. Несмотря на то, что для целей прогноза осадков важны приземные и аэрологические наблюдения, лишь системы дистанционного зондирования могут адекватным образом обеспечить пространственный охват с высоким разрешением. Для краткосрочного прогноза осадков наилучшим образом подходят радиолокационные наблюдения, позволяющие отображать поля облачности и осадков в реальном масштабе времени с достаточной дискретностью. Однако на значительных территориях Сибири и Дальнего Востока России на текущий момент необходимые оперативные радиолокационные системы отсутствуют. Альтернативной возможностью количественной и качественной оценки осадкообразующей облачности с высоким пространственно-временным разрешением на этих территориях могут быть измерения геостационарных и высокоэллиптических космических аппаратов (КА). Развитие в последние годы возможностей российских и зарубежных КА гидрометеорологического назначения [3] позволило рассматривать эти данные в качестве базовой информации для создания прогностических алгоритмов. Так, в рамках российской космической программы производится пополнение спутниковой группировки новыми КА гидрометеорологического назначения. В частности, запланированы запуски новых высокоэллиптических КА серии «Артика-М» для наблюдений за Арктическим регионом, а также геостационарных КА серии «Электро-Л». Запуск подобных аппаратов позволит значительно расширить горизонты наблюдений, особенно для труднодоступных регионов. Для эффективного усвоения этих

спутниковых данных требуется создание и обработка технологий получения информационной продукции. В связи с этим в настоящей работе используются данные зарубежного геостационарного КА Himawari-8/9.

Для прогноза интенсивности осадков, полученных с помощью спутниковых наблюдений, чаще всего применяются методы на основе алгоритмов оптического потока и машинного обучения. Первые из них, в силу своей линейности [4], недостаточно эффективны для прогноза конвективной осадкообразующей облачности [5, 6, 7], быстро меняющейся во времени [8]. С другой стороны, появление современных нейросетевых архитектур для обработки изображений позволило значительно продвинуться в развитии альтернативных методов прогноза [9, 10]. Искусственные нейронные сети (ИНС) хорошо учитывают нелинейные зависимости на основе данных и не требуют явного задания физических закономерностей и моделей, по которым осуществляется прогноз [6]. Это дает возможность принять во внимание большое количество второстепенных факторов, которые сложно учесть классическими методами экстраполяции. Но при всем многообразии стремительно развивающихся нейросетевых архитектур спутниковые данные, в отличие от радарных, недостаточно точно позволяют учесть трехмерную (3-D) структуру облака, что требует от исследователей поиска комбинированных подходов, с помощью которых можно не только максимально точно определять перемещение потенциально осадкообразующей облачности, но и прогнозировать на требуемые сроки динамику высоковариативных элементов такой облачности.

В настоящей работе предлагается алгоритм RainCast для краткосрочного прогноза осадков по спутниковым данным на примере КА Himawari-8/9 на территории Дальневосточного региона России. В качестве источника информации об осадках авторы используют результаты расчета собственного алгоритма по данным этого спутника. Предлагаемый алгоритм RainCast сочетает преимущества статистического и детерминированного подходов и основан на идее постобработки с помощью диффузионной модели ИНС. Полученные результаты показали высокое качество прогноза, достигающее уровня точности лучших из представленных ранее решений, что подтверждается численными и визуальными оценками.

2. Текущее состояние исследований. С появлением первых нейросетевых архитектур для прогноза осадков на примере сверточнорекуррентных моделей ConvLSTM [11] и TrajGRU [12] было показано, что ИНС имеют высокий потенциал в решении данной задачи.

Отмечается, что сверточные сети обладают высокой скоростью обработки, однако недостаточно хорошо справляются с обработкой временных последовательностей [7]. С другой стороны, рекуррентные сети лучше выявляют временные связи между кадрами, но имеют высокую вычислительную сложность [13]. Чтобы увеличить продолжительность прогноза, рассматриваются различные варианты архитектур специальных ячеек памяти для хранения контекстуальной информации в рекуррентных архитектурах, таких как PredRNN2 [14]. Кроме того, рассматриваются варианты прогноза путем его декомпозиции на составляющие изменения интенсивности пикселей (transient variation) и общего перемещения облачности (motion trend), как, например, в модели MotionRNN [15]. Чтобы компенсировать недостаток рекуррентных архитектур в виде потери контекста при большой продолжительности прогноза, применяются архитектуры трансформеров, такие как Earthformer [16] и Rainformer [17]. Ключевым элементом подобных архитектур является встроенный механизм внимания, позволяющий эффективно использовать информацию из любого кадра всей входной последовательности, что в свою очередь положительно влияет на качество прогноза и позволяет достичь наилучших показателей точности [7, 13].

Вышеперечисленные модели относятся к классу детерминированных [18, 19]. С их помощью можно достаточно точно прогнозировать общее направление движения осадкообразующей облачности, однако прогноз локальных явлений затруднен из-за их высокой вариативности [20]. Под локальными явлениями понимается осадкообразующая облачность малого масштаба, чаще всего состоящая из групп отдельных облаков с линейным размером порядка 10 пикселей. Для таких явлений подобные модели предсказывают усредненные значения интенсивности осадков, из-за чего прогноз получается размытым, а детализация изображений быстро снижается [7, 18, 21]. С другой стороны, статистические методы, такие как генеративно-состязательные (Generative Adversarial Network, GAN) и диффузионные сети, осуществляют прогноз, проецируя его из вероятностного распределения всех возможных значений интенсивности [16, 20]. Такой подход позволяет повысить детализацию и достичь реалистичности изображений, однако, в силу стохастической природы подобных моделей, прогноз может несколько варьироваться от случая к случаю для одних и тех же исходных данных [6, 19].

Диффузионные и GAN модели позволяют достичь наилучших результатов с точки зрения заблаговременности прогнозов, а также

резкости и детализации спрогнозированных изображений [22, 23]. Идея GAN-моделей основана на теории игр и заключается в использовании двух ИНС – генератора и дискриминатора, конкурирующих между собой. GAN позволяют создавать реалистичные изображения, однако их обучение является достаточно трудоемким процессом из-за возникающей нестабильности и коллапса мод (mode collapse) [22, 24, 25], а генерируемые изображения нередко содержат артефакты [18, 24]. В качестве примера GAN-моделей можно привести модель DGMR (Deep Generative Model of Rainfall) [20], имеющую одни из лучших показателей точности вместе с высокой детализацией и четкостью изображений. При этом DGMR использует одновременно два дискриминатора при обучении для контроля временной и пространственной компоненты прогноза. Для повышения заблаговременности прогноза осадков высокой интенсивности в некоторых случаях в модель интегрируются элементы физики. Например, в работе [26] предложена модель NowcastNet на основе GAN, ключевой идеей которой стало использование вспомогательной ИНС EvolutionNet, реализующей физический закон непрерывности потока. Здесь можно отметить, что использование физических принципов или данных численного прогноза погоды (ЧПМ) [24, 27, 28] в комбинации с ИНС в том или ином виде положительно влияет на качество прогнозов не только GAN-моделей [1, 23], но и других нейросетевых архитектур в целом [29, 30].

Недостатки моделей на основе GAN удастся преодолеть путем использования диффузионных ИНС [24, 30], достигающих при этом схожих результатов [22, 25]. Однако проблемами диффузионных моделей являются высокая вычислительная сложность, необходимость тонкой настройки архитектуры ИНС и подбора гиперпараметров обучения [32], вследствие чего получаемые изображения имеют низкое разрешение. Чтобы улучшить разрешение, можно использовать диффузионные ИНС для уточнения имеющегося прогноза, сделанного с помощью другой модели. Например, авторами работы [18] предложена модель CasCast, используемая для постобработки исходного детерминированного прогноза на примере сети-трансформера Earthformer. CasCast достигает уровня точности, схожего с GAN-моделью NowcastNet, но с учетом преимуществ диффузионного подхода. Подобный принцип использовали авторы [25], представившие свой вариант

диффузионной модели DiffCast для постобработки детерминированного прогноза.

Хотя предлагаемые различными авторами нейросетевые модели, как правило, обучаются на общедоступных данных, в основном на основе наземных радиолокационных измерений (НКО-7 [12]) и реже – с использованием спутниковых наблюдений (SEVIR [33]), их практическое применение сопряжено с рядом ограничений. Точность прогнозов таких моделей во многом зависит от соответствия обучающих данных характеристикам конкретных региональных условий и используемых средств наблюдений. В связи с этим в данной работе предложена методика формирования специализированной обучающей выборки, адаптированной под особенности целевого региона и тип оборудования, что позволяет повысить качество прогноза осадков для заданных условий.

3. Используемые данные. В работе рассматривается Азиатско-Тихоокеанский регион с учетом морской акватории, ограниченный $35\text{--}60^\circ$ с.ш. и $100\text{--}160^\circ$ в.д. На большей части региона характерен муссонный климат, характерными чертами которого являются суровая малоснежная зима и относительно теплое, обильное осадками лето. Годовое количество осадков, за исключением Приморья, где оно местами достигает 800 мм, мало отличается от средней полосы Европейской части России и составляет порядка 450–600 мм в год. Однако распределение осадков в течение года имеет существенные особенности по сравнению с указанным регионом: они обильны в период теплой части года и исключительно малы в зимнее время. Дожди на Дальнем Востоке носят ливневой характер, охватывают большие территории и отличаются высокой интенсивностью. В некоторых случаях за одни сутки выпадает до 150–250 мм [34].

Для решения задачи краткосрочного прогноза осадков для рассматриваемой территории требуются данные об осадках с высокой временной дискретностью. Поскольку на территории Дальнего Востока плотность сети наземных наблюдений остается крайне невысокой, в работе используются спутниковые измерения спектрорадиометра АНИ (Advanced Himawari Imager), установленного на геостационарном КА Himawari-8/9. Данный прибор предоставляет изображения облачности и подстилающей поверхности с частотой 10 минут в 16 различных диапазонах электромагнитного спектра. Пространственное разрешение составляет 0.5 км/пиксель для канала с центральной длиной волны 0.645 мкм, 1 км/пиксель для видимых и коротковолновых инфракрасных каналов (до 3,85 мкм) и 2

км/пиксель в надире для остальных инфракрасных каналов (в настоящей работе все каналы приведены к единому разрешению 2 км/пиксель).

Для получения мгновенных значений интенсивности осадков по измерениям прибора АНІ авторами ранее был разработан алгоритм [35], основанный на использовании двух ИНС трансформерной и сверточной архитектур, одна из которых предназначена для обнаружения осадкообразующей облачности и расчета маски осадков, а другая позволяет оценить значение интенсивности осадков в каждом пикселе маски. Для обучения алгоритма использовались данные, предоставляемые в рамках проекта GPM (Global Precipitation Measurements) по глобальному мониторингу осадков путем комбинирования измерений различных КА [36]. Использование данных GPM непосредственно для краткосрочного прогноза осадков затруднено по причине задержки в их получении от 4 часов, что связано с необходимостью накопления измерений различных приборов. Разработанный авторами алгоритм для оценки осадков позволяет получать их оперативно с исходным пространственно-временным разрешением прибора АНІ практически сразу после получения спектральных измерений, что крайне важно для решения рассматриваемой задачи, где требуются наиболее актуальные данные для прогноза.

Для разработки прогностического алгоритма авторами были сформированы обучающий, тестовый и валидационный наборы данных, охватывающие временной период с 2020 по 2024 год. Данные представляют собой последовательности изображений интенсивности осадков размером 512×512 пикселей для всей рассматриваемой территории. Каждая последовательность состоит из 21 изображения, из которых входными являются первые 9. Интенсивности осадков выражены в мм/ч и нормированы в диапазоне $[0, 1]$. Таким образом, общая продолжительность последовательности с учетом временного разрешения прибора АНІ составляет 210 минут. Исходная выборка, состоящая из 20 тысяч последовательностей, была разделена на наборы данных, предназначенные для обучения нейросетевых моделей (80% от общего объема выборки), их проверки и настройки (2%), а также итогового тестирования (18%).

4. Метод. Предлагаемый алгоритм RainCast объединяет преимущества детерминированного [25] и статистического [18] подходов, что в равной степени позволяет акцентировать внимание на прогнозе осадков как на мезомасштабном, так и на локальном уровне.

На рисунке 1 представлена блок-схема алгоритма прогноза, детали которого будут описаны ниже.

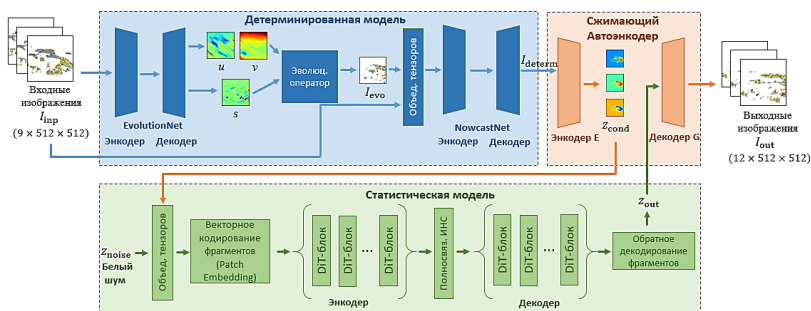


Рис. 1. Блок-схема алгоритма наукастинга осадков RainCast

4.1. Детерминированная модель. Работа алгоритма RainCast начинается с подготовки общего прогноза осадков без учета локальной облачности при помощи детерминированной модели. Для этого в настоящей работе используется модель-генератора из GAN-архитектуры NowcastNet [26]. Данная модель имеет одни из наиболее высоких показателей точности в задаче прогноза осадков [24, 29]. В ее состав интегрирована вспомогательная модель EvolutionNet, являющаяся нейросетевой реализацией физического принципа непрерывности потока, что делает модель NowcastNet физически обусловленной.

Обучение модели NowcastNet осуществляется в два этапа. На первом этапе происходит обучение подсети EvolutionNet, выходной информацией которой являются карты признаков s , u и v , первая из которых связана с изменением интенсивности осадков, а две другие – с перемещением облачного массива в двух проекциях. Эволюционный оператор, действуя на последний входной кадр, перемещает пиксели в соответствии с полями u и v , а затем прибавляет интенсивность s , в результате чего формируется спрогнозированный кадр I^{evo} [26]. Следуя работе [26], для обучения модели EvolutionNet использовалась следующая целевая функция потерь:

$$L_{evonet} = L_{accum} + 0.01 \cdot L_{motion}, \quad (1)$$

$$L_{\text{accum}} = \sum_{t=1}^T (L_{\text{wdis}}(I_t, I_t^{\text{evo bili}}) + L_{\text{wdis}}(I_t, I_t^{\text{evo}})), \quad (2)$$

$$L_{\text{motion}} = \sum_{t=1}^T (\|\nabla u_t \odot \sqrt{w(I_t)}\|_{L_2}^2 + \|\nabla v_t \odot \sqrt{w(I_t)}\|_{L_2}^2), \quad (3)$$

$$L_{\text{wdis}}(I_t, I_t^{\text{evo}}) = \|(I_t - I_t^{\text{evo}}) \odot w(I_t)\|_{L_1}, \quad (4)$$

где L_{accum} – ошибка прогноза значений интенсивности, L_{motion} – ошибка прогноза положения пикселей с осадками, L_{wdis} – взвешенное значение ошибки с учетом интенсивности осадков, представляющее собой L_1 норму, T – количество спрогнозированных кадров, I_t – фактическое значение интенсивности из обучающей выборки для кадра с номером t , I_t^{evo} и $I_t^{\text{evo bili}}$ – интенсивность, полученная в результате действия эволюционного оператора с интерполяцией по методу ближайшего соседа и билинейной интерполяцией соответственно, $\odot$ – операция поэлементного умножения тензоров, $w(I_t)$ – весовой коэффициент, применяемый для нормированной в диапазоне $[0,1]$ интенсивности при ее максимальном значении, равном 50 мм/ч:

$$w(I_t) = \begin{cases} 0.5, & \text{если } I_t < 0.2 \\ 1 + I_t, & \text{если } I_t \geq 0.2 \end{cases}. \quad (5)$$

Градиент полей u и v рассчитывается приближенно путем применения операции свертки с фильтром Собеля [26]. Обучение модели EvolutionNet осуществлялось с использованием метода оптимизации Adam [37] с коэффициентом скорости обучения 10^{-3} на протяжении первых 100 эпох, после чего производилось дообучение сети с уменьшенным коэффициентом, равным 10^{-4} . Далее весовые коэффициенты модели EvolutionNet фиксировались, а выход модели I^{evo} объединялся с входными кадрами I^{inp} , после чего полученный тензор подавался на вход автоэнкодера NowcastNet (рисунок 1). Для обучения детерминированной модели NowcastNet использовалась

взвешенная функция потерь в виде суммы L_I и L_2 норм, обеспечивающая сбалансированное значение ошибки между различными уровнями интенсивности [7]:

$$L_{\text{determ}} = \frac{1}{N} \sum_{t=1}^T (\|(I_t - I_t^{\text{determ}}) \odot w(I_t)\|_{L_1} + \|(I_t - I_t^{\text{determ}}) \odot w(I_t)\|_{L_2}), \quad (6)$$

где N – суммарное количество пикселей в кадре. Аналогично EvolutionNet, обучение осуществлялось с использованием метода Adam с коэффициентом скорости обучения, равным 10^{-3} , уменьшенным на этапе дообучения до 10^{-4} спустя 100 обучающих эпох. Выходной прогноз I_t^{determ} модели NowcastNet использовался в дальнейшем в качестве входной информации для статистической модели.

4.2. Статистическая модель. Вторым шагом работы алгоритма RainCast является применение статистического подхода на основе диффузионной ИНС. Это позволяет уточнить исходный детерминированный прогноз, недостатком которого является низкая точность прогнозирования локальных областей осадков из-за применения детерминированной функции потерь, приводящей к потере детализации [20]. Диффузионная ИНС позволяет реконструировать изображение таким образом, чтобы оно выглядело максимально похожим на фактическое изображение осадков. Статистическая модель обучается находить характерные паттерны на изображениях на основе обучающей выборки и дотраивать изображение, дополняя его недостающими деталями.

4.2.1. Сжимающий автоэнкодер. Для уменьшения вычислительной сложности при использовании диффузионного подхода в настоящей работе используется обученный по методу GAN вариационный автоэнкодер VQGAN (Vector Quantised Variational Autoencoder trained as GAN), применяемый для сжатия изображений с минимальными потерями в точности [38, 39].

Автоэнкодер состоит из двух частей – энкодера E и декодера G . Энкодер сжимает исходное изображение из пространства $I_t \in \mathbb{R}^{H \times W \times C}$ в новое скрытое представление $z_t = E(I_t) \in \mathbb{R}^{h \times w \times c}$, где H, W, C и h, w, c – размерности длины, ширины и количества каналов исходного ($512 \times 512 \times 12$) и сжатого ($64 \times 64 \times 8$) изображений. Декодер осуществляет обратное преобразование: $\bar{I}_t = G(z_t) = G(E(I_t))$. В процессе осуществления прогноза (рисунок 1) энкодеру E передается

результат работы детерминированной модели I^{determ} , в результате чего формируется тензор z^{cond} . Данный тензор задает начальные условия для прогноза с помощью статистической модели, описанной ниже. Выходной информацией энкодера является тензор z^{out} . Его декодирование осуществляется при помощи декодера G, результатом работы которого являются спрогнозированные изображения I^{out} .

Процесс обучения автоэнкодера заключается в демонстрации эталонных изображений I_t из исходной выборки. Целевой функцией минимизации при этом является комбинированная функция потерь [40]:

$$L_{\text{AE}} = \min_{E,D} \max_D (L_1(I_t, \bar{I}_t) - L_{\text{adv}}(\bar{I}_t) + \log D(I_t) + L_{\text{reg}}(I_t; \bar{I}_t)), \quad (7)$$

где L_{adv} – генеративно-состязательная функция потерь [39], L_{reg} – регуляризационный параметр на основе дивергенции Кульбака-Лейблера [41], умноженный на коэффициент 10^{-6} , D – модель дискриминатора из работы [18]. Первые 25 тыс. итераций обучения проходят при условии $L_{\text{adv}}(\bar{I}_t) + \log D(I_t) = 0$. Общее количество обучающих итераций составляет 200 тыс.

4.2.2. Диффузионная модель. В качестве диффузионной модели в настоящей работе используется модель CasFormer [18], основанная на архитектуре трансформера DiT (Diffusion Transformer) [42]. Модель принимает в качестве входных данных результаты детерминированного прогноза, сжатые с использованием энкодера E в скрытое представление z^{cond} . Выходной информацией модели является скорректированный прогноз z^{out} , который затем реконструируется до полного изображения осадков с помощью декодера G.

Принцип работы диффузионной модели состоит из прямого и обратного процесса. При прямом процессе к исходному изображению $I_0 = I_t^{\text{determ}}$, преобразованному в скрытое представление $z_0 = E(I_0)$ итеративно добавляется случайный шум $q(z_k | z_{k-1}) = N(\sqrt{\alpha_k} z_{k-1}, \beta_k \sigma^2), 1 \leq k \leq K$, где $\beta_k = 1 - \alpha_k$, k – номер итерации процесса зашумления, K – количество итераций зашумления, α – коэффициент, регулирующий количество добавляемого шума на каждой итерации, N – распределение Гаусса с

параметрами математического ожидания и дисперсии соответственно, σ – величина дисперсии значений пикселей изображения. При обратном процессе полученное совместное распределение факторизуется на произведение условных распределений $p_{\theta}(z_{0:K}) = p(z_K) \prod_{i=K}^{i=1} p_{\theta}(z_{i-1} | z_i)$. Диффузионная модель Ψ с параметрами θ обучается прогнозировать шум q на каждом шаге k из распределения $p_{\theta}(z_{k-1} | z_k)$. Более подробно принцип работы диффузионных моделей описан в работе [43].

Следуя [18], для обучения диффузионной модели CasFormer использовался оптимизатор AdamW с коэффициентом скорости обучения, линейно меняющимся в диапазоне от 10^{-3} до 10^{-6} на протяжении 300 тыс. итераций. Целевой функцией минимизации является:

$$L_{diff} = \|N(0, 1) - \Psi_{\theta}(z_k, k, z_{cond})\|, \quad (8)$$

где $z_k = \sqrt{\alpha_k} z_{k-1} + \sqrt{1 - \alpha_k} \sigma$, $z_{cond} = E(\bar{I}_0)$.

Для обучения всех описанных моделей использовалось 8 графических ускорителей NVIDIA A6000 с объемом видеопамати каждого ускорителя 48 Гб. Количество примеров, демонстрируемое ИНС за одну обучающую итерацию (batch size) в расчете на одну видеокарту, принимался равным 20, 4, 3 и 8 для моделей EvolutionNet, NowcastNet, автоэнкодера и CasFormer соответственно. Итоговый расчет прогноза с использованием обученного алгоритма осуществлялся на видеокарте RTX2080 с объемом видеопамати 8 Гб. Полное время расчета карты осадков для рассматриваемой территории занимает менее 10 минут, что позволяет использовать данный алгоритм для оперативного расчета прогноза осадков.

5. Оценка результатов. Для оценки точности полученных результатов рассчитывался ряд оценочных метрик с использованием предварительно сформированной тестовой выборки, не использовавшейся при обучении или настройке ИНС. При оценке качества прогноза рассматривался ряд показателей. В частности, показатель корня среднеквадратичной ошибки (Root Mean Square Error, RMSE) и коэффициент корреляции Пирсона (Pearson Correlation Coefficient, PCC) позволяют оценить близость спрогнозированных значений интенсивностей осадков к их фактическим наблюдениям для всех пикселей изображения:

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}, \quad (9)$$

$$\text{PCC} = \frac{\sum_{i=1}^n (x_i - \langle x \rangle)(y_i - \langle y \rangle)}{\sqrt{\sum_{i=1}^n (x_i - \langle x \rangle)^2 \sum_{i=1}^n (y_i - \langle y \rangle)^2}}, \quad (10)$$

где x_i и y_i – спрогнозированный и эталонный пиксели, n – общее количество пикселей одного изображения (512×512), $\langle x \rangle$ – среднее значение по всему изображению.

Бинарные показатели качества прогноза, такие как коэффициент ложного обнаружения (False Alarm Ratio, FAR), вероятность обнаружения (Probability of Detection, POD) и критический индекс успешности (Critical Success Index, CSI), лучше подходят для оценки способности модели прогнозировать наличие или отсутствие осадков в каждом пикселе. Для получения бинарных изображений применялось пороговое значение 0.1 мм/ч. Чтобы оценить точность прогноза осадков средней и высокой интенсивности выше 5 мм/ч, в работе дополнительно рассматривается метрика $\text{CSI}_{5\text{мм/ч}}$ с соответствующим пороговым значением интенсивности.

$$\text{FAR} = \text{FP} / (\text{TP} + \text{FP}), \quad (11)$$

$$\text{CSI} = \text{TP} / (\text{TP} + \text{FP} + \text{FN}), \quad (12)$$

где FP – количество ложноположительных, TP – истинно положительных и FN – ложноотрицательных случаев (к положительному классу относятся пиксели, где присутствуют осадки).

Для оценки структурного сходства спрогнозированных и эталонных изображений использовались показатель соотношения сигнал/шум (Peak-to-Signal Noise Ratio, PSNR) и индекс структурной схожести (Structural Similarity Index, SSIM).

$$\text{PSNR} = 20 \log_{10}(y^{\max}) - 10 \log_{10}(\text{MSE}), \quad (13)$$

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (14)$$

где $y^{\max}$ – максимально возможное значение пикселя, равное 50 мм/ч, μ_x и μ_y – средние значения для спрогнозированного и эталонного изображения, σ_x и σ_y – их дисперсия, σ_{xy} – ковариация, $c_1 = (0.01L)^2$, $c_2 = (0.03L)^2$, $L = 2^b - 1 = 255$ – динамический диапазон интенсивности пикселей, b – разрядность (8 бит на пиксель).

Ниже перечислены модели, для которых проводилось сравнение. Выбор этих моделей основывается на том, что они являются последними достижениями в области краткосрочного прогноза осадков и, по утверждениям их авторов, обладают высокой точностью. Помимо этого, архитектуры этих моделей либо частично использовались в настоящей работе в качестве основы для создания модернизированного алгоритма, либо схожи по принципу действия и работы.

а. Оригинальная модель NowcastNet-GAN [26]. Данная статистическая модель имеет в своем составе два дискриминатора для контроля пространственной и временной компонент прогноза и демонстрирует одни из лучших результатов [24, 29].

б. Модель генератора NowcastNet из работы [26]. С целью сравнения детерминированного и статистического подхода, авторы отдельно обучили модель сверточного генератора с использованием детерминированной функции потерь на основе суммы среднеквадратичной и средней абсолютной ошибок (формула 6). Как и в оригинальной работе, в состав модели включена вспомогательная ИНС EvolutionNet.

в. Детерминированная трансформерная модель Earthformer [16]. Исходя из сравнительного анализа различных архитектур [7], обладающая одним из лучших результатов в задаче прогноза осадков.

г. Диффузионная модель CasFormer + Earthformer, описанная в работе [18]. В оригинальной работе авторы использовали данную модель в качестве демонстрации работы своей модели CasCast [15], предназначенной для улучшения детерминированных результатов прогноза.

В рамках эксперимента все перечисленные модели были предварительно обучены на полученном авторами обучающем наборе

данных в соответствии с описаниями в оригинальных работах. Тестовая выборка, полученная авторами, состоит из 3.6 тысяч примеров последовательностей изображений интенсивностей осадков на протяжении с 2022 по 2024 гг. на территории Дальневосточного региона. В таблице 1 представлены значения метрик, рассчитанные для двенадцатого выходного кадра каждой модели, что соответствует прогнозу на 120 минут.

Таблица 1. Сравнение моделей для прогноза осадков на 120 минут (кадр № 12)

Model	RMSE	PCC	POD	FAR	CSI	CSI _{5 мм/ч}	SSIM	PSNR
Earthformer	1.09	0.56	0.80	0.34	0.57	0.57	0.88	34.49
NowcastNet	0.88	0.73	0.73	0.13	0.68	0.70	0.92	36.73
NowcastNet-GAN	0.94	0.71	0.77	0.29	0.63	0.71	0.90	35.87
Earthformer + Casformer	1.20	0.51	0.74	0.38	0.52	0.60	0.85	32.95
RainCast	0.88	0.75	0.78	0.27	0.65	0.72	0.91	36.63

На рисунке 2 представлены значения метрик для каждого спрогнозированного кадра.

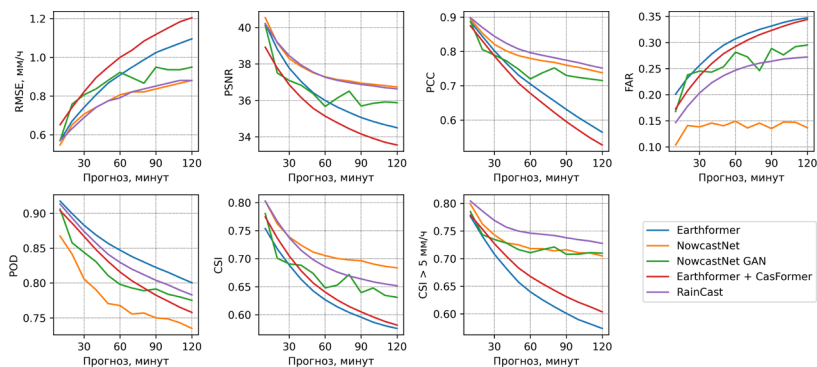


Рис. 2. Изменение показателей качества прогнозов тестируемых моделей с течением времени

Анализ полученных результатов позволяет сделать вывод, что наилучшие показатели точности были достигнуты детерминированной моделью NowcastNet и предложенной в данной работе моделью RainCast. На рисунках 3 и 4 приведены примеры прогноза осадков, сделанные с использованием тестируемых моделей.

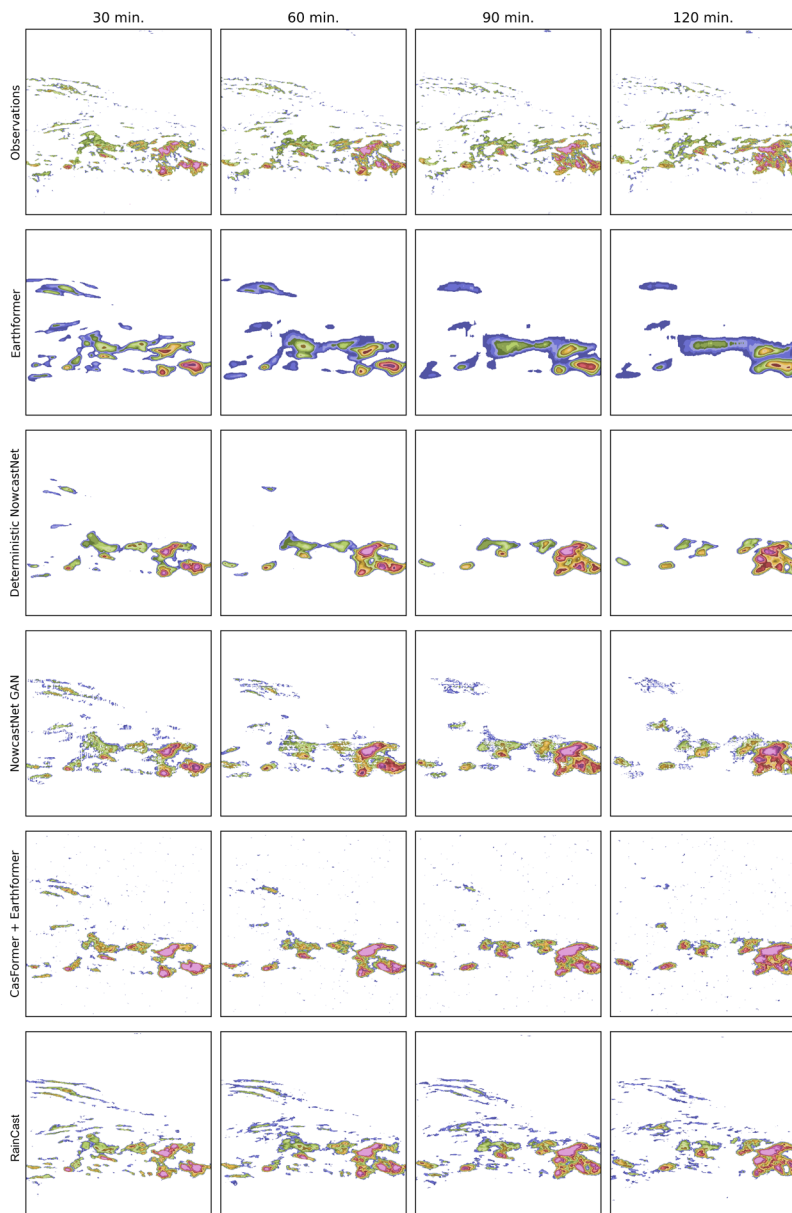


Рис. 3. Пример прогноза осадков от 01-08-2023 16:20 UTC, сделанный с использованием тестируемых моделей

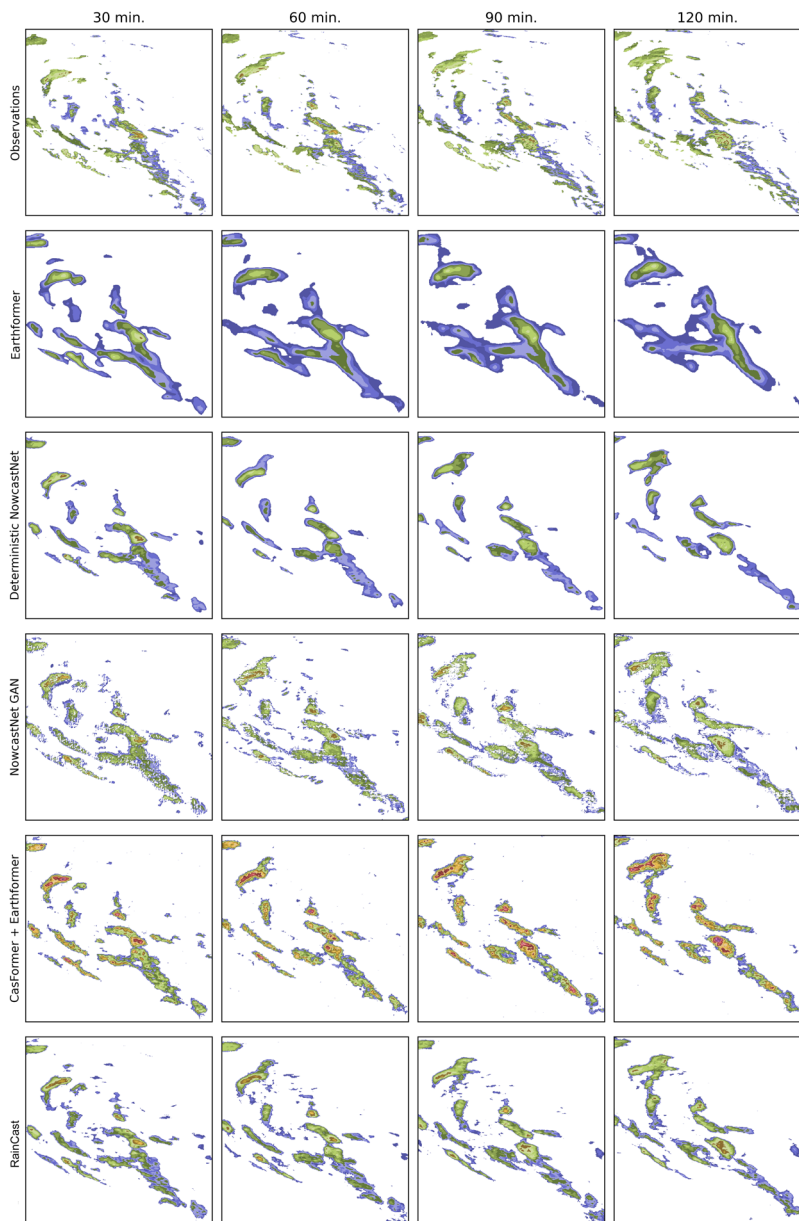


Рис. 4. Пример прогноза осадков от 10-12-2023 6:20 UTC, сделанный с использованием тестируемых моделей

Прогноз модели RainCast имеет наилучшие показатели среди всех моделей по метрикам PCC и $CSI_{5\text{мм/ч}}$, равным 0.75 и 0.72 соответственно. По таким показателям, как FAR, SSIM и PSNR, он незначительно уступает прогнозу детерминированной модели NowcastNet, а по показателю RMSE эти модели близки друг к другу. При этом показатель POD для NowcastNet имеет наихудшее значение, равное 0.73, в то время как наилучшие результаты по этому показателю были достигнуты предложенной гибридной моделью RainCast и детерминированной моделью Earthformer (0.78 и 0.80 соответственно).

Исходя из полученных данных и визуального анализа, можно сделать вывод о том, что детерминированная модель NowcastNet имеет склонность к недооценке осадков, тогда как Earthformer, напротив, существенно переоценивает их количество. В частности, показатель FAR для NowcastNet имеет значение 0.13, а для Earthformer он равен 0.34. На рисунках 3 и 4 данная ситуация также хорошо прослеживается.

Если сравнить прогноз двух лучших моделей – детерминированного NowcastNet и гибридного RainCast, можно отметить, что для первой модели наблюдается существенное снижение детализации при прогнозе локальной облачности (рисунок 5).

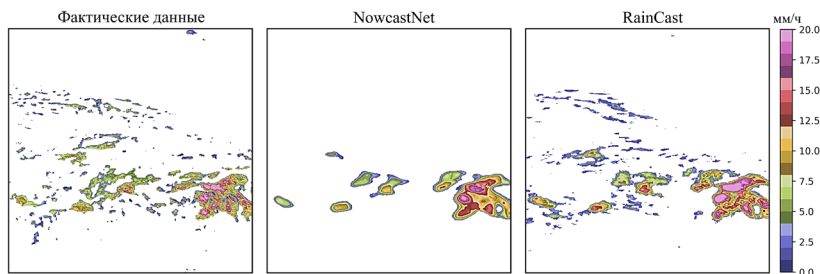


Рис. 5. Сравнение моделей NowcastNet и RainCast, прогноз на 120 минут

Одновременно с этим, несмотря на лучшее визуальное соответствие полей осадков у модели RainCast, значения рассчитанных показателей качества у нее несколько ниже. Это связано с тем, что поля осадков, спрогнозированные с помощью этой модели, немного смещены от их фактического положения. Как следствие, это приводит к ухудшению показателей, учитывающих подобные различия. Если обратить внимание на показатель $CSI_{5\text{мм/ч}}$, то он является наилучшим среди всех моделей, поскольку в этом случае оцениваются только средние и крупные осадки.

По большинству показателей (таблица 1) и исходя из визуального анализа прогнозов среди тестируемых моделей на рисунках 3 и 4 наименее точный прогноз имеет модель Earthformer, что проявляется в виде большого размытия и низкой детализации изображений. Прогноз модели NowcastNet-GAN имеет некоторое количество артефактов, проявляющихся в виде зернистости, но в целом локальная облачность спрогнозирована достаточно точно. Прогноз, полученный с помощью гибридной модели RainCast также обладает большим количеством деталей и высокой резкостью изображений, как и в случае с NowcastNet-GAN. Прогноз детерминированной модели NowcastNet хорошо совпадает с фактическими данными для крупных облачных массивов, что подтверждают рассчитанные оценки точности из таблицы 1. Однако, как было отмечено выше, детерминированные модели гораздо хуже справляются с прогнозом локальной облачности, что видно на примере модели Earthformer. Таким образом, принимая во внимание результаты проведенного численного и визуального анализа прогнозов, можно сделать вывод о том, что наиболее близкие к фактическим наблюдениям результаты имеет предложенная гибридная модель RainCast.

8. Заключение. В работе предложен алгоритм RainCast краткосрочного прогноза интенсивности осадков на срок до 2-х часов с использованием наблюдений геостационарного КА Himawari-8/9. В основе алгоритма лежит подход, суть которого заключается в получении предварительного прогноза с помощью детерминированной модели и его последующего уточнения с использованием диффузионной ИНС. Особенностью предложенного алгоритма является использование данных геостационарного спутникового прибора в региональных условиях при отсутствии наземных источников информации об осадках. Алгоритм построен на основе актуальных решений и объединяет в себе преимущества детерминированного и статистического прогноза. В частности, в качестве источника детерминированного прогноза использовалась физически обусловленная модель NowcastNet, обученная с использованием детерминированного подхода. Чтобы повысить детализацию прогноза, использовалась статистическая диффузионная модель CasFormer. Для обучения и тестирования вышеуказанных моделей были сформированы соответствующие наборы данных с применением разработанного ранее алгоритма для оценки интенсивности осадков по данным прибора АНН Himawari-8/9.

Сравнение разработанного алгоритма RainCast с аналогичными детерминированными и статистическими моделями на основе GAN и диффузионного подхода показало, что статистический подход, основанный на использовании GAN и диффузий, позволяет существенно улучшить качество прогнозов, выполненных с помощью детерминированных моделей, особенно для локальной осадкообразующей облачности. Однако процедура обучения этих моделей достаточно трудоемка. В частности, для ИНС на основе GAN требуется тонкая настройка гиперпараметров модели в процессе обучения. Диффузионные модели нуждаются в применении дополнительной ИНС, предназначенной для сжатия данных.

При этом статистические модели позволяют лишь приблизительно оценить расположение локальной облачности. Как следствие, несмотря на заметно более высокое качество статистического прогноза при визуальном сравнении изображений, значения численных показателей качества, таких как RMSE, POD, FAR и др., могут ухудшиться в сравнении с детерминированным прогнозом. Это объясняется тем, что ввиду малых размеров локальной облачности, даже небольшая ошибка в прогнозе ее положения приводит к снижению показателей качества. Чтобы решить эту проблему, в работе использовался показатель CSI для значений интенсивности осадков выше 5 мм/ч при рассмотрении основных облачных массивов на мезомасштабном уровне. Для более объективной численной оценки качества прогноза необходимо исследовать другие показатели.

Дополнительное сравнение двух актуальных моделей, сверточной NowcastNet и трансформерной Earthformer, показало, что интеграция физических принципов в ИНС положительно влияет на точность прогноза осадков. В частности, в NowcastNet используется нейросетевая реализация закона непрерывности потока в виде подсети EvolutionNet. При этом, как было отмечено ранее, трансформеры и рекуррентные сети более эффективны с точки зрения обработки временных последовательностей, чем сверточные ИНС. Поэтому в дальнейших работах целесообразно исследовать возможности по улучшению качества прогнозов путем реализации других физических принципов, а также использования данных численного моделирования в качестве начальных условий для прогноза.

Разработанный алгоритм RainCast предназначен для прогноза осадков на территории Дальневосточного региона по данным КА Himwari-8/9. Однако он может быть адаптирован для использования измерений других космических систем, таких как КА серии «Электро-

Л» и «Арктика-М». Для этого требуется формирование соответствующих обучающих наборов данных, учитывающих особенности измерений данных КА. Поскольку входными данными алгоритма RainCast являются изображения интенсивностей осадков, для их получения требуется разработка соответствующих алгоритмов оценки осадков, учитывающих территориально-климатические особенности рассматриваемого региона.

Литература

1. Browning K.A. Review Lecture: Local weather forecasting // Proc. of the Royal Society A. Mathematical and Physical Sciences. 1980. vol. 371. pp. 179–211. DOI: 10.1098/rspa.1980.0076.
2. Киктев Д.Б., Муравьев А.В., Смирнов А.В. Наукастинг метеорологических параметров и опасных явлений: опыт реализации и перспективы развития // Гидрометеорологические исследования и прогнозы. 2019. № 4(374). С. 92–111.
3. Observing Systems Capability Analysis and Review Tool // Geneva, Switzerland: World Meteorological Organization. 2021. URL: <http://www.wmo-sat.info/oscar/> (дата обращения: 01.04.2021).
4. Simonenko E. V., Chudin A. O., Davidenko A. N. The differential method for calculation of cloud motion vectors // Russian Meteorology and Hydrology. 2017. vol. 42. no. 3. pp. 159–167. DOI: 10.3103/S1068373917030037.
5. Marrocu M., Massidda L. Performance comparison between deep learning and optical flow-based techniques for nowcast precipitation from radar images // Forecasting. 2020. vol. 2. no. 2. pp. 194–210. DOI: 10.3390/forecast2020011.
6. Prudden R., Adams S., Kangin D., Robinson N., Ravuri S., Mohamed S., Arribas A. A review of radar-based nowcasting of precipitation and applicable machine learning techniques // arXiv preprint arXiv:2005.04988. 2020. DOI: 10.48550/arXiv.2005.04988.
7. An S., Oh T.-J., Sohn E., Kim D. Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting // Expert Systems with Applications. 2025. vol. 268. pp. 126301. DOI: 10.1016/j.eswa.2024.126301.
8. Tao W. K., Moncrieff M. W. Multiscale cloud system modeling // Reviews of Geophysics. 2009. vol. 47. no. 4. DOI: 10.1029/2008RG000276.
9. Tan J., Huang Q., Chen S. Deep learning model based on multi-scale feature fusion for precipitation nowcasting // Geoscientific Model Development. 2024. vol. 17. no. 1. pp. 53–69. DOI: 10.5194/gmd-17-53-2024.
10. Kong D., Zhi X., Ji Y., Yang C., Wang Y., Tian Y., Li G., Zeng X. Precipitation Nowcasting Based on Deep Learning over Guizhou, China // Atmosphere. 2023. vol. 14. no. 5. DOI: 10.3390/atmos14050807.
11. Shi X., Chen Z., Wang H., Yeung D.-Y., Wong W., Woo W. Convolutional LSTM network: A machine learning approach for precipitation nowcasting // Advances in Neural Information Processing Systems. 2015. vol. 28. pp. 1–9.
12. Shi X., Gao Z., Lausen L., Wang H., Yeung D.-Y., Wong W., Woo W. Deep learning for precipitation nowcasting: A benchmark and a new model // Advances in Neural Information Processing Systems: 31st Conference on Neural Information Processing Systems (NIPS 2017). 2017. vol. 30.
13. Han K., Wang Y., Chen H., Chen X., Guo J., Liu Z. A Survey on Vision Transformer // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2023. vol. 45. no. 1. pp. 87–110. DOI: 10.1109/TPAMI.2022.3152247.

14. Wang Y., Wu H., Zhang J., Gao Z., Wang J., Philip S.Y., Long M. PredRNN: A Recurrent Neural Network for Spatiotemporal Predictive Learning // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2023. vol. 45. no. 2. pp. 2208–2225. DOI: 10.1109/TPAMI.2022.3165153.
15. Wu H., Yao Z., Wang J., Long M. MotionRNN: A Flexible Model for Video Prediction with Spacetime-Varying Motions // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA, 2021. pp. 15430–15439. DOI: 10.1109/CVPR46437.2021.01518.
16. Gao Z., Shi X., Wang H., Zhu Y., Wang Y.B., Li M., Yeung D.Y. Earthformer: Exploring space-time transformers for earth system forecasting // Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems (NeurIPS 2022). 2022. vol. 35. pp. 25390–25403.
17. Bai C., Sun F., Zhang J., Song Y., Chen S. Rainformer: Features extraction balanced network for radar-based precipitation nowcasting // IEEE Geoscience and Remote Sensing Letters. 2022. vol. 19. pp. 1–5. DOI: 10.1109/LGRS.2022.3162882.
18. Gong J., Bai L., Ye P., Xu W., Liu N., Dai J., Yang X., Ouyang W. CasCast: Skillful high-resolution precipitation nowcasting via cascaded modelling // arXiv preprint arXiv:2402.04290. 2024. DOI: 10.48550/arXiv.2402.04290.
19. Wen P., Bai L., He M., Filippi P., Zhang F., Bishop T.F., Wang Z., Hu K. DuoCast: Duo-Probabilistic Meteorology-Aware Model for Extended Precipitation Nowcasting // arXiv preprint arXiv:2412.01091. 2024. DOI: 10.48550/arXiv.2412.01091.
20. Ravuri S., Lenc K., Willson M., Kangin D., Lam R., Mirowski P., Fitzsimons M., Athanassiadou M., Kashem S., Madge S., Prudden R., Mandhane A., Clark A., Brock A., Simonyan K., Hadsell R., Robinson N., Clancy E., Arribas A., Mohamed S. Skillful precipitation nowcasting using deep generative models of radar // Nature. 2021. vol. 597. pp. 672–677. DOI: 10.1038/s41586-021-03854-z.
21. Кучма М.О., Мальковский С.И., Андреев А.И., Блощинский В.Д. Нейросетевые методы наукастинга осадков: обзор и апробация существующих решений // Современные проблемы дистанционного зондирования Земли из космоса. 2023. Т. 20. № 6. С. 35–48. DOI: 10.21046/2070-7401-2023-20-6-35-48.
22. Nai C., Pan B., Chen X., Tang Q., Ni G., Duan Q., Lu B., Xiao Z., Liu X. Reliable precipitation nowcasting using probabilistic diffusion models // Environmental Research Letters. 2024. vol. 19. no. 3. DOI: 10.1088/1748-9326/ad2891.
23. Yin J., Meo C., Cher Z.B., Lica M., Wang Y., Imhoff R., Uijlenhoet R., Dauwels J. Precipitation Nowcasting Using Physics Informed Discriminator Generative Models // 32nd European Signal Processing Conference (EUSIPCO). Lyon, France, 2024. pp. 967–971. DOI: 10.23919/EUSIPCO63174.2024.10715141.
24. Wang R., Fung J.C.H., Lau A.K.H. Skillful precipitation nowcasting using physical-driven diffusion networks // Geophysical Research Letters. 2024. vol. 51. no. 24. DOI: 10.1029/2024GL110832.
25. Yu D., Li X., Ye Y., Zhang B., Luo C., Dai K., Wang R., Chen X. DiffCast: A Unified Framework via Residual Diffusion for Precipitation Nowcasting // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, 2024. pp. 27758–27767. DOI: 10.1109/CVPR52733.2024.02622.
26. Zhang Y., Long M., Chen K., Xing L., Jin R., Jordan M.I., Wang J. Skillful nowcasting of extreme precipitation with NowcastNet // Nature. 2023. vol. 619. pp. 526–532. DOI: 10.1038/s41586-023-06184-4.
27. Kim W., Jeong C.H., Kim S. Improvements in deep learning-based precipitation nowcasting using major atmospheric factors with radar rain rate // Computers & Geosciences. 2024. vol. 184. DOI: 10.1016/j.cageo.2024.105529.
28. Rahimi R., Ravirathinam P., Ebtehaj A., Behrang A., Tan J., Kumar V. Global precipitation nowcasting of Integrated Multi-satellitE Retrievals for GPM: A U-Net

- convolutional LSTM architecture // *Journal of Hydrometeorology*. 2024. vol. 25. no. 6. pp. 947–963. DOI: 10.1175/JHM-D-23-0119.1.
29. Das P., Posch A., Barber N., Hicks M., Duffy K., Vandal T., Singh D., Werkhoven K., van Werkhoven K., Ganguly A.R. Hybrid physics-AI outperforms numerical weather prediction for extreme precipitation nowcasting // *NPJ Climate and Atmospheric Science*. 2024. vol. 7. no. 1. DOI: 10.1038/s41612-024-00834-8.
30. Xu Q., Shi Y., Bamwig R., Tuo Y., Ludwig R., Zhu X.X. Physics-aware machine learning revolutionizes scientific paradigm for machine learning and process-based hydrology // *arXiv preprint arXiv:2310.05227*. 2023. DOI: 10.48550/arXiv.2310.05227.
31. Leinonen J., Hamann U., Nerini D., Germann U., Franch G. Latent diffusion models for generative precipitation nowcasting with accurate uncertainty quantification // *arXiv preprint arXiv:2304.12891*. 2023. DOI: 10.48550/arXiv.2304.12891.
32. Chen T. On the importance of noise scheduling for diffusion models // *arXiv preprint arXiv:2301.10972*. 2023. DOI: 10.48550/arXiv.2301.10972.
33. Veillette M., Samsi S., Mattioli C. SEVIR: A storm event imagery dataset for deep learning applications in radar and satellite meteorology // *Advances in Neural Information Processing Systems: 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*. 2020. vol. 33. pp. 22009–22019.
34. Бикбулатова Г.Г. Исследование распределения атмосферного увлажнения на территории Сибири и Дальнего Востока // *Омский научный вестник*. 2011. Т. 1(104). С. 232–235.
35. Андреев А.И., Филей А.А., Мальковский С.И. Метод количественной оценки осадков на основе измерений Himawari-8/9 // *Современные проблемы дистанционного зондирования Земли из космоса*. 2024. Т. 21. № 5. С. 20–35. DOI: 10.21046/2070-7401-2024-21-5-20-35.
36. Huffman G.J., Bolvin D.T., Braithwaite D., Hsu K.L., Joyce R.J., Kidd C., Nelkin E.J., Sorooshian S., Stocker E.F., Tan J., Wolff D.B., Xie P. Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG) // *Satellite Precipitation Measurement // Advances in Global Change Research*. 2020. vol. 67. pp. 343–353. Springer, Cham. DOI: 10.1007/978-3-030-24568-9_19.
37. Kingma D.P., Ba J. Adam: A method for stochastic optimization // *arXiv preprint arXiv:1412.6980*. 2017. DOI: 10.48550/arXiv.1412.6980.
38. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models // *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022. pp. 10674–10685. DOI: 10.1109/CVPR52688.2022.01042.
39. Esser P., Rombach R., Ommer B. Taming Transformers for High-Resolution Image Synthesis // *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. pp. 12868–12878. DOI: 10.1109/CVPR46437.2021.01268.
40. Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets // *Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems*. 2014. vol. 27. pp. 2672–2680.
41. Kingma D.P., Welling M. Auto-encoding variational bayes // *arXiv preprint arXiv:1312.6114*. 2013. DOI: 10.48550/arXiv.1312.6114.
42. Peebles W., Xie S. Scalable Diffusion Models with Transformers // *IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France, 2023. pp. 4172–4182. DOI: 10.1109/ICCV51070.2023.00387.
43. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models // *Advances in neural information processing systems*. 2020. vol. 33. pp. 6840–6851.

Андреев Александр Иванович — научный сотрудник, лаборатория информационных и вычислительных систем, Вычислительный центр Дальневосточного отделения Российской академии наук; младший научный сотрудник, отдел разработки и внедрения НИОКР, ДЦ ФГБУ «НИЦ «Планета». Область научных интересов: компьютерная оптика, обработка изображений, машинное обучение, дистанционное зондирование Земли. Число научных публикаций — 51. a.andreev@dvrcpod.ru; улица Ленина, 18, 680000, Хабаровск, Россия; р.т.: +7(4212)214-311.

Кучма Михаил Олегович — канд. техн. наук, научный сотрудник, лаборатория информационных и вычислительных систем, Вычислительный центр Дальневосточного отделения Российской академии наук; младший научный сотрудник, отдел разработки и внедрения НИОКР, ДЦ ФГБУ «НИЦ «Планета». Область научных интересов: дистанционное зондирование Земли, обработка данных, геоданные, машинное обучение, искусственный интеллект. Число научных публикаций — 55. m.kuchma@dvrcpod.ru; улица Ленина, 18, 680000, Хабаровск, Россия; р.т.: +7(4212)214-311.

Мальковский Сергей Иванович — канд. техн. наук, старший научный сотрудник, лаборатория информационных и вычислительных систем, Вычислительный центр Дальневосточного отделения Российской академии наук. Область научных интересов: высокопроизводительные вычисления, гибридные вычислительные системы, GRID, параллельные алгоритмы, искусственный интеллект, компьютерное моделирование. Число научных публикаций — 90. sergey.malkovsky@ccfebras.ru; улица Ким Ю Чена, 65, 680000, Хабаровск, Россия; р.т.: +7(4212)703-913.

Филей Андрей Александрович — канд. физ.-мат. наук, старший научный сотрудник, лаборатория информационных и вычислительных систем, Вычислительный центр Дальневосточного отделения Российской академии наук; научный сотрудник, отдел разработки и внедрения НИОКР, ДЦ ФГБУ «НИЦ «Планета». Область научных интересов: дистанционное зондирование Земли, обработка данных, физическое моделирование, машинное обучение. Число научных публикаций — 55. andreymf61@mail.ru; улица Ким Ю Чена, 65, 680000, Хабаровск, Россия; р.т.: +7(4212)214-311.

Поддержка исследований. Исследование выполнено за счет гранта Российского научного фонда № 23-77-00011, <https://rscf.ru/project/23-77-00011/>. В работе использовались ресурсы Центра коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений Института космических исследований Российской академии наук для решения задач изучения и мониторинга окружающей среды (ЦКП «ИКИ-Мониторинг»).

A. ANDREEV, M. KUCHMA, S. MALKOVSKY, A. FILEI
**RAINCAS: A HYBRID PRECIPITATION NOWCASTING
ALGORITHM USING THE HIMAWARI-8/9 SATELLITE
MEASUREMENTS**

Andreev A., Kuchma M., Malkovsky S., Filei A. RainCast: A Hybrid Precipitation Nowcasting Algorithm Using the Himawari-8/9 Satellite Measurements.

Abstract. This paper proposes an algorithm for short-term rain rate forecasting, RainCast (Rain Rate NowCast), for up to two hours. This area of meteorology, known as 'nowcasting', is one of the most important tools in many areas of human activity. However, its availability may be severely limited by existing ground infrastructure. In this paper, the authors aim to create a precipitation forecasting algorithm for one such territory using the Asia-Pacific region as an example, based on satellite measurements from the Himawari-8/9 spacecraft. The algorithm combines the advantages of deterministic and statistical approaches to solve the forecasting problem and is based on two neural networks. The first model, a modified version of the physically constrained neural network NowcastNet, generates a preliminary forecast of the general direction of precipitation movement at the mesoscale level. The second model, based on the CasFormer architecture, employs diffusion methods to post-process the initial forecast, refining fine-scale details. The resulting hybrid algorithm, named RainCast, enables short-term precipitation forecasting (up to 2 hours) with high spatiotemporal resolution (2 km, updated every 10 minutes), utilizing solely infrared satellite measurements. Satellite data are converted into precipitation intensity using the algorithm previously developed by the authors. Based on precipitation maps, training, validation, and test datasets were compiled for the algorithm development and forecast quality assessment. The proposed RainCast algorithm was trained on these datasets and compared with other state-of-the-art solutions such as NowcastNet, Casformer, and Earthformer. Analysis of performance metrics demonstrated that the hybrid RainCast algorithm achieves comparable accuracy. For a 2-hour forecast, the Root Mean Square Error (RMSE) was 0.88, the Probability of Detection (POD) was 0.78, the Pearson Correlation Coefficient (PCC) was 0.75, the Structural Similarity Index Measure (SSIM) was 0.91, and the Peak Signal-to-Noise Ratio (PSNR) was 36.63. Visual analysis of the forecasts confirmed that RainCast produces results closest to actual observations, primarily due to the diffusion model's ability to refine fine-scale spatial and temporal precipitation patterns.

Keywords: nowcasting, precipitation, rain rate, rainfall forecasting, Himawari, RainCast, diffusion models.

References

1. Browning K.A. Review Lecture: Local weather forecasting // Proc. of the Royal Society A. Mathematical and Physical Sciences. 1980. vol. 371. pp. 179–211. DOI: 10.1098/rspa.1980.0076.
2. Kiktev D.B., Muravyov A.V., Smirnov A.V. [Nowcasting of meteorological parameters and hazardous phenomena: implementation experience and development prospects]. *Gidrometeorologicheskie issledovaniya i prognozy – Hydrometeorological research and forecasts*. 2019. no. 4(374). pp. 92–111. (In Russ.).
3. Observing Systems Capability Analysis and Review Tool. Geneva, Switzerland: World Meteorological Organization. 2021. Available at: <http://www.wmo-sat.info/oscar/> (data accessed 01.04.2021).

4. Simonenko E. V., Chudin A. O., Davidenko A. N. The differential method for calculation of cloud motion vectors. *Russian Meteorology and Hydrology*. 2017. vol. 42. no. 3. pp. 159–167. DOI: 10.3103/S1068373917030037.
5. Marrocu M., Massidda L. Performance comparison between deep learning and optical flow-based techniques for nowcast precipitation from radar images. *Forecasting*. 2020. vol. 2. no. 2. pp. 194–210. DOI: 10.3390/forecast2020011.
6. Prudden R., Adams S., Kangin D., Robinson N., Ravuri S., Mohamed S., Arribas A. A review of radar-based nowcasting of precipitation and applicable machine learning techniques. *arXiv preprint arXiv:2005.04988*. 2020. DOI: 10.48550/arXiv.2005.04988.
7. An S., Oh T.-J., Sohn E., Kim D. Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting. *Expert Systems with Applications*. 2025. vol. 268. pp. 126301. DOI: 10.1016/j.eswa.2024.126301.
8. Tao W. K., Moncrieff M. W. Multiscale cloud system modeling. *Reviews of Geophysics*. 2009. vol. 47. no. 4. DOI: 10.1029/2008RG000276.
9. Tan J., Huang Q., Chen S. Deep learning model based on multi-scale feature fusion for precipitation nowcasting. *Geoscientific Model Development*. 2024. vol. 17. no. 1. pp. 53–69. DOI: 10.5194/gmd-17-53-2024.
10. Kong D., Zhi X., Ji Y., Yang C., Wang Y., Tian Y., Li G., Zeng X. Precipitation Nowcasting Based on Deep Learning over Guizhou, China. *Atmosphere*. 2023. vol. 14. no. 5. DOI: 10.3390/atmos14050807.
11. Shi X., Chen Z., Wang H., Yeung D.-Y., Wong W., Woo W. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems*. 2015. vol. 28. pp. 1–9.
12. Shi X., Gao Z., Lausen L., Wang H., Yeung D.-Y., Wong W., Woo W. Deep learning for precipitation nowcasting: A benchmark and a new model. *Advances in Neural Information Processing Systems: 31st Conference on Neural Information Processing Systems (NIPS 2017)*. 2017. vol. 30.
13. Han K., Wang Y., Chen H., Chen X., Guo J., Liu Z. A Survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2023. vol. 45. no. 1. pp. 87–110. DOI: 10.1109/TPAMI.2022.3152247.
14. Wang Y., Wu H., Zhang J., Gao Z., Wang J., Philip S.Y., Long M. PredRNN: A Recurrent Neural Network for Spatiotemporal Predictive Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2023. vol. 45. no. 2. pp. 2208–2225. DOI: 10.1109/TPAMI.2022.3165153.
15. Wu H., Yao Z., Wang J., Long M. MotionRNN: A Flexible Model for Video Prediction with Spacetime-Varying Motions. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA, 2021. pp. 15430–15439. DOI: 10.1109/CVPR46437.2021.01518.
16. Gao Z., Shi X., Wang H., Zhu Y., Wang Y.B., Li M., Yeung D.Y. Earthformer: Exploring space-time transformers for earth system forecasting. *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems (NeurIPS 2022)*. 2022. vol. 35. pp. 25390–25403.
17. Bai C., Sun F., Zhang J., Song Y., Chen S. Rainformer: Features extraction balanced network for radar-based precipitation nowcasting. *IEEE Geoscience and Remote Sensing Letters*. 2022. vol. 19. pp. 1–5. DOI: 10.1109/LGRS.2022.3162882.
18. Gong J., Bai L., Ye P., Xu W., Liu N., Dai J., Yang X., Ouyang W. CasCast: Skillful high-resolution precipitation nowcasting via cascaded modelling. *arXiv preprint arXiv:2402.04290*. 2024. DOI: 10.48550/arXiv.2402.04290.
19. Wen P., Bai L., He M., Filippi P., Zhang F., Bishop T.F., Wang Z., Hu K. DuoCast: Duo-Probabilistic Meteorology-Aware Model for Extended Precipitation Nowcasting. *arXiv preprint arXiv:2412.01091*. 2024. DOI: 10.48550/arXiv.2412.01091.

20. Ravuri S., Lenc K., Willson M., Kangin D., Lam R., Mirowski P., Fitzsimons M., Athanassiadou M., Kashem S., Madge S., Prudden R., Mandhane A., Clark A., Brock A., Simonyan K., Hadsell R., Robinson N., Clancy E., Arribas A., Mohamed S. Skilful precipitation nowcasting using deep generative models of radar. *Nature*. 2021. vol. 597. pp. 672–677. DOI: 10.1038/s41586-021-03854-z.
21. Kuchma M.O., Malkovsky S.I., Andreev A.I., Bloschchinskiy V.D. [Neural network approaches to precipitation nowcasting: A review and test of existing methods]. *Sovremennyye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa – Current problems in remote sensing of the Earth from space*. 2023. vol. 20. no. 6. pp. 35–48. DOI: 10.21046/2070-7401-2023-20-6-35-48. (In Russ.).
22. Nai C., Pan B., Chen X., Tang Q., Ni G., Duan Q., Lu B., Xiao Z., Liu X. Reliable precipitation nowcasting using probabilistic diffusion models. *Environmental Research Letters*. 2024. vol. 19. no. 3. DOI: 10.1088/1748-9326/ad2891.
23. Yin J., Meo C., Cher Z.B., Lica M., Wang Y., Imhoff R., Uijlenhoet R., Dauwels J. Precipitation Nowcasting Using Physics Informed Discriminator Generative Models. 32nd European Signal Processing Conference (EUSIPCO). Lyon, France, 2024. pp. 967–971. DOI: 10.23919/EUSIPCO63174.2024.10715141.
24. Wang R., Fung J.C.H., Lau A.K.H. Skillful precipitation nowcasting using physical-driven diffusion networks. *Geophysical Research Letters*. 2024. vol. 51. no. 24. DOI: 10.1029/2024GL110832.
25. Yu D., Li X., Ye Y., Zhang B., Luo C., Dai K., Wang R., Chen X. DiffCast: A Unified Framework via Residual Diffusion for Precipitation Nowcasting. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA, 2024. pp. 27758–27767. DOI: 10.1109/CVPR52733.2024.02622.
26. Zhang Y., Long M., Chen K., Xing L., Jin R., Jordan M.I., Wang J. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*. 2023. vol. 619. pp. 526–532. DOI: 10.1038/s41586-023-06184-4.
27. Kim W., Jeong C.H., Kim S. Improvements in deep learning-based precipitation nowcasting using major atmospheric factors with radar rain rate. *Computers & Geosciences*. 2024. vol. 184. DOI: 10.1016/j.cageo.2024.105529.
28. Rahimi R., Ravirathinam P., Ebtehaj A., Behrangi A., Tan J., Kumar V. Global precipitation nowcasting of Integrated Multi-satellitE Retrievals for GPM: A U-Net convolutional LSTM architecture. *Journal of Hydrometeorology*. 2024. vol. 25. no. 6. pp. 947–963. DOI: 10.1175/JHM-D-23-0119.1.
29. Das P., Posch A., Barber N., Hicks M., Duffy K., Vandal T., Singh D., Werkhoven K., van Werkhoven K., Ganguly A.R. Hybrid physics-AI outperforms numerical weather prediction for extreme precipitation nowcasting. *NPJ Climate and Atmospheric Science*. 2024. vol. 7. no. 1. DOI: 10.1038/s41612-024-00834-8.
30. Xu Q., Shi Y., Bamber J., Tuo Y., Ludwig R., Zhu X.X. Physics-aware machine learning revolutionizes scientific paradigm for machine learning and process-based hydrology. *arXiv preprint arXiv:2310.05227*. 2023. DOI: 10.48550/arXiv.2310.05227.
31. Leinonen J., Hamann U., Nerini D., Germann U., Franch G. Latent diffusion models for generative precipitation nowcasting with accurate uncertainty quantification. *arXiv preprint arXiv:2304.12891*. 2023. DOI: 10.48550/arXiv.2304.12891.
32. Chen T. On the importance of noise scheduling for diffusion models. *arXiv preprint arXiv:2301.10972*. 2023. DOI: 10.48550/arXiv.2301.10972.
33. Veillette M., Samsi S., Mattioli C. SEVIR: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. *Advances in Neural Information Processing Systems: 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*. 2020. vol. 33. pp. 22009–22019.

34. Bikbulatova G.G. [A study of the distribution of atmospheric hydration in Siberia and the Far East]. *Omskij nauchnyj vestnik – Omsk scientific messenger*. 2011. vol. 1(104). pp. 232–235. (In Russ.).
35. Andreev A.I., Filey A.A., Malkovsky S.I., Korolev S.P. [Method of quantitative rainfall estimation based on Himawari-8/9 measurements]. *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa – Current problems in remote sensing of the Earth from space*. 2024. vol. 21. no. 5. pp. 20–35. DOI: 10.21046/2070-7401-2024-21-5-20-35. (In Russ.).
36. Huffman G.J., Bolvin D.T., Braithwaite D., Hsu K.L., Joyce R.J., Kidd C., Nelkin E.J., Sorooshian S., Stocker E.F., Tan J., Wolff D.B., Xie P. Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). *Satellite Precipitation Measurement. Advances in Global Change Research*. 2020. vol. 67. pp. 343–353. Springer, Cham. DOI: 10.1007/978-3-030-24568-9_19.
37. Kingma D.P., Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. 2017. DOI: 10.48550/arXiv.1412.6980.
38. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022. pp. 10674–10685. DOI: 10.1109/CVPR52688.2022.01042.
39. Esser P., Rombach R., Ommer B. Taming Transformers for High-Resolution Image Synthesis. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. pp. 12868–12878. DOI: 10.1109/CVPR46437.2021.01268.
40. Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets. *Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems*. 2014. vol. 27. pp. 2672–2680.
41. Kingma D.P., Welling M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*. 2013. DOI: 10.48550/arXiv.1312.6114.
42. Peebles W., Xie S. Scalable Diffusion Models with Transformers. *IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France, 2023. pp. 4172–4182. DOI: 10.1109/ICCV51070.2023.00387.
43. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models. *Advances in neural information processing systems*. 2020. vol. 33. pp. 6840–6851.

Andreev Aleksandr — Researcher, Laboratory of information and computing systems, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences; Junior researcher, R& D development and implementation department, Far-Eastern Center of the Federal State Budgetary Institution «State Research Center of Space Hydrometeorology «Planeta». Research interests: computer optics, image processing, machine learning, remote sensing. The number of publications — 51. a.andreev@dvrcpod.ru; 18, Lenina St., 680000, Khabarovsk, Russia; office phone: +7(4212)214-311.

Kuchma Mikhail — Ph.D., Researcher, Laboratory of information and computing systems, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences; Junior researcher, R& D development and implementation department, Far-Eastern Center of the Federal State Budgetary Institution «State Research Center of Space Hydrometeorology «Planeta». Research interests: remote sensing, data processing, geodata, machine learning, artificial intelligence. The number of publications — 55. m.kuchma@dvrcpod.ru; 18, Lenina St., 680000, Khabarovsk, Russia; office phone: +7(4212)214-311.

Malkovsky Sergey — Ph.D., Senior researcher, Laboratory of information and computing systems, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences.

Research interests: high-performance computing, hybrid computing systems, GRID, parallel algorithms, artificial intelligence, computer modeling. The number of publications — 90. sergey.malkovsky@ccfebras.ru; 65, Kim Yu Chen St., 680000, Khabarovsk, Russia; office phone: +7(4212)703-913.

Filei Andrei — Ph.D., Senior researcher, Laboratory of information and computing systems, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences; Researcher, R&D development and implementation department, Far-Eastern Center of the Federal State Budgetary Institution «State Research Center of Space Hydrometeorology «Planeta». Research interests: remote sensing, data processing, physical modeling, machine learning. The number of publications — 55. andreyvm-61@mail.ru; 65, Kim Yu Chen St., 680000, Khabarovsk, Russia; office phone: +7(4212)214-311.

Acknowledgements. The research was supported by the Russian Science Foundation (grant 23-77-00011, <https://rscf.ru/en/project/23-77-00011/>). The resources of the Satellite Data Archiving, Processing, and Analysis System Common Use Center for Solving Problems of Environmental Research and Monitoring of the Space Research Institute of the Russian Academy of Sciences were used.

А.Г. ЛОСЕВ, И.Е. ПОПОВ, А.С. РЕЗНИКОВА
**ПОВЫШЕНИЕ РЕПРЕЗЕНТАТИВНОСТИ ОБУЧАЮЩЕГО
НАБОРА ДАННЫХ ЗА СЧЕТ ПРОСТРАНСТВЕННОЙ
БАЛАНСИРОВКИ**

Лосев А.Г., Попов И.Е., Резникова А.С. Повышение репрезентативности обучающего набора данных за счет пространственной балансировки.

Аннотация. В работе исследуются некоторые проблемы машинного обучения, связанные с анализом данных малого объема. А именно, решается задача формирования обучающего набора, обладающего высокой репрезентативностью в задачах многоклассовой классификации. Известно, что межклассовая и внутриклассовая несбалансированность, свойственная малым наборам данных, негативно влияет на качество работы алгоритмов. С целью их устранения в машинном обучении разработаны некоторые методы синтеза данных, дополняющие имеющийся набор и уравнивающие количество объектов каждого из классов. Однако такие методы не всегда решают проблему недостаточной репрезентативности. В данной статье предлагается метод построения репрезентативного обучающего набора данных за счет задания распределения, в наибольшей степени соответствующего действительности. Распределение формируется по каждому признаку в пределах информативных областей. Информативные области содержат характерные значения признаков, наиболее значимые для различения классов объектов. Предложенный метод построения областей основан на идее поэтапного расширения, сопровождающегося ростом информативности областей. При этом под информативностью понимается мера, отражающая то, насколько хорошо с помощью рассматриваемой области можно разделить объекты различных классов. С целью формирования дополняющего набора данных разработан метод генерации. В результате его применения дополняющий набор данных объединяется с исходным и образует в информативной области указанное распределение. Это распределение может быть задано либо на основе экспертных знаний о предметной области, если известно истинное распределение, либо получено в результате вычислительных экспериментов, направленных на поиск наиболее эффективного варианта. Применимость метода продемонстрирована на примере решения задачи определения уровня температурных аномалий молочных желёз. Показано, что для рассматриваемых температурных признаков характерно нормальное распределение. Повышение репрезентативности обучающего набора позволило обучить классический алгоритм классификации – логистическую регрессию – с точностью, сопоставимой с многослойной нейронной сетью. Такой подход к формированию обучающего набора данных открывает возможность создания более прозрачных и интерпретируемых систем искусственного интеллекта.

Ключевые слова: машинное обучение, малые наборы данных, репрезентативность данных, синтез данных, нейронные сети, логистическая регрессия.

1. Введение. Одной из ключевых проблем машинного обучения является разработка методов правильного формирования обучающего набора данных [1–3]. Особенно остро данный вопрос стоит при работе с "малыми" данными, хотя и в интеллектуальном анализе "больших" данных указанная задача тоже является актуальной.

В подавляющем большинстве исследований формирование обучающей выборки производится на основе "натурных" данных, т.е. данных, полученных в результате каких-то измерений и т.п. При этом известно, что натурные данные могут обладать следующими недостатками.

1. Наличие ошибок результатов измерений (пропуски, некорректные измерения и т.п. [4])

2. Наличие ошибок в метках [5,6]. Например, по мнению некоторых специалистов, в медицинских картах пациентов может содержаться более 30 процентов неточных диагнозов, что порождает неверные метки [7].

3. Наличие персональных данных. Последнее может породить трудности с распространением набора данных и валидацией полученных результатов.

4. Малый объем набора данных. В зависимости от предметной области сложность составления набора данных варьируется. Тем не менее исследователи сталкиваются с проблемой недостатка данных во многих дисциплинах [8–14]. В то время как большие данные подразумевают терабайты или петабайты объема, малые наборы данных не столь велики в объеме и обычно измеряются сотнями или тысячами записей. Однако понятие малых данных определяется не только объемом записей. Так, например, достаточно много данных может быть в национальной переписи населения, тем не менее такой набор данных не может быть определен понятием большие данные. Ключевые отличия малых наборов данных от больших заключаются в том, что малые данные накапливаются достаточно медленно, имеют однообразную структуру, низкую детализацию и поступают из одного или небольшого числа источников. В химии, физике и материаловедении трудности с получением больших выборок данных обусловлены высокими затратами и сложностью проведения экспериментов. Кроме того исследователи сталкиваются с малыми наборами данных и при решении задачи по обработке естественного языка в области программной инженерии (классификация отзывов о работе приложений, классификация комментариев к коду и т.д.). При этом объемы используемых наборов данных варьируются по объему в диапазоне от 300 до 65 000 экземпляров. В медицине и биологии редкость некоторых заболеваний и состояний также приводит к недостаточному количеству клинических случаев для обучения моделей машинного обучения и проведения статистически значимых исследований. Однако несмотря на свою редкость, эти явления могут существенно влиять на здоровье пациентов, поэтому их изучение становится необходимым. Так в одном из исследований, посвященном решению задачи прогнозирования

развития сахарного диабета, итак изначально малое количество равное 3000 записей, было снижено до 162 единиц после выбора полностью заполненных данных [38]. В исследовании по классификации МРТ снимков головного мозга количество записей тоже было мало и составило 380 единиц [8].

Следовательно, проблема малого набора данных охватывает многие сферы деятельности и имеет междисциплинарный характер. При этом работа с малым набором данных имеет свои особенности: высокая вероятность переобучения, повышенная чувствительность к шуму в данных, отсутствие разнообразия наблюдений, ограничения в проведении качественной валидации и тестирования.

5. Несбалансированность данных. Работа с наборами данных малого объема практически всегда обостряет проблему их несбалансированности. В ситуациях, когда количество объектов одного или нескольких классов значительно превосходит количество объектов других классов, производительность и точность большинства стандартных алгоритмов классификации снижается [15]. По умолчанию эти алгоритмы рассчитаны на равномерное распределение данных по классам. Однако, при решении различных задач, модели машинного обучения должны оставаться непредвзятыми и одинаково эффективно определять объекты как мажоритарного, так и миноритарного классов. Кроме проблемы несбалансированного соотношения данных различных классов существует также проблема внутриклассовой несбалансированности, проявляющаяся в неравномерном распределении данных внутри каждого отдельного класса. Таким образом, важность учета расположения объектов в признаковом пространстве и равнозначности их оценки становится особенно значимой.

Обозначенные проблемы в известной степени можно решать на основе синтезирования набора данных различными методами. Например, статистическими методами в какой-то мере можно пытаться решить проблемы № 3, № 4 и № 5, а проблемы № 1 и № 2 стандартными методами можно решить до степени распространенности соответствующих ошибок в обучающем наборе данных.

Весьма перспективным выглядит подход синтеза на основе создания компьютерных моделей [16]. Однако такой подход не нацелен напрямую на решение проблемы несбалансированности данных. Компьютерные модели работают с теми данными, которые им предоставляются, и если эти данные изначально несбалансированы, то набор данных, полученный в результате моделирования, также может сохранить этот недостаток.

Другое решение проблем, связанных с малыми данными и несбалансированностью выборок, заключается в выполнении стандартной количественной балансировки. При таком подходе на основе статистического анализа создаются дополнительные объекты, которые имитируют свойства исходного набора данных. Например, при использовании алгоритма SMOTE, для каждого объекта x_i из класса меньшинства новый объект синтезируется вдоль отрезков, соединяющих его с соседями x_{ni} из этого же класса [17]:

$$x_{new} = x_i + \lambda(x_i - x_{ni}),$$

здесь λ – некоторое случайное число, $\lambda \in [0; 1]$.

В случае использования алгоритма ADASYN синтезация производится преимущественно для тех объектов миноритарного класса, которые располагаются в области с повышенной концентрацией объектов мажоритарного класса. Это позволяет увеличить плотность объектов миноритарного класса именно в областях наибольшей неопределенности.

Таким образом, синтез данных является мощным подходом, способствующим улучшению качества выборок и решению некоторых проблем, связанных с их формированием.

Тем не менее, синтезирование набора данных порождает свои проблемы, требующие дополнительного внимания. Целью синтеза должно быть не просто увеличение объема данных за счет создания подобных данных (что наблюдается при количественной балансировке, например, при использовании SMOTE и ADASYN), а именно улучшение их репрезентативности. Сравнение различных методов синтеза приведено в таблице 1.

Таким образом, такие методы синтеза как SMOTE и ADASYN демонстрируют ограниченные возможности, создавая объекты похожие на уже имеющиеся в данных, основываясь на подходе поиска ближайшего соседа, что может также привести к увеличению количества объектов-выбросов. Особенностью Gaussian Copula Synthesizer является учет зависимостей между объектами на основе всего набора данных, однако отсутствие приоритета синтезирования в определенной области может привести к сохранению или даже усилению шума. Предлагаемый в данной работе метод пространственная балансировка, напротив, целенаправленно синтезирует объекты в значимых областях, что способствует устранению шума и выявлению новых закономерностей в данных и существенно повышает качество и репрезентативность данных.

Таблица 1. Сравнение методов синтеза

SMOTE	ADASYN	Gaussian Copula Synthesizer	Пространственная балансировка
Создание синтетических объектов путём линейной интерполяции между соседними объектами миноритарного класса (синтез на локальном уровне ближайших соседей).	Адаптивная синтезация объектов в области трудноклассифицируемых объектов миноритарного класса (адаптивный синтез на локальном уровне ближайших соседей).	Преобразование исходных данных в заданное распределение. Синтез с учетом зависимостей между характеристиками объектов на основе всего набора данных.	Синтез данных в определенных значимых с точки зрения классификации областях путем задания нужного распределения
Объекты миноритарного класса, на основе которых осуществляется синтез, выбираются случайным равномерным образом, поэтому исходное распределение данных практически не меняется.	Увеличение числа объектов на границе классов и в проблемных зонах (незначительное повышение репрезентативности).	Синтез по всей области значений, характеризующих состояние объектов.	Синтез в заданной области значений, способствующий выявлению новых закономерностей с учетом предметной области (предметно обоснованный выбор области синтеза, способствующий повышению репрезентативности данных).
Возможно усиление шума при создании синтетических объектов возле объектов-выбросов			Снижение влияния шума за счет учета логики размещения объектов согласно предметной области.

Репрезентативность – это соответствие распределения признаков объектов выборки аналогичному истинному распределению в генеральной совокупности. Репрезентативность характеризуется такими понятиями, как полнота и разнообразие данных [18, 19] и свидетельствует о способности выборки достоверно отражать структуру и особенности исходной совокупности. При обнаружении недостаточно представленных областей исходного распределения выборка дополняется синтетическими данными, что позволяет существенно повысить её репрезентативность по сравнению с чистым набором натуральных эмпирических данных. При этом оценка репрезентативности является нетривиальной задачей ввиду

того, что зачастую в задачах машинного обучения нам неизвестно истинное распределение данных. В работе оценка репрезентативности производится опосредованно через модели машинного обучения. В условиях, когда меняются только характеристики обучающего набора данных, а не модели, логично предполагать, что на изменения точности классификации в первую очередь повлияли изменения в данных и их качественных характеристиках. Также осуществляется оценка репрезентативности через энтропию распределения объектов в признаковом пространстве.

2. Термометрические данные как класс медицинских данных.

Как уже отмечалось ранее, одной из распространенных сфер деятельности, где ведется работа преимущественно с малыми наборами данных, является медицина. При этом особое значение в медицине имеют термометрические данные, получаемые после проведения соответствующих обследований, например микроволновой радиотермометрии.

Микроволновая радиотермометрия (РТМ метод) – это метод диагностики различных заболеваний, основанный на измерении температурного поля $T(x, y, z)$ обследуемого органа, где x, y, z – координаты точек органа. Обследование РТМ методом является неинвазивной и совершенно безопасной для пациента процедурой. Обследование проводится в инфракрасном и микроволновом диапазоне. Инфракрасный диапазон измеряет температуру подкожного слоя, а микроволновый диапазон позволяет выявлять тепловые аномалии внутри организма на глубине нескольких сантиметров. Далее на основе полученных измерений выполняется построение, визуализация и анализ температурного поля. При визуализации данных, характерными особенностями могут быть горячие области, различие температурных полей парных органов. К примеру, при раке молочной железы, в скрининговых обследованиях которого данный метод применяется наиболее активно, у пациента наблюдается повышенная температура в области нахождения опухоли [22]. Диагностику молочных желез методом микроволновой радиотермометрии впервые начали проводить в 70-ых годах [23–26]. В итоге было установлено, что благодаря данному методу рак молочной железы можно косвенно выявлять на ранней стадии.

Однако РТМ метод находит применение не только в диагностике молочных желез, но и в выявлении других заболеваний, таких как патологии головного мозга, заболевания сердечно-сосудистой системы, вен нижних конечностей, болезни костно-мышечной системы и др. [27, 28]. Для каждого типа заболевания существует своя схема обследования, согласно которой измерение температур производится в определенных

точках. Таким образом, термометрические данные представляют собой значения температур, измеренных в соответствующих точках. Структура данных в таком случае может быть задана таблицей. Колонки таблицы соответствуют номерам точек, в которых производились измерения. Строки - значениям измерений обследуемого органа. При этом схема обследования ограничена заданным количеством точек измерения, в таком случае функционалы, описываемые далее, заменяются функциями. Помимо этого термометрические данные также можно представить в виде изображения температурных полей, которые строятся путем цветового кодирования схематично обозначенной области измерения температур от среднего значения температуры (зеленый цвет) к максимальной (красный цвет) и минимальной (синий цвет).

Термометрические данные, получаемые при использовании метода РТМ, можно отнести к отдельному классу медицинских данных. Проведение анализа таких данных и постановка на их основе диагноза во многом субъективны и значимо зависят от опыта и искусства врача. Поэтому на практике анализ таких данных доступен только высококвалифицированным специалистам достаточно узких специальностей. Данное обстоятельство нивелирует потенциальную возможность широкого использования термометрического оборудования. Решение данной проблемы видится в использовании методов математического и компьютерного моделирования, возможностей современных информационных технологий и, прежде всего, методов машинного обучения. Применение машинного обучения позволит сделать РТМ-диагностику доступной для максимально широкого круга специалистов и пациентов. Интеллектуальный анализ данных упростит процесс анализа результатов обследования, что позволит расширить доступ к качественной диагностике при массовых скрининговых обследованиях и снизить нагрузку на опытных высококвалифицированных врачей, направив их усилия на решение наиболее сложных задач.

3. Предшествующие результаты. Важный вклад в развитие методов интеллектуального анализа термометрических данных внесли работы [16, 29–36], в которых были созданы математические модели описания диагностического состояния пациента, представлена оригинальная концепция построения признаковых пространств на основе дескриптивных моделей тепловых полей биологических тканей и реализованы эффективные алгоритмы классификации.

Первым шагом к созданию математических моделей для описания диагностического состояния пациента послужило использование не самих температурных данных, а функционалов, т.е. отображений из множества

температурных полей на числовую прямую. Отметим, что на практике производится аппроксимация температурных полей путем измерения их значений в определенных точках. Выбор точек осуществляется врачами на основе знаний предметной области. В итоге при аппроксимации поля конечным множеством температур функционалы заменяются функциями. Так, поля глубинных температур i -го пациента представимы в виде векторов:

$$T_c^{i,mw} = (t_{c,0}^{i,mw}, t_{c,1}^{i,mw}, \dots, t_{c,9}^{i,mw});$$

$$T_p^{i,mw} = (t_{p,0}^{i,mw}, t_{p,1}^{i,mw}, \dots, t_{p,9}^{i,mw}),$$

где первый нижний индекс обозначает температуры текущей ("с") или парной молочной железы ("р"), второй нижний индекс – номер точки согласно заранее заданной схеме обследования. Тогда значения функции 1 считается как разность среднего значения множества глубинных температур текущей молочной железы и парной. Для температур подкожного слоя вид векторов аналогичен:

$$T_c^{i,ir} = (t_{c,0}^{i,ir}, t_{c,1}^{i,ir}, \dots, t_{c,9}^{i,ir});$$

$$T_p^{i,ir} = (t_{p,0}^{i,ir}, t_{p,1}^{i,ir}, \dots, t_{p,9}^{i,ir}),$$

В таблицах 2 – 5 показан пример записи температур для нескольких молочных желез.

Таблица 2. Значения глубинных температур рассматриваемой молочной железы

i	$t_{c,0}^{i,mw}$	$t_{c,1}^{i,mw}$	$t_{c,2}^{i,mw}$	$t_{c,3}^{i,mw}$	$t_{c,4}^{i,mw}$	$t_{c,5}^{i,mw}$	$t_{c,6}^{i,mw}$	$t_{c,7}^{i,mw}$	$t_{c,8}^{i,mw}$	$t_{c,9}^{i,mw}$
0	35.6	35.4	35.9	35.8	35.6	35.6	35.3	35.3	35.6	35.6
1	33.6	33.8	33.2	33.2	33.0	34.0	33.2	33.8	34.3	34.1

Таблица 3. Значения температур подкожного слоя рассматриваемой молочной железы

i	$t_{c,0}^{i,ir}$	$t_{c,1}^{i,ir}$	$t_{c,2}^{i,ir}$	$t_{c,3}^{i,ir}$	$t_{c,4}^{i,ir}$	$t_{c,5}^{i,ir}$	$t_{c,6}^{i,ir}$	$t_{c,7}^{i,ir}$	$t_{c,8}^{i,ir}$	$t_{c,9}^{i,ir}$
0	33.9	33.9	34.6	34.5	33.9	33.3	32.7	33.2	33.2	32.8
1	32.7	32.7	31.7	31.6	30.5	32.1	31.4	32.9	33.3	32.9

Таблица 4. Значения глубинных температур молочной железы парной к рассматриваемой

i	$t_{p,0}^{i,mw}$	$t_{p,1}^{i,mw}$	$t_{p,2}^{i,mw}$	$t_{p,3}^{i,mw}$	$t_{p,4}^{i,mw}$	$t_{p,5}^{i,mw}$	$t_{p,6}^{i,mw}$	$t_{p,7}^{i,mw}$	$t_{p,8}^{i,mw}$	$t_{p,9}^{i,mw}$
0	34.9	35.4	35.6	35.7	35.6	35.2	35.0	35.1	35.6	34.9
1	35.2	34.2	34.4	34.3	34.5	34.3	34.5	34.7	34.9	34.5

Таблица 5. Значения температур подкожного слоя молочной железы парной к рассматриваемой

i	$t_{p,0}^{i,ir}$	$t_{p,1}^{i,ir}$	$t_{p,2}^{i,ir}$	$t_{p,3}^{i,ir}$	$t_{p,4}^{i,ir}$	$t_{p,5}^{i,ir}$	$t_{p,6}^{i,ir}$	$t_{p,7}^{i,ir}$	$t_{p,8}^{i,ir}$	$t_{p,9}^{i,ir}$
0	33.2	33.5	33.8	34.0	33.6	32.9	32.5	32.7	33.5	32.7
1	33.8	32.8	32.7	33.2	32.8	32.4	31.8	32.6	33.3	32.8

При создании математических моделей рассматривались качественные характеристики патологий, основанные на базовых медицинских принципах. Описание данных характеристик возможно с помощью математических функционалов, приведенных в работах [20, 32, 33]. В частности, известно, что здоровые железы обладают «зеркальной» симметрией температур в одноименных точках. Данное свойство может быть описано, например, следующим функционалом

$$sym = \overline{T_c^{mw}} - \overline{T_p^{mw}}, \quad (1)$$

Здесь T_c^{mw} – поле глубинных температур молочной железы, T_p^{mw} – поле глубинных температур парной молочной железы, $\overline{T}$ – среднее значение температур.

Для здоровых молочных желез характерны невысокие меры разброса температур. Данное свойство может быть описано, например, следующим функционалом

$$osc = \max_{t \in T_c^{mw}} t - \min_{t \in T_c^{mw}} t, \quad (2)$$

Кроме того отметим, что температура внутри органа почти всегда выше, чем на поверхности. Разность внутренней и поверхностной температур ($t_{c,b}^{i,g} = t_{c,b}^{i,mw} - t_{c,b}^{i,ir}$, b – номер точки измерения) принято называть внутренним градиентом, величина которого является важной

характеристикой состояния пациентов. Пример записи о внутреннем градиенте молочных желез приведён в таблице 6.

Таблица 6. Значения внутреннего градиента рассматриваемой молочной железы

i	$t_{c,0}^{i,g}$	$t_{c,1}^{i,g}$	$t_{c,2}^{i,g}$	$t_{c,3}^{i,g}$	$t_{c,4}^{i,g}$	$t_{c,5}^{i,g}$	$t_{c,6}^{i,g}$	$t_{c,7}^{i,g}$	$t_{c,8}^{i,g}$	$t_{c,9}^{i,g}$
0	1.7	1.5	1.3	1.3	1.7	2.3	2.6	2.1	2.4	2.8
1	0.9	1.1	1.5	1.6	2.5	1.9	1.8	0.9	1.0	1.2

Обозначенную характеристику разности температур внутри и на поверхности органа можно описать, например, функционалом

$$\max_g = \max_{t \in T_c^{mw} - T_c^{ir}} |t|, \quad (3)$$

где T_c^{ir} – температуры подкожного слоя молочной железы.

Таким образом, на основе приведённых формул (1) – (3) для каждой из рассматриваемых молочных желез были найдены соответствующие значения функционалов (таблица 7).

Таблица 7. Найденные значения функционалов

i	$\overline{T_c^{mw}}$	$\overline{T_p^{mw}}$	$\max_{t \in T_c^{mw}} t$	$\min_{t \in T_c^{mw}} t$	sym	osc	$\max_g$
0	35.6	35.3	35.9	35.3	0.3	0.6	2.8
1	33.6	34.6	34.3	33.0	-1	1.3	2.5

Итоговое признаковое пространство для решения задачи классификации составляется из наиболее информативных признаков и для различных органов и различных задач классификации может меняться.

Отметим, что информативные области найденные в работе [20] не были напрямую задействованы в алгоритме классификации, а использовались как инструмент для обоснования решения. Благодаря этим данным эксперт может установить, по каким признакам и насколько сильно термометрические показатели пациентки отклоняются от нормы.

Созданный в работе [20] алгоритм выполнял классификацию двух классов: здоровых молочных желез и больных раком. Полученное значение эффективности классификации $eff = 0.85$, чувствительности $sens = 0.89$, специфичности $spec = 0.81$. Здесь

$$sens = \frac{TP}{TP + FN},$$

$$spec = \frac{TN}{TN + FP},$$

$$eff = \sqrt{sens \cdot spec},$$

а TP – количество истинно-положительных результатов, или объектов класса «Больные», классифицированных как «Больные»; TN – количество истинно-отрицательных результатов, или объектов класса «Здоровые», классифицированных как «Здоровые»; FP – количество ложно-положительных результатов, или объектов класса «Больные», не классифицированных как «Больные»; FN – количество ложно-отрицательных результатов, или объектов класса «Здоровые», классифицированных как «Больные».

Учитывая, что микроволновая радиотермометрия является, так называемым, скрининговым методом, решалась и другая задача классификации молочных желез по термометрическим данным. А именно, производилось не выявление рака, а принадлежность пациентки к группе риска. Группа риска выделяется на основе тепловых аномалий по специальной термометрической шкале. Данная шкала содержит шесть групп, которые характеризуют уровень выраженности тепловых процессов и присваиваются молочным железам, исходя из РТМ-заключения. Далее эти группы будем именовать th-классами, которые принимают следующие значения: th0, th1, th2, th3, th4, th5. Согласно термометрической шкале чем выше уровень аномалий (значение th-класса), тем выше риск наличия рака молочной железы. Основные характеристики каждого из значений th-классов приведены в таблице 8.

Руководствуясь рекомендациями онкологов-маммологов, th-классы были разделены на 2 группы. Пациентки, температурные поля молочных желез которых определяются как th0-, th1- или th2-класс, относятся к группе пациенток, у которых отсутствуют онкологические заболевания. Пациентки, температурные поля молочных желез которых определяются как th3-, th4- или th5-класс, относятся к группе риска развития злокачественного новообразования.

Определение группы риска рака молочной железы по термометрическим признакам и интерпретация полученных результатов на языке понятном врачам выполнялась в работе [37]. Эффективность классификации достигала 0.92, специфичность – 0.90 и чувствительность – 0.93. Но в ходе проведения детального анализа результатов, несмотря на

высокую точность отделения этих двух групп данных, было замечено, что точность определения некоторых из th-классов оказалась критически низкой.

Таблица 8. Характеристики уровней выраженности температурных аномалий

th	Характеристика
0	Тепловых изменений практически нет.
1	Характерных признаков рака молочной железы на текущий момент нет. Но присутствует область пониженной температуры.
2	Характерных признаков рака молочной железы на текущий момент нет. Но присутствует область повышенной температуры.
3	Высокие глубинные температуры и температуры подкожного слоя, значение разности между ними не соответствует норме.
4	Температурные отклонения соответствуют одному из нижеперечисленных признаков: – заметно отличающаяся термоасимметрия (разность температурных полей молочных желез) при этом (глубинные температуры органа) t_{mw} и t_{ir} (температуры подкожного слоя) на одном уровне; – $t_{ir} \geq t_{mw}$; – сочетание повышенного уровня температур в одной из молочных желез с сохранением соотношения глубинных температур и температур подкожного слоя на одном уровне.
5	Температурные отклонения соответствуют одному или нескольким нижеперечисленным признакам: – заметно отличающаяся термоасимметрия при этом t_{mw} и t_{ir} на одном уровне; – на поле глубинных температур наблюдается «колокол»; – повышенный уровень температур в одной из желез, при этом $t_{ir} \geq t_{mw}$;

Также стоит принять во внимание тот факт, что th-классы, которые определяют молочные железы без онкологии вовсе не означают, что железы абсолютно здоровы. Так в молочных железах, определенных как th1, обнаруживаются следующие доброкачественные процессы: локализованный фиброз, липома, грубый послеоперационный рубец или фиброаденома без признаков роста. В случае th2 врачи диагностируют мастопатию, воспаленные кисты, участки ожога или воспаления на коже, а также фиброаденомы с признаками роста. Следовательно, обнаружение температурных отклонений, присущих всем этим 6 классам, будет способствовать своевременной диагностике различных заболеваний.

Исходя из написанного выше становится понятно, что возникает новая задача по определению, описанных th-классов, т.е. проведению многоклассовой классификации для выявления различных температурных отклонений в молочных железах. Решение данной задачи также рассматривается в настоящей работе. Признаки, по которым проводилась

классификация, брались из работы [20]. Обоснование выбора таких признаков заключается в схожести предметных областей данных исследований.

4. Постановка задачи. Предположим, что у нас имеется множество объектов обучающей выборки, разделенное на k классов. Будем считать, что объем выборки мал, а количество объектов некоторых классов значимо превосходит количество объектов в других классах. В связи с малым объемом данных, нам нежелательно решать проблему несбалансированности путем удаления объектов обучающей выборки. Следовательно, вместо удаления будем добавлять объекты. Далее считаем, что добавление объектов обучающей выборки возможно только путем их синтеза.

Пусть состояние каждого объекта выборки описывается некоторой функцией (или вектор-функцией). Далее будем считать, что в качестве элементов признакового пространства используется набор функционалов $(F_1, \dots, F_n)$, определённых на множестве функций, описывающих состояние объектов [20]. Подобный подход к построению признакового пространства в последние годы стал активно применяться и уже доказал свою эффективность.

Рассмотрим функционал F_j , $j \in \{1, \dots, n\}$. Эффективность разделения объектов на классы с помощью данного функционала определяется информативностью W . Традиционно под информативностью понимают величину, количественно характеризующую степень уменьшения неопределенности при классификации объектов с помощью данного функционала, т.е. меру того, насколько хорошо он помогает различать объекты разных классов. Для оценки информативности существует множество критериев. Так в работе [20] использовался критерий статистической информативности, но также рассматривались энтропийный, эвристический критерий и др. В данной работе информативность функционала оценивалась для каждого класса $y \in \{0, \dots, k\}$ отдельно и основывалась на том, насколько хорошо по значениям F_j можно отделить объекты класса y от объектов остальных классов (метод "один против всех"). Область значений F_j , которая является типичной для класса объектов y , называется информативной областью и обозначается I_{jy} . Информативная область – это область, на которой обеспечивается достижение величины:

$$W_{jy} = \sup_{I_{j\alpha}} W(F_j, I_{jy}),$$

где $I_{j\alpha}$ – всевозможные подобласти значений F_j , на которых указанная информативность имеет смысл.

Информативность $W(F_j, I_{jy})$ является функцией от рассматриваемой информативной области I_{jy} по функционалу F_j . В настоящем исследовании информативность определяется следующим образом:

$$W(F_j, I_{jy}) = \sqrt{(1 - l_{j0}) \cdot l_{jy}}, \quad (4)$$

где:

$$l_{jy} = \frac{|\{x \in F_{jy} \mid i_{jy,start} < x < i_{jy,end}\}|}{|F_{jy}|}$$

– доля объектов класса y (множество F_{jy} содержит значения функционала F_j объектов, соответствующих этому классу), попавших в интервал значений $[i_{jy,start}; i_{jy,end}]$, характерный для данного класса; здесь $|\cdot|$ – мощность множества;

$$l_{j0} = \frac{|\{x \in F_{j0} \mid i_{jy,start} < x < i_{jy,end}\}|}{|F_{j0}|}$$

– доля объектов других классов (множество F_{j0} включает значения функционала F_j объектов, не относящихся к классу y), также попавших в указанный интервал $[i_{jy,start}; i_{jy,end}]$.

Выбор формулы определен устоявшейся формой определения качества разделения объектов разных классов в данной предметной области.

На основе введенных обозначений определим геометрическое строение подмножеств признакового пространства на основе определения признака из работы [20]:

$$\phi_j = (F_j, (I_{j0}, \dots, I_{jk}), (W_{j0}, \dots, W_{jk})).$$

Здесь F_j – j -ый функционал; I_j – вектор информативных областей функционала F_j по классам объектов; W_j – вектор информативностей j -ого признака по классам объектов.

Учитывая геометрическое строение признакового пространства и несбалансированность классов в обучающей выборке, сформулируем задачу исследования.

Задача: повысить репрезентативность обучающего набора данных путем синтеза дополняющего набора, который при объединении с исходным набором данных, меняет расположение элементов обучающей выборки в признаковом пространстве в соответствии с некоторым распределением.

5. Метод решения задачи. Пусть в имеющейся обучающей выборке содержится m объектов, определенных с помощью матрицы функционалов F и вектор-столбца Y , содержащего классы, к которым принадлежат соответствующие объекты:

$$F = \begin{bmatrix} F_{11} & \dots & F_{1n} \\ \vdots & \ddots & \vdots \\ F_{m1} & \dots & F_{mn} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

Здесь $(F_{i1}, \dots, F_{in})$ – описание объекта i из обучающей выборки; y_i – класс объекта i из обучающей выборки. Матрица F строится на основе измеренных температур и выбранных функционалов (например, 1, 3, 2). Пример расчета ее значений приведен в таблице 7. Элементы вектор-столбца Y как правило заранее определены врачом-диагностом и являются числом от 0 до 5, которое соответствует температурным аномалиям из таблицы 8.

Каждому классу объектов y по каждому функционалу F_j соответствует свое значение информативности W_{jy} и своя информативная область I_{jy} (в данной работе в качестве информативной области рассматривается числовой промежуток - информативный интервал). Это можно отразить в виде соответствующих матриц W^* и I размерностью $n \times k$. Строки матриц соответствуют функционалам, а столбцы – классам объектов:

$$W^* = \begin{bmatrix} W_{11} & \dots & W_{1k} \\ \vdots & \ddots & \vdots \\ W_{n1} & \dots & W_{nk} \end{bmatrix},$$

$$I = \begin{bmatrix} (i_{11,start}; i_{11,end}) & \dots & (i_{1k,start}; i_{1k,end}) \\ \vdots & \ddots & \vdots \\ (i_{n1,start}; i_{n1,end}) & \dots & (i_{nk,start}; i_{nk,end}) \end{bmatrix}.$$

Здесь $i_{j,start}$ – левая граница (начало) интервала соответствующего F_j ; $i_{j,end}$ – правая граница (конец) интервала соответствующего F_j .

Построение информативных интервалов I_{jy} признакового пространства выполняется в соответствии с предметной областью исследования, согласно которой у каждого термометрического класса должна существовать единственная связная информативная область. Опишем метод построения информативных областей (интервалов) подробнее:

1. Определяется стартовая точка, от нее будет проводиться расширение интервала. В качестве стартовой точки для большинства классов бралась их медиана. Для $th0$ -класса, которому характерно отсутствие температурных аномалий, в качестве стартовой точки по некоторым признакам выбиралось нулевое значение. Так, например, исходя из медицинского положения о симметричности температур в одноименных точках парных молочных желез значение стартовой точки интервала $th0$ -класса было равно 0 для признака, характеризующего отличие температурных полей правой и левой молочных желез.

2. Определяется способ расширения интервала. Исходя из особенностей выбора стартовой точки, расширение интервала происходило либо в обе стороны от нее, либо только в одну. Расширение в одну сторону применялось, когда известно, что значения левее точки, от которой начинается расширение, признак принимать не может.

3. Осуществляется расширение интервала от стартовой точки на шаг δ до тех пор, пока не уменьшается информативность интервала W_{ij} , а границы не выходят за пределы области распределения класса. Информативность интервалов определяется формулой 4, ей же задается матрица W^* .

Выбор шага δ влияет на скорость решения и точность границ интервала. При низких значениях шага необходимо большее количество итераций для сходимости метода, но будет найдено наибольшее значение величины 4. Так как на малых наборах данных метод не является ресурсоемким, рекомендуемым значением δ является минимальное значение расстояния между элементами множества F_{jy} , по которому строится интервал:

$$\min_{\substack{x, y \in F_{jy} \\ x \neq y}} |x - y|.$$

Выбор меньшего значения шага не имеет смысла, так как при расширении интервала может не изменяться соотношение количество элементов в нем, т.е. информативность будет оставаться прежней. В работе шаг δ определялся данным способом.

Следует отметить, что интервалы различных классов по ряду функционалов могут пересекаться, что указывает на их общие характеристики по ним.

Можно считать, что выбор начального варианта обучающей выборки для каждого класса y по каждому функционалу F_j задал нам некоторый закон распределения данных d_{jy} внутри информативного интервала I_{jy} . Под законом распределения понимается некоторое соотношение, которое устанавливает связь между возможными значениями $\xi = F_j$ и соответствующей им вероятностью. Задать такой закон можно, например, через функцию распределения или плотность распределения. Функция распределения выражает вероятность того, что ξ примет значение, меньше некоторого x : $f_\xi(x) = P(\xi < x)$. Плотность распределения в свою очередь находится как f'_ξ [21].

Таким образом исходная матрица распределений D имеет следующий вид:

$$D = \begin{bmatrix} d_{11} & \dots & d_{1k} \\ \vdots & \ddots & \vdots \\ d_{n1} & \dots & d_{nk} \end{bmatrix}.$$

Распределение объектов выборки в информативных интервалах может не соответствовать распределению, имеющемуся в реальности. Следовательно, возникает задача восстановления репрезентативности обучающей выборки данных с помощью дополнения имеющегося набора данных синтезированным набором и приведения закона распределения D к некоторому D^* :

$$D^* = \begin{bmatrix} d_{11}^* & \dots & d_{1k}^* \\ \vdots & \ddots & \vdots \\ d_{n1}^* & \dots & d_{nk}^* \end{bmatrix},$$

где d_{jy}^* определяется реальным распределением изучаемых объектов.

Закон распределения D^* может быть заранее известен, исходя из предметной области. В противном случае рассматривается конечное

множество различных законов распределений, из которых выбирается наилучшее по каждому функционалу. Наилучшим считается то распределение, по которому было получено максимальное значение меры эффективности на валидационном наборе данных. Множество законов распределений можно задавать как исходя из гипотез, сформированных на основе предметной области, так и выбирая из стандартных законов, например, нормальное, равномерное, биномиальное и т.п.

Для решения поставленной задачи предлагается метод управляемой синтезаций данных. В результате получим новый набор данных, который, помимо натуральных данных, также будет содержать и синтетические данные: $(F^*, Y^*) = (F \cup F_{syn}, Y \cup Y_{syn})$. Далее производится обучение на новом наборе данных алгоритма классификации и проверка его эффективности на валидационном наборе. Увеличение эффективности классификации при использовании нового обучающего набора позволит подтвердить правильность выбора распределения и, соответственно, правильность пространственной балансировки.

Приступим к синтезу на основе метода пространственной балансировки. Для каждого класса объектов y , интервала i_{jy} и соответствующего функционала F_j проведем следующие действия.

1. Разделим интервал i_{jy} на p подинтервалов. Значение p задается изначально. Его оптимальное значение определяется экспериментальным путем. Так, например, в случае нормального распределения, p может быть задано таким образом, чтобы ширина каждого подинтервала была не меньше среднеквадратического отклонения σ . В некоторых случаях, например в гауссовой смеси распределений с несколькими модами, имеет смысл увеличить количество подинтервалов, в работе интервал разбивался на 5 подинтервалов.

2. Определим долю объектов обучающего набора данных в каждом подинтервале i_{jy}^e , $e = \{1, \dots, p\}$ на основе закона распределения d_{jy} .

В непрерывном случае доля объектов определяется формулой:

$$p_{jy}^e = P(i_{jy,start}^e < x < i_{jy,end}^e) = \int_{i_{jy,start}^e}^{i_{jy,end}^e} d(x) dx.$$

Здесь $i_{jy,start}^e$ и $i_{jy,end}^e$ – левая и правая границы подинтервала соответственно, $d(x)$ – функция плотности распределения.

В дискретном случае для известного набора данных доля объектов определяется по формуле:

$$p_{jy}^e = P(i_{jy,start}^e < x < i_{jy,end}^e) = \frac{|\{x \in F_{jy} \mid x \in i_{jy}^e\}|}{|\{x \in F_{jy} \mid x \in i_{jy}\}|}.$$

Таким образом имеется два вектора, соответствующих долям объектов в подинтервалах для распределений d_{jy} и d_{jy}^* соответственно:

$$P_{jy} = (p_{jy}^1, p_{jy}^2, \dots, p_{jy}^p), \quad P_{jy}^* = (p_{jy}^{1*}, p_{jy}^{2*}, \dots, p_{jy}^{p*}).$$

3. Определим количество имеющихся в каждом подинтервале объектов:

$$a_{jy}^e = |\{x \in F_{jy} \mid x \in i_{jy}^e\}|.$$

4. Определим количество объектов, которое должно быть в интервале i_{jy} . Поскольку обе плотности d_{ij} и d_{jy}^* на интервале i_{jy} непрерывны, неотрицательны и нормированы, найдется хотя бы один подинтервал, в котором $p_{jy}^e \geq p_{jy}^{e*}$.

Чтобы уменьшить долю объектов в таких подинтервалах, нужно либо удалить некоторое количество объектов в них, либо синтезировать новые объекты в других подинтервалах. Из-за малого объема данных нельзя удалять исходные объекты, поэтому имеет смысл сохранить в данном подинтервале имеющееся количество объектов и определить необходимое количество объектов на всем интервале i_{jy} следующим образом:

$$A^* = \left[\frac{a_{jy}^e}{p_{jy}^{e*}} \right].$$

В качестве значения e выбирается наименьшая разница среди элементов P_{jy} и P_{jy}^* :

$$e_{min} = \arg \min_{i \in \{0,1,\dots,p\}} (p_{jy}^i - p_{jy}^{i*}).$$

На основе этого рассчитаем общее количество объектов:

$$A^* = \left[\frac{a_{jy}^{e_{min}}}{p_{jy}^{e_{min}^*}} \right],$$

где $[\cdot]$ – целая часть числа.

5. Определим количество объектов t_e , которых не хватает в каждом из подинтервалов для получения итогового распределения d_{jy}^* по формуле:

$$t_e = [A^* \cdot p_{jy}^{e^*}] - a_{jy}^e.$$

6. Выполним синтез t_e объектов, таким образом, чтобы значения $F_{j, syn}$ попадали в подинтервал i_{jy}^e .

7. Добавим синтезированные объекты в F_{syn} и укажем для них соответствующее значение класса y . После завершения цикла синтеза, объединим изначальную и синтезированную выборки.

Пространственная балансировка позволяет получить более репрезентативный (с точки зрения заданного распределения D^*) набор данных. Тем самым данные начинают лучше отражать ключевые характеристики предметной области, минимизируется влияние пространственных аномалий, повышается точность прогнозов и устойчивость модели относительно распределения данных. Все это особенно актуально в условиях малых выборок.

Псевдокод методов приведен в Приложении.

Пример 1. На рассматриваемом интервале содержится 1000 объектов. При его разбиении на 5 равных подинтервалов получается следующий вектор долей p и количество объектов a :

$$p = (0.1, 0.2, 0.4, 0.2, 0.1), \quad a = (100, 200, 400, 200, 100).$$

Нужно синтезировать объекты, приводящие распределение на данном интервале к нормальному, т.е. к следующим долям p^* :

$$p^* = (0.027, 0.184, 0.578, 0.184, 0.027).$$

Подинтервалом e_{min} с наименьшей разницей долей будет являться первый (или последний, величины разностей долей в обоих случаях совпадают):

$$p^0 - p^{0*} = -0.073.$$

Тогда:

$$A^* = \left\lceil \frac{100}{0.027} \right\rceil = 3073,$$

и количество объектов, необходимых для синтеза в каждом подинтервале:

$$t = (0, 481, 1741, 481, 0).$$

Пример 2. С целью выявления взаимосвязи между эффективностью классификации и репрезентативность набора данных рассмотрим следующий пример. Зададим самостоятельно генеральную совокупность объектов $F_g \times Y_g$ объемом 10000 записей для каждого класса ($m = 30000$).

$$F_g = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ \vdots & \ddots & \vdots \\ F_{m1} & F_{m2} & F_{m3} \end{bmatrix}, Y_g = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, y = \{0; 1; 2\}.$$

При этом пусть распределение объектов генеральной совокупности в информативных интервалах близким к нормальному (статистика Андерсона-Дарлинга принимает значения меньше критического для уровня значимости в 5%). Пример распределения для функционала F_3 и объектов из генеральной совокупности с классом 1 приведен на рисунке 1. В данном случае границы информативного интервала имели значения 9 и 10 соответственно.

Предположим, что на основе генеральной совокупности составлена некоторая выборка объектов. Причем составлена таким образом, что нормальное распределение значений функционалов в информативных интервалах было нарушено (рисунок 2).

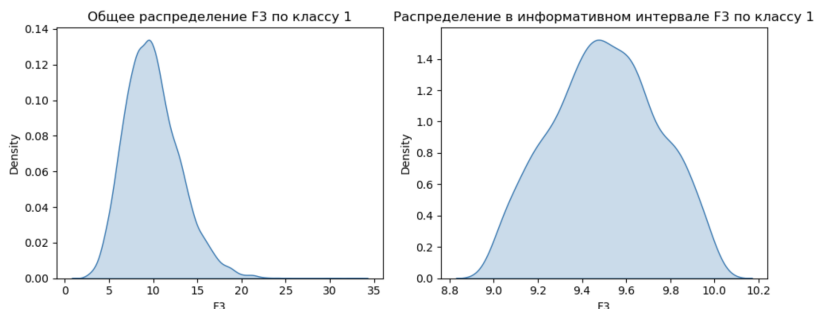


Рис. 1. Плотность распределения значений функционала $F3$ по объектам класса 1 из генеральной совокупности



Рис. 2. Плотность распределения значений функционала $F3$ по объектам класса 1 из выборочной совокупности

Следовательно, данная выборка будет обладать пониженной репрезентативностью. Кроме того объем данных в выборке по сравнению с генеральной совокупностью был существенно уменьшен и составил 1374 записи для класса 0, 888 - для класса 1 и 536 - для класса 2; также в выборке наблюдается классовая несбалансированность.

Применим метод пространственной балансировки с целью синтеза данных и восстановления распределения в информативных интервалах.

Пример распределения значений для функционала $F3$ в информативном интервале после синтеза объектов класса 1 приведен на рисунке 3.

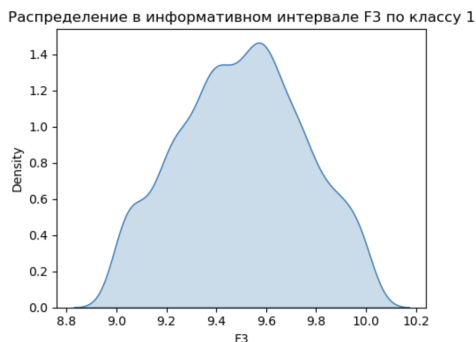


Рис. 3. Плотность распределения значений для функционала $F3$ в информативном интервале после синтеза объектов класса 1

Оценим значение дивергенции Кульбака-Лейблера (D_{KL_F}) по распределениям p генеральной совокупности и q – выборочной совокупности для каждого функционала F в пределах информативных интервалов до и после проведения синтеза:

$$D_{kl}(p||q) = \int p(x) \ln \left(\frac{p(x)}{q(x)} \right) dx .$$

Также до и после синтеза вычислим эффективность классификации (применяемый алгоритм классификации к-ближайших соседей), исходя из точности определения каждого из классов выборочной совокупности (таблица 9):

$$m_{eff} = \sqrt[3]{\prod_{i=0}^2 y(i)},$$

где $y(i)$ – точность определения класса i .

Таблица 9. Точность классификации и значения дивергенции до и после синтеза, осуществляемого с помощью пространственной балансировки

Этап	$D_{KL_{F1}}$	$D_{KL_{F2}}$	$D_{KL_{F3}}$	m_{eff}
до синтеза	1.1	0.83	0.45	0.84
после синтеза	0.18	0.11	0	0.94

Уменьшение значения дивергенции после проведения синтеза с помощью пространственной балансировки свидетельствует о повышении репрезентативности выборочной совокупности. Такой эффект наблюдается из-за того, что распределение выборочной совокупности в информативных интервалах после синтеза больше соответствует тому распределению, которое наблюдалось в генеральной совокупности. Кроме того с повышением репрезентативности увеличивается и эффективность классификации. При этом увеличение меры эффективности составляет 10% (с 84% до 94%).

Для большей наглядности проиллюстрируем метод на примере решения задачи многоклассовой классификации пациенток с заболеваниями молочных желез на основе данных микроволновой радиотермометрии.

6. Повышение сбалансированности. При создании модели машинного обучения, обучаемой для решения задачи многоклассовой классификации, а именно, определения уровня выраженности температурных аномалий, было выявлено, что простые модели зачастую склонны игнорировать объекты некоторых th -классов. Точность определения каждого из классов моделью логистической регрессии представлена на рисунке 4.

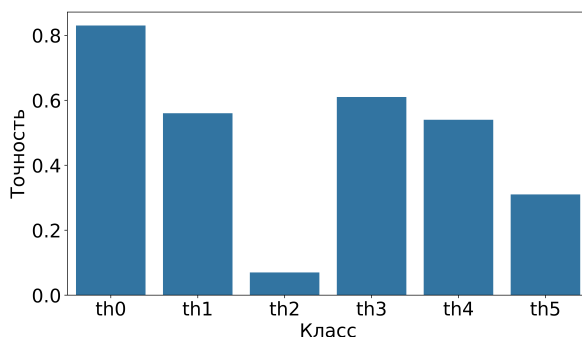


Рис. 4. Точность определения th классов моделью логистической регрессии

Отметим, что наименьшая точность наблюдается по объектам $th2$ -класса – этот класс определяется лишь в 7% случаев, но для специалистов предметной области может представлять не меньший интерес, чем другие классы.

Основная причина неудовлетворительных результатов заключается в несбалансированности начального набора данных – распределение объектов по классам значительно отличается. Кроме того работа ведется

с выборкой данных малого объема (рисунок 5), что также влечет проблемы снижения репрезентативности данных.

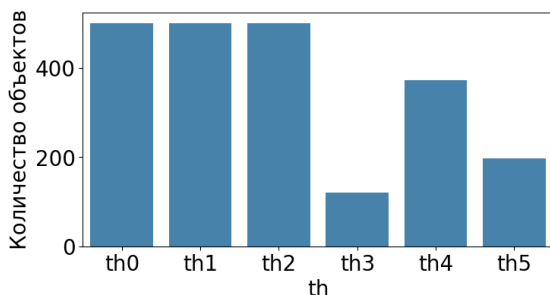


Рис. 5. Начальное распределение данных по th -классам

Заметим, что в оригинальном наборе данных объектов $th0$ -, $th1$ - и $th2$ -классов значительно больше, чем остальных. В общей сумме их количество составляет 8428 объектов. Для упрощения задачи балансировки и в связи с низкой размерностью данных было произведено снижение их количества до 500 в каждом классе с сохранением оригинального распределения.

Для устранения всех вышеперечисленных недостатков предлагается применение метода балансировки, основанном на синтезании объектов, который позволит повысить точность многоклассовой классификации. Выполнять балансировку будем в соответствии с методом пространственной балансировки, описанном ранее в разделе 3. Перед ее осуществлением должны быть заданы:

1. матрица F значений функционалов по объектам (функционалы получены из работ [20, 32, 33]);
2. классы объектов Y , определенные специалистами;
3. матрица I с информативными интервалами, построенные описанным ранее алгоритмом;
4. распределение данных D , которое задано внутри интервалов по классам объектов.

В данном исследовании в качестве объектов рассматриваются температурные поля молочных желез пациенток. В качестве функционалов для описания термометрических процессов используются sym , osc , max_g , описывающие основные характеристические особенности классов. Данные функционалы были определены равенствами (1), (2), (3) выше. Матрица информативности, используемых признаков, представлена в таблице 10.

Таблица 10. Матрица информативности W^*

F	th0	th1	th2	th3	th4	th5
sym	0.765	0.699	0.652	0.598	0.735	0.689
max_g	0.529	0.544	0.531	0.685	0.512	0.593
osc	0.621	0.551	0.551	0.543	0.516	0.512

Метками объектов выступают значения th-классов. Задание информативных интервалов и определение распределения данных в них становится возможным после построения интервалов. Таким образом, в соответствии с постановкой задачи выполнение пространственной балансировки также связано с построением информативных интервалов признакового пространства. Полученные интервалы позволяют выделить, во-первых, области характерных значений для объектов различных классов, а также выявить области, в которых наблюдается нехватка объектов. Последние области будут дополняться синтетическими объектами.

Для синтезаций объектов использовалась модель Gaussian Copula Synthesizer. Она позволяет проводить генерацию новых данных двумя способами: создание всего набора данных с заданным количеством объектов и функцией распределения; создание группы объектов с заданными характеристиками. Благодаря второму способу создания объектов, становится возможной их генерация в найденных информативных интервалах I .

Обучение модели Gaussian Copula Synthesizer предполагает преобразование значений $(F_{i1}, \dots, F_{in})$, описывающих объекты из обучающей выборки, к нормальному распределению с помощью Гауссовой копулы, получая таким образом значения в виде

$$[\Phi^{-1}(R_0(F_0)), \dots, \Phi^{-1}(R_n(F_n))],$$

где R – это функция распределения (в случае использования нормального распределения R определяется через μ и σ); $\Phi^{-1}(R_i(F_i))$ – обратная функция распределения. Затем на основе вычисленных значений рассчитывается матрица ковариаций Q .

При создании синтетического объекта случайным образом генерируется вектор v размерностью n , компоненты которого распределены по нормальному закону. Последующие вычисления происходят с помощью разложения Холецкого, которое позволяет генерировать новые данные с заданной структурной зависимостью,

основанной на рассчитанной ковариационной матрицей: $LL^T = Q$. Таким образом, полученный вектор $u = Lv$ используется для синтеза нового объекта со значениями $[R_i^{-1}(\Phi(u_i)), \dots, R_n^{-1}(\Phi(u_n))]$. Синтезация новых объектов проводится для каждого функционала, при этом объект добавляется в синтетический набор только в том случае, если значения рассматриваемого функционала попадают в информативный интервал.

7. Результаты. По представленным признакам для каждого класса были построены информативные интервалы, максимизирующие элементы матрицы W^* , результаты показаны на рисунке 6.

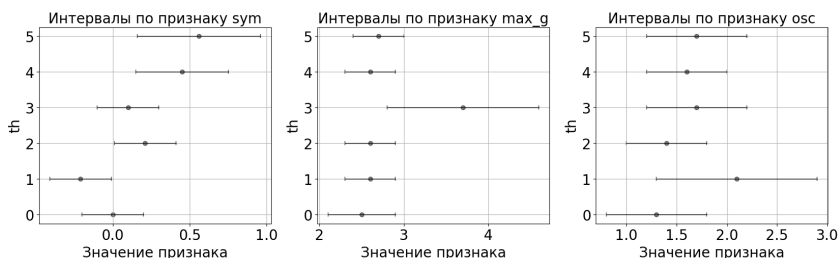


Рис. 6. Найденные информативные интервалы

По первому признаку (*sym*) наблюдается, что *th4*-, *th5*-классы имеют повышенную разность температурных полей молочных желез. Интервал по *th1*-классу заметно смещен влево. Это объясняется тем, что при вычитании из температур обследуемой молочной железы температур по парной молочной железе, найденная разность в большинстве случаев имеет отрицательный знак. Такое поведение температурных полей подтверждается положением о том, что температуры обследуемой молочной железы из *th1*-класса понижены по сравнению с парной железой.

По второму признаку (*max_g*) заметно, что интервал по *th3*-классу значительно смещён в сторону более высоких значений по сравнению с другими *th*-классами, что отражает аномалии в соотношении температур подкожного слоя и глубинных температурах.

По третьему признаку (*osc*) особенно выделяется интервал *th1*-класса. Вычисленные значения осцилляции свидетельствуют о существенном перепаде температур в пределах одной молочной железы, который не так заметен в *th2*- и особенно *th0*-класссах. Таким образом, в объектах *th1*-класса наблюдается температурная неоднородность. Немного менее выраженные аномалии по признаку *osc* соответствуют классам *th3*, *th4* и *th5*. Однако их отличие от классов с наименьшим

риском развития злокачественной опухоли (*th*-классы до *th3*) все же имеется и заметно также по признакам, описанным ранее.

На рисунке 7 показано сравнение распределений по первым двум признакам на изначальной выборке и сбалансированной для объектов *th2*-класса. Пунктирной областью обозначены границы информативных интервалов по данным признакам. Как видно, концентрация объектов в области увеличилась, а распределение объектов стало ближе к нормальному. Аналогичным образом была изменена характеристика распределения *th4*- и *th5*-классов.

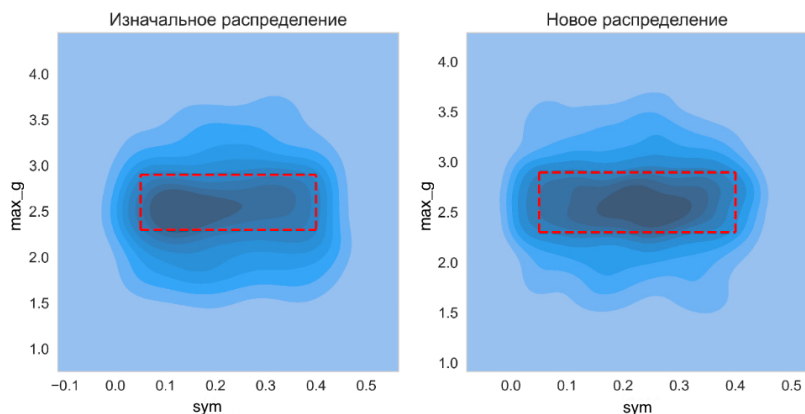


Рис. 7. Распределение объектов класса *th2*

Класс медицинских данных характеризуется высокой сложностью, обусловленной множеством факторов, влияющих на измеряемые характеристики. Одной из ключевых особенностей подобных данных является отсутствие информации о реальном распределении признаков в генеральной совокупности, что делает невозможным применение классических метрических критериев, таких как дивергенция Кульбака–Лейблера, для оценки степени репрезентативности выборки. Кроме того, применение таких дивергенционных мер требует знания априорных распределений, что в условиях ограниченных и неоднородных медицинских данных практически недостижимо.

В связи с этим в данной работе для количественной оценки репрезентативности используется мера информационной энтропии:

$$H = - \sum_{i=1}^N p_i \ln p_i,$$

где p_i – доля объектов в i -м подинтервале признакового пространства, N – число подинтервалов (пять). Энтропия, как мера неопределённости служит косвенной оценкой приближённости распределения к предполагаемому истинному распределению. Высокое значение энтропии может свидетельствовать о разнообразии выборки и охвате различных подгрупп в пределах исследуемой совокупности, тогда как низкая энтропия может указывать на смещённость или ограниченность представленных данных. Таким образом, энтропийный подход предоставляет инструмент для сравнения различных выборок по степени их потенциальной репрезентативности без необходимости априорного знания полной структуры генеральной совокупности.

Оценка энтропии по каждому интервалу и классу приведена в таблице 11. Как видно, количество информации в каждом интервале увеличилось.

Таблица 11. Сравнение энтропии в интервалах при количественной и пространственной балансировках. A^* – количество синтезируемых объектов в интервалах I

th-класс	Признак	Энтропия интервала по количественно сбалансированным данным	Энтропия интервала по пространственно сбалансированным данным	A^*	I
3	sym	1.28	1.29	173	(−0.06; 0.14)
3	max_g	1.14	1.22	357	(2.8; 4.6)
3	osc	1.08	1.15	422	(1.2; 2)
4	sym	1.25	1.29	366	(0.13; 0.73)
4	max_g	1.01	1.12	462	(2.3; 2.9)
4	osc	1.05	1.14	518	(1.2; 2)
5	sym	1.24	1.28	227	(0.16; 0.96)
5	max_g	1.12	1.20	281	(2.4; 3)
5	osc	1.03	1.15	205	(1.25; 2.25)

Для проверки качества полученного набора данных проводилось сравнение моделей классификации, обученных на пяти различных выборках:

- изначальной, не подвергавшейся изменениям;
- количественно сбалансированной с помощью применения метода SMOTE, в результате чего количество объектов каждого класса в выборке стало равным;
- количественно сбалансированной с помощью применения метода ADASYN;

- количественно сбалансированной с помощью применения метода Gaussian Copula;
- пространственно сбалансированной.

После обучения проводилась классификация одного и того же тестового набора данных. В результате определялись точность определения каждого из классов, а также мера эффективности – среднегеометрическое всех точностей:

$$m_{eff} = \sqrt[6]{\prod_{i=0}^5 th_i},$$

где th_i – точность определения th_i -класса. Определим точность, как долю верноопределенных объектов рассматриваемого класса. Для выявления наилучшего набора данных результаты классификации сравнивались по двум параметрам: минимальное значение точности (чем оно выше, тем модель классификации считается более сбалансированной), а также по эффективности, определяющей усреднённое качество классификации.

В ходе экспериментов были рассмотрены различные типы распределений, среди которых наиболее эффективным по всем признакам оказалось нормальное распределение, выбранное для дальнейшей работы в качестве D^* (таблица 12).

Таблица 12. Оценка распределений

Распределение	Равномерное	Нормальное	Гамма	Gaussian KDE
Эффективность	0.27	0.48	0.47	0.475

Также рассматривалась задача объективного задания параметра p , направленного на максимизацию меры эффективности классификации. В качестве эвристики был протестирован подход, основанный на правиле Стёрджеса:

$$p = 1 + \log_2[N],$$

где N – общее количество объектов выборки.

В рамках вычислительного эксперимента использовался набор данных, ограниченный 500 объектами каждого класса, что соответствовало значению $p = 9$ для каждого подкласса. Результаты показали, что оптимальное значение параметра p зависит не только от

состава обучающей выборки, но и от используемой классификационной модели (таблица 13).

Таблица 13. Зависимость меры эффективности классификаторов от параметра p

p	m_{eff}^{df}	m_{eff}^{lr}	A^*
3	0.42	0.382	1994
4	0.441	0.404	2491
5	0.417	0.414	3011
6	0.421	0.426	3022
7	0.35	0.407	3414
8	0.411	0.407	3489
9	0.434	0.409	3857
10	0.426	0.418	2993
11	0.409	0.41	3420
12	0.436	0.406	3341
13	0.411	0.401	3497
14	0.41	0.42	3616

Более того, наивысшие значения меры эффективности m_{eff} были достигнуты при значениях $p < 1 + \log_2 N$. Это наблюдалось как для классификатора на основе решающего леса ($m_{eff}^{df}, p = 4$), так и для логистической регрессии ($m_{eff}^{lr}, p = 6$).

Дополнительно установлено, что увеличение значения параметра p приводит к росту числа синтезируемых объектов A^* , что, в свою очередь, увеличивает время выполнения процедуры синтеза данных.

Следует отметить, что параметр p по сути выполняет роль гиперпараметра, значение которого должно подбираться с учётом характеристик обучающей выборки, архитектуры модели, целей синтеза и допустимого времени выполнения. Это означает, что несмотря на попытки его объективного задания, оптимальный выбор p остаётся задачей, зависящей от множества факторов и требует валидации на конкретных данных.

В таблице 14 показаны результаты вычислительных экспериментов моделью логистической регрессии.

Как видно, количественная балансировка методом ADASYN понизила эффективность классификации, а при балансировке методами SMOTE и Gaussian Copula снизила точность определения $th4$ -класса и повысила точность определения $th2$ -класса, который практически никогда не определялся корректно. В результате эффективность классификации повысилась на 4-5%. Однако в целом качество классификации осталось на низком уровне, так как на двух классах модель оказывается верна

лишь в 17-20% случаев. Пространственная балансировка, в свою очередь, повысила точность определения каждого из этих классов более чем на 14% относительно количественной балансировки. Из-за чего наименьшая точность определения повысилась до 31%, а эффективность – на 7%, что согласуется с увеличением энтропии также на 0,07. Это связано с тем, что увеличение энтропии приводит к повышению информационной насыщенности информативных зон, определяющих характерные значения признаков для каждого класса.

Таблица 14. Результаты вычислительных экспериментов, проведенных логистической регрессией

Набор данных	m_{eff}	th_0	th_1	th_2	th_3	th_4	th_5
Начальный набор данных	0.38	0.83	0.56	0.07	0.61	0.54	0.31
Количественно сбалансированный методом SMOTE	0.42	0.61	0.71	0.19	0.68	0.17	0.69
Количественно сбалансированный методом ADASYN	0.36	0.74	0.54	0.06	0.74	0.41	0.37
Количественно сбалансированный методом Gaussian Copula	0.43	0.61	0.7	0.2	0.68	0.18	0.67
Пространственно сбалансированный	0.49	0.66	0.6	0.41	0.59	0.31	0.5

В случае более сложной модели машинного обучения, а именно полносвязной нейронной сети, разница между количественно и пространственно сбалансированными наборами данных, представленная в таблице 15, не столь значительна. Однако пространственная балансировка все же улучшает показатели количественной. Так можно отметить, что пространственная балансировка повысила минимальную точность определения th_5 -класса (относительно методов SMOTE и Gaussian Copula. ADASYN имеет высокую точность определения th_5 -класса, но практически никогда не определяет корректно th_4 -класс). Рост составил 42% (с 20% до 62%) относительно начального набора данных и 2% (с 60% до 62%) относительно количественно сбалансированного набора методом SMOTE. Устойчивость различных методов балансировки данных оценивалась на модели нейронной сети. Для этого использовалась процедура 5-кратной кросс-валидации: исходный набор данных

разделялся на 5 равных по объёму подвыборок с сохранением исходного соотношения классов в каждой из них. На четырёх подвыборках (объединённых в обучающую часть) выполнялось обучение модели нейронной сети и применялись соответствующие методы балансировки. Полученная модель затем тестировалась на оставшейся подвыборке. Процедура повторялась для всех пятёрок, в результате чего выполнялось 5 вычислительных экспериментов.

Таблица 15. Результаты вычислительных экспериментов, проведенных нейронной сетью

Набор данных	m_{eff}	th_0	th_1	th_2	th_3	th_4	th_5
Начальный набор данных	0.46	0.65	0.72	0.35	0.51	0.58	0.2
Количественно сбалансированный методом SMOTE	0.47	0.66	0.68	0.24	0.58	0.31	0.6
Количественно сбалансированный методом ADASYN	0.12	0.66	0.71	0.31	0.51	0.005	0.81
Количественно сбалансированный методом Gaussian Copula	0.48	0.7	0.65	0.27	0.6	0.34	0.53
Пространственно сбалансированный	0.48	0.66	0.65	0.28	0.58	0.31	0.62

Для количественной оценки устойчивости методов балансировки использовалась дисперсия величины m_{eff} , рассчитанная по результатам кросс-валидации. Наименьшую дисперсию (наиболее высокую устойчивость) продемонстрировал метод *Gaussian Copula*, для которого дисперсия составила 0,023. Методы количественной балансировки *ADASYN* и *SMOTE* показали сопоставимую устойчивость – дисперсия m_{eff} составила 0,25.

Наиболее чувствительным к составу выборки оказался метод пространственной балансировки. При параметре $p = 7$ дисперсия достигала значения 0,3. При этом устойчивость данного метода существенно зависела от значения параметра p : при уменьшении p наблюдалось снижение дисперсии. Так, при $p = 4$ она составила 0,25.

Полученные результаты подтверждают, что метод пространственной балансировки требует дополнительной настройки параметров и валидации для обеспечения устойчивости, в то время как

генеративные методы, такие как Gaussian Copula, демонстрируют более стабильное поведение.

8. Заключение. Таким образом, разработанный метод формирования обучающего набора данных на основе малой несбалансированной выборки обеспечил повышение качества данных и результатов многоклассовой классификации как по отношению к исходной выборке, так и к количественно сбалансированной. Такой эффект объясняется тем, что синтез объектов проводился именно в информативно значимых областях, что отличает данную методику от других существующих подходов синтеза.

Отметим, что обе рассматриваемые в работе модели классификации, а именно логистическая регрессия и полносвязная нейронная сеть, в результате пространственной балансировки достигли схожей эффективности в определении th -классов. Таким образом, при пространственной балансировке происходит снижение требований к моделям машинного обучения, благодаря чему становится возможным использование более интерпретируемых моделей в системах искусственного интеллекта. При этом повышение репрезентативности позволяет также улучшить качество обоснования по интерпретируемым моделям за счет более адекватного знания о предметной области, имеющегося у моделей.

Существенным фактором, влияющим на эффективность метода, является выбор распределений, к которым приводится обучающая выборка. В данной работе наилучшим распределением оказалось нормальное. Мы предполагаем, что в силу своего широкого распространения, зачастую именно оно будет являться адекватным выбором, если из предметной области не следует что-то другое.

Отметим, что в результате применения метода повышения репрезентативности также происходит снижение требований к изначальному обучающему набору данных, а именно к его объему и исходной репрезентативности. Описанный в работе метод позволяет повышать качество наборов данных по этим характеристикам на основе знаний предметной области. Таким образом, потенциально это может способствовать расширению применимости методов машинного обучения в областях, для которых характерны малые объемы данных, а именно в медицине, физике, химии и т.п.

Приложение

Алгоритм 1. Метод построения информативных интервалов

```

1: for all  $F$  in  $Functional\_set$  do
2:   for all  $y$  in  $class\_set$  do
3:     if  $y = 0$  then
4:        $start\_point \leftarrow 0$ 
5:     else
6:        $start\_point \leftarrow$  медианное значение  $F$  в классе  $y$ 
7:     end if
8:      $\delta \leftarrow$  минимальное расстояние между значениями  $F$  в классе  $y$ 
9:      $direction \leftarrow$  направление расширения интервала для  $F$  по классу  $y$ 
10:     $previos\_interval\_W \leftarrow 0$ 
11:     $interval\_W \leftarrow 0$ 
12:     $i\_start \leftarrow start\_point$ 
13:     $i\_end \leftarrow start\_point$ 
14:     $min\_F \leftarrow$  минимальное значение  $F$  в классе  $y$ 
15:     $max\_F \leftarrow$  максимальное значение  $F$  в классе  $y$ 
16:    while ( $interval\_W \geq previos\_interval\_W$ ) and ( $i\_start \geq min\_F$ ) and ( $i\_end \leq max\_F$ ) do
17:      if  $direction = 'one\ side'$  then
18:         $i\_end \leftarrow i\_end + \delta$ 
19:      else
20:         $i\_start \leftarrow i\_start - \delta$ 
21:         $i\_end \leftarrow i\_end + \delta$ 
22:      end if
23:       $interval \leftarrow [i\_start; i\_end]$ 
24:       $interval\_W \leftarrow$  вычисленная информативность интервала
25:    end while
26:  end for
27: end for

```

Алгоритм 2. Метод пространственной балансировки

```

1: for all  $F$  in  $Functional\_set$  do
2:   for all  $y$  in  $class\_set$  do
3:      $i \leftarrow$  информативный интервал значений  $F$  по классу  $y$ 
4:      $min\_diff \leftarrow 1$ 
5:     for  $subinterval = 1, \dots, m$  do
6:        $p \leftarrow$  имеющаяся доля значений  $F$  в подинтервале
7:        $required\_p \leftarrow$  требуемая доля значений  $F$  в подинтервале
8:       if  $required\_p - p < min\_diff$  then
9:          $min\_diff \leftarrow required\_p - p$ 
10:         $a\_e \leftarrow$  имеющееся число значений  $F$  в подинтервале
11:         $required\_p\_e \leftarrow required\_p$ 
12:      end if
13:    end for
14:     $A \leftarrow a\_e \div required\_p\_e$ 
15:    for  $subinterval = 1, \dots, m$  do
16:       $a \leftarrow$  имеющееся число значений  $F$  в подинтервале
17:       $t \leftarrow int(A \cdot required\_p) - a$ 
18:      синтезируем  $t$  объектов с классом  $y$  и значениями  $F$  внутри
        подинтервала
19:    end for
20:  end for
21: end for

```

Литература

1. Jakubik J., Vossing M., Kuhl N., Walk J., Satzger G. Data-Centric Artificial Intelligence // Business & Information Systems Engineering. 2024. vol. 66. no. 4. pp. 507–515.
2. Kumar S., Datta S., Singh V., Singh S.K., Sharma R. Opportunities and Challenges in Data-Centric AI // IEEE Access. 2024. vol. 12. pp. 33173–33189.
3. Hamid O.H. Data-Centric and Model-Centric AI: Twin Drivers of Compact and Robust Industry 4.0 Solutions // Applied Sciences. 2023. vol. 13. no. 5. DOI: 10.3390/app13052753.
4. Emmanuel T., Maupong T., Mpoeleng D., Semong T., Mphago B., Tabona O. A survey on missing data in machine learning // Journal of Big Data. 2021. vol. 8. no. 1. DOI: 10.1186/s40537-021-00516-9.
5. Gerber S., Pospisil L., Sys S., Hewel C., Torkamani A., Horenko I. Co-Inference of Data Mislabeledings Reveals Improved Models in Genomics and Breast Cancer Diagnostics // Frontiers in Artificial Intelligence 2022. vol. 4. DOI: 10.3389/frai.2021.739432.
6. Feng W., Quan Y., Dauphin G. Label Noise Cleaning with an Adaptive Ensemble Method Based on Noise Detection Metric // Sensors. 2020. no. 20(23). DOI: org/10.3390/s20236718.
7. Чехомова О.В. О проблемах формирования административной ответственности за нарушения прав пациента // Юрист - Правоведъ. 2010. № 3. С. 120–123.

8. Собянин К.В., Куликова С.П. Unet-boosted classifier – мультизадачная архитектура для малых выборок на примере классификации МРТ снимков головного мозга // Информатика и автоматизация. 2024. Т. 23. № 4. С. 1022–1046.
9. Зуев Ю.Ф. Сравнительная оценка достоверности биологических данных при малых выборках методами дисперсионного, кластерного и графового анализа (нестохастический подход) // Актуальные проблемы гуманитарных и естественных наук. 2015. № 1-1. С. 36–45.
10. Моргунова А.В., Сажина О.С. Использование непараметрических методов математической статистики для поиска космических струн // Вестник Московского университета. Серия 3. Физика. Астрономия. 2019. № 5. С. 86–92.
11. Мокеев В.В., Томилов С.В. О решении проблемы выборки малого размера при использовании линейного дискриминантного анализа в задачах распознавания лиц // Бизнес-информатика. 2013. № 1(23). С. 37–43.
12. Алексахина К.А., Жуков В.В. Особенности построения моделей машинного обучения на малых наборах данных в биомедицине // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. Российский университет дружбы народов, 2021. С. 151–155.
13. Hu Y., Sandt R., Spatschek R. Practical feature filter strategy to machine learning for small datasets in chemistry // Scientific Reports. 2024. vol. 14. no. 1. DOI: 10.1038/s41598-024-71342-1.
14. Safonova A., Ghazaryan G., Stiller S., Main-Knorn M., Nendel C., Ryo M. Ten deep learning techniques to address small data problems with remote sensing // International Journal of Applied Earth Observation and Geoinformation. 2023. vol. 125. DOI: 10.1016/j.jag.2023.103569.
15. He H., Garcia E.A. Learning from Imbalanced Data // IEEE Transactions on Knowledge and Data Engineering. 2009. vol. 21. no. 9. pp. 1263–1284. DOI: 10.1109/TKDE.2008.239.
16. Polyakov M., Sirotin D. A Realistic Breast Phantom for Investigating the Features of the Microwave Radiometry Method Using Mathematical and Physical Modelling // Technologies. 2025. vol. 13(3). DOI: 10.3390/technologies13030106.
17. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique // Journal of Artificial Intelligence Research. 2002. vol. 16. pp. 321–357. DOI: 10.1613/jair.953.
18. Hauptmann T., Fellenz S., Nathan L., Tuschler O., Kramer S. Discriminative Machine Learning for Maximal Representative Subsampling // Scientific Reports. 2023. vol. 13. no. 1. DOI: 10.1038/s41598-023-48177-3.
19. Priestley M., O'Donnell F., Simperl E. A Survey of Data Quality Requirements That Matter in ML Development Pipelines // ACM Journal of Data and Information Quality. 2023. vol. 15. no. 2. pp. 1–39. DOI: 10.1145/3592616.
20. Levshinskii V. Mathematical models for analyzing and interpreting microwave radiometry data in medical diagnosis // Journal of computational and engineering mathematics. 2021. vol. 8. no. 1. pp. 3–14.
21. Лебедев А.В., Фадеева Л.Н. Теория вероятностей и математическая статистика (4-е издание) // М.: Наука. 2018. 408 с.
22. Kwon S., Lee S. Recent Advances in Microwave Imaging for Breast Cancer Detection // International journal of biomedical imaging. 2016. vol. 2016. pp. 1–25. DOI: 10.1155/2016/5054912.
23. Barrett A.H., Myers P.C. Subcutaneous Temperature: A method of Noninvasive Sensing // Science. 1975. vol. 190. pp. 669–671. DOI: 10.1126/science.1188361.

24. Barrett A.H., Myers Ph.C., Sadovsky N.L. Microwave Thermography in the Detection of Breast Cancer // AJR. American journal of roentgenolog. 1980. vol. 134(2). pp. 365–368. DOI: 10.2214/ajr.134.2.365.
25. Gautherie M. Temperature and Blood Flow Patterns in Breast Cancer During Natural Evolution and Following Radiotherapy // Progress in Clinical and Biological Research. 1982. pp. 21–64.
26. Троицкий В.С. К теории контактных радиометрических измерений внутренней температуры тел // Радиофизика. 1981. Т. 2. № 9. С.1054–1061.
27. Goryanin I., Karbainov S., Shevelev O., Tarakanov A., Redpath K., Vesnin S., Ivanov Yu. Passive microwave radiometry in biomedical studies // Drug Discovery Today. 2020. vol. 25. no. 4. pp. 757–763.
28. Попов И.Е., Крылова А.Е. Анализ термометрических данных головного мозга, полученных методом микроволновой радиотермометрии // Математическая физика и компьютерное моделирование. 2023. Т. 26. № 2. С. 32–42. DOI: 10.15688/mpcm.jvolsu.2023.2.3.
29. Sedankin M.K., Gudkov A.G., Nelin I.V., Skuratov V.A., Leushin V.Yu, Chizhikov S.V., Sidorov I.A. Mathematical modeling of heat exchange processes in the brain in the presence of pathology for the design of a microwave radiothermograph // Bio-Medical Engineering. 2023. vol. 57. no. 4. pp. 276–279. DOI: 10.1007/s10527-023-10314-7.
30. Li J., Galazis C., Popov L., Ovchinnikov L., Kharybina T., Vesnin S., Losev A., Goryanin I Dynamic Weight Agnostic Neural Networks and Medical Microwave Radiometry (MWR) for Breast Cancer Diagnostics // Diagnostics. 2022. vol. 12. no. 9. DOI: 10.3390/diagnostics12092037.
31. Веснин С.Г., Каплан М.А., Авакян Р.С. Современная микроволновая радиотермометрия молочных желез // Опухоли женской репродуктивной системы. 2008. № 3. С. 28–33. DOI: 10.17650/1994-4098-2008-0-3-28-33.
32. Лосев А.Г., Левшинский В.В. Интеллектуальный анализ данных микроволновой радиотермометрии в диагностике рака молочной железы // Математическая физика и компьютерное моделирование. 2017. Т. 20. № 5. С. 49–62. DOI: 10.15688/mpcm.jvolsu.2017.5.6.
33. Лосев А.Г., Левшинский В.В. Интеллектуальный анализ термометрических данных в диагностике молочных желез // Управление большими системами: сборник трудов. 2017. № 70. С. 113–135.
34. Levshinskii V., Galazis C., Ovchinnikov L., Vesnin S., Goryanin I. Application of Data Mining and Machine Learning in Microwave Radiometry (MWR) // Communications in Computer and Information Science (CCIS). 2020. vol. 1211. pp. 265–288. DOI: 10.1007/978-3-030-46970-2_13.
35. Гермашев И.В., Дубовская В.И. Применение моделей нечеткой математики для решения задач медицинской диагностики // Математическая физика и компьютерное моделирование. 2021. Т. 24. № 4. С. 53–66. DOI: 10.15688/mpcm.jvolsu.2021.4.4.
36. Osmonov B., Ovchinnikov L., Galazis C., Emilov B., Karaibragimov M., Seitov M., Vesnin S., Losev A., Levshinskii V., Popov I., Mustafin C., Kasymbekov T., Goryanin I. Passive Microwave Radiometry for the Diagnosis of Coronavirus Disease 2019 Lung Complications in Kyrgyzstan // Diagnostics. 2021. vol. 11. no. 2. DOI: 10.3390/diagnostics11020259.
37. Лосев А.Г., Попов И.Е., Гудков А.Г., Чижиков С. В. Интеллектуальный анализ данных микроволновой радиотермометрии в медицинской диагностике // Нанотехнологии: разработка, применение – XXI век. 2023. Т. 15. № 1. С. 5–22. DOI: 10.18127/j22250980-202301-01.

38. Ahmad H.F., Mukhtar H., Alaqail H., Seliaman M., Alhumam A. Investigating Health-Related Features and Their Impact on the Prediction of Diabetes Using Machine Learning // Applied Sciences. 2021. vol. 11. no. 3. DOI: 10.3390/app11031173.

Лосев Александр Георгиевич — д-р физ.-мат. наук, профессор, директор института, институт математики и информационных технологий, Волгоградский государственный университет. Область научных интересов: искусственный интеллект, интеллектуальный анализ данных, математическое моделирование, эллиптические уравнения в частных производных на римановых многообразиях. Число научных публикаций — 191. alexander.losev@volsu.ru; проспект Университетский, 100, 400062, Волгоград, Россия; р.т.: +7(904)774-6821.

Попов Илларион Евгеньевич — аспирант, кафедра математического анализа и теории функций, Волгоградский государственный университет. Область научных интересов: искусственный интеллект, интеллектуальный анализ данных, математическое моделирование. Число научных публикаций — 28. popov.larion@volsu.ru; проспект Университетский, 100, 400062, Волгоград, Россия; р.т.: +7(937)088-8938.

Резникова Анастасия Сергеевна — магистрант, кафедра математического анализа и теории функций, Волгоградский государственный университет. Область научных интересов: искусственный интеллект, интеллектуальный анализ данных, математическое моделирование. Число научных публикаций — 11. a.reznikova@volsu.ru; проспект Университетский, 100, 400062, Волгоград, Россия; р.т.: +7(917)845-2535.

Поддержка исследований. Исследование выполнено за счет гранта Российского научного фонда № 25-21-00330, <https://rscf.ru/project/25-21-00330/>.

A. LOSEV, I. POPOV, A. REZNIKOVA
**IMPROVING THE REPRESENTATIVENESS OF THE TRAINING
DATASET BY MEANS OF SPATIAL BALANCING**

Losev A., Popov I., Reznikova A. Improving the Representativeness of the Training Dataset by Means of Spatial Balancing.

Abstract. The study examines the specifics of training machine learning algorithms on small datasets and addresses the task of forming a training set with high representativeness. It is known that class imbalance in objects, typical for small datasets, negatively affects the performance of algorithms. To mitigate this issue, various data synthesis methods have been developed in machine learning to supplement existing datasets and equalize the number of objects per class. However, these methods do not solve the problem of insufficient representativeness. This article proposes a method for constructing a representative training dataset by specifying the distribution that best corresponds to reality. The distribution is formed for each feature within the informative areas. Informative areas contain characteristic values of features that are most significant for distinguishing classes of objects. The proposed method of constructing areas is based on the idea of gradual expansion, accompanied by an increase in the informativeness of the areas. At the same time, informativeness is understood as a measure reflecting how well objects of different classes can be separated using the considered area. To form a complementary dataset, a generation method has been developed. As a result of its application, the complementary dataset is combined with the original one and forms the specified distribution in the informative area. This distribution can be determined either based on expert knowledge about the subject area, if the true distribution is known, or obtained as a result of computational experiments aimed at finding the most effective option. The applicability of the method is demonstrated by solving the problem of determining the level of temperature anomalies of the mammary glands. It is shown that the considered temperature features are characterized by a normal distribution. Increasing the representativeness of the training set allowed training a classic classification algorithm – logistic regression – with an accuracy comparable to a multilayer neural network. This approach to the formation of a training dataset opens up the possibility of creating more transparent and interpretable artificial intelligence systems.

Keywords: machine learning, small datasets, data representativeness, data synthesis, neural networks, logistic regression.

References

1. Jakubik J., Vossing M., Kuhl N., Walk J. Satzger G. Data-Centric Artificial Intelligence. Business & Information Systems Engineering. 2024. vol. 66. no. 4. pp. 507–515.
2. Kumar S., Datta S., Singh V., Singh S.K., Sharma R. Opportunities and Challenges in Data-Centric AI. IEEE Access. 2024. vol. 12. pp. 33173–33189.
3. Hamid O.H. Data-Centric and Model-Centric AI: Twin Drivers of Compact and Robust Industry 4.0 Solutions. Applied Sciences. 2023. vol. 13. no. 5. DOI: 10.3390/app13052753.
4. Emmanuel T., Maupong T., Mpoeleng D., Semong T., Mphago B., Tabona O. A survey on missing data in machine learning. Journal of Big Data. 2021. vol. 8. no. 1. DOI: 10.1186/s40537-021-00516-9.
5. Gerber S, Pospisil L, Sys S, Hewel C, Torkamani A, Horenko I Co-Inference of Data Mislabelings Reveals Improved Models in Genomics and Breast Cancer Diagnostics. Frontiers in Artificial Intelligence 2022. vol. 4. DOI: 10.3389/frai.2021.739432.

6. Feng W., Quan Y., Dauphin G. Label Noise Cleaning with an Adaptive Ensemble Method Based on Noise Detection Metric. *Sensors*. 2020. no. 20(23). DOI: [org/10.3390/s20236718](https://doi.org/10.3390/s20236718).
7. Chehomova O.V. [On the problems of forming administrative responsibility for violations of patient rights]. *Jurist - Pravoved – Jurist - Lawyer*. 2010. no. 3. pp. 120–123. (In Russ.).
8. Sobjanin K.V., Kulikova S.P. [Unet-boosted classifier – a multi-task architecture for small samples using the example of classifying MRI images of the brain]. *Informatics and automation*. 2024. vol. 23. no. 4. pp. 1022–1046. (In Russ.).
9. Zuev Ju.F. [Comparative assessment of the reliability of biological data in small samples using methods of dispersion, cluster and graph analysis (non-stochastic approach)]. *Aktual'nye problemy gumanitarnykh i estestvennykh nauk – Current issues in the humanities and natural sciences*. 2015. no. 1-1. pp. 36–45. (In Russ.).
10. Morgunova A.V., Sazhina O.S. [Using Nonparametric Methods of Mathematical Statistics to Search for Cosmic Strings]. *Vestnik Moskovskogo universiteta. Seriya 3. Fizika. Astronomiya – Bulletin of Moscow University. Series 3. Physics. Astronomy*. 2019. no. 5. pp. 86–92. (In Russ.).
11. Mokeev V.V., Tomilov S.V. [On the solution of the problem of small sample size when using linear discriminant analysis in face recognition problems]. *Biznes-informatika – Business informatics*. 2013. no. 1(23). pp. 37–43. (In Russ.).
12. Aleksahina K.A., Zhukov V.V. Osobennosti postroeniya modelej mashinnogo obucheniya na mal'nykh naborakh dannykh v biomedicine [Features of building machine learning models on small data sets in biomedicine]. *Informacionno-telekommunikacionnye tehnologii i matematicheskoe modelirovanie vysokotekhnologichnykh system [Information and telecommunication technologies and mathematical modeling of high-tech system]*. Peoples' Friendship University of Russia, 2021. pp. 151–155. (In Russ.).
13. Hu Y., Sandt R., Spatschek R. Practical feature filter strategy to machine learning for small datasets in chemistry. *Scientific Reports*. 2024. vol. 14. no. 1. DOI: [10.1038/s41598-024-71342-1](https://doi.org/10.1038/s41598-024-71342-1).
14. Safonova A., Ghazaryan G., Stiller S., Main-Knorn M., Nendel C., Ryo M. Ten deep learning techniques to address small data problems with remote sensing. *International Journal of Applied Earth Observation and Geoinformation*. 2023. vol. 125. DOI: [10.1016/j.jag.2023.103569](https://doi.org/10.1016/j.jag.2023.103569).
15. He H., Garcia E.A. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*. 2009. vol. 21. no. 9. pp. 1263–1284. DOI: [10.1109/TKDE.2008.239](https://doi.org/10.1109/TKDE.2008.239).
16. Polyakov M., Sirotin D. A Realistic Breast Phantom for Investigating the Features of the Microwave Radiometry Method Using Mathematical and Physical Modelling. *Technologies*. 2025. vol. 13(3). DOI: [10.3390/technologies13030106](https://doi.org/10.3390/technologies13030106).
17. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*. 2002. vol. 16. pp. 321–357. DOI: [10.1613/jair.953](https://doi.org/10.1613/jair.953).
18. Hauptmann T., Fellenz S., Nathan L., Tuscher O., Kramer S. Discriminative Machine Learning for Maximal Representative Subsampling. *Scientific Reports*. 2023. vol. 13. no. 1. DOI: [10.1038/s41598-023-48177-3](https://doi.org/10.1038/s41598-023-48177-3).
19. Priestley M., O'Donnell F., Simperl E. A Survey of Data Quality Requirements That Matter in ML Development Pipelines. *ACM Journal of Data and Information Quality*. 2023. vol. 15. no. 2. pp. 1–39. DOI: [10.1145/3592616](https://doi.org/10.1145/3592616).
20. Levshinskii V. Mathematical models for analyzing and interpreting microwave radiometry data in medical diagnosis. *Journal of computational and engineering mathematics*. 2021. vol. 8. no. 1. pp. 3–14.

21. Lebedev A.V., Fadeeva L.N. Teorija verojatnostej i matematicheskaja statistika (4-e izdanie) [Probability Theory and Mathematical Statistics (4th edition)]. M: Nauka, 2018. 408 p. (In Russ.).
22. Kwon S., Lee S. Recent Advances in Microwave Imaging for Breast Cancer Detection. International journal of biomedical imaging. 2016. vol. 2016. pp. 1–25. DOI: 10.1155/2016/5054912.
23. Barrett A.H., Myers P.C. Subcutaneous Temperature: A method of Noninvasive Sensing. Science. 1975. vol. 190. pp. 669–671. DOI: 10.1126/science.1188361.
24. Barrett A.H., Myers Ph.C., Sadovsky N.L. Microwave Thermography in the Detection of Breast Cancer. AJR. American journal of roentgenolog. 1980. vol. 134(2). pp. 365–368. DOI: 10.2214/ajr.134.2.365.
25. Gautherie M. Temperature and Blood Flow Patterns in Breast Cancer During Natural Evolution and Following Radiotherapy. Progress in Clinical and Biological Research. 1982. pp. 21–64.
26. Troickij V.S. [On the theory of contact radiometric measurements of the internal temperature of bodies]. Radiofizika – Radiophysics. 1981. vol. 2. no. 9. pp. 1054–1061. (In Russ.).
27. Goryanin I., Karbainov S., Shevelev O., Tarakanov A., Redpath K., Vesnin S., Ivanov Yu. Passive microwave radiometry in biomedical studies. Drug Discovery Today. 2020. vol. 25. no. 4. pp. 757–763.
28. Popov I.E., Krylova A.E. [Analysis of brain thermometric data obtained by microwave radiothermometry]. Matematicheskaja fizika i kompyuternoe modelirovanie – Mathematical Physics and Computer Modeling. 2023. vol. 26. no. 2. pp. 32–42. DOI: 10.15688/mpcm.jvolsu.2023.2.3. (In Russ.).
29. Sedankin M.K., Gudkov A.G., Nelin I.V., Skuratov V.A., Leushin V.Yu., Chizhikov S.V., Sidorov I.A. Mathematical modeling of heat exchange processes in the brain in the presence of pathology for the design of a microwave radiothermograph. Bio-Medical Engineering. 2023. vol. 57. no. 4. pp. 276–279. DOI: 10.1007/s10527-023-10314-7.
30. Li J., Galazis C., Popov L., Ovchinnikov L., Kharybina T., Vesnin S., Losev A., Goryanin I. Dynamic Weight Agnostic Neural Networks and Medical Microwave Radiometry (MWR) for Breast Cancer Diagnostics. Diagnostics. 2022. vol. 12. no. 9. DOI: 10.3390/diagnostics12092037.
31. Vesnin S.G., Kaplan M.A., Avakjan R.S. [Modern microwave radiothermometry of mammary glands]. Opuholi zhenskoj reproduktivnoj sistemy – Female reproductive system tumors. 2008. no. 3. pp. 28–33. DOI: 10.17650/1994-4098-2008-0-3-28-33. (In Russ.).
32. Losev A.G., Levshinskii V.V. [Microwave radiothermometry data mining in breast cancer diagnostics]. Matematicheskaja fizika i komp'yuternoe modelirovanie – Mathematical Physics and Computer Modeling. 2017. vol. 20. no. 5. pp. 49–62. DOI: 10.15688/mpcm.jvolsu.2017.5.6. (In Russ.).
33. Losev A.G., Levshinskii V.V. [Intelligent analysis of thermometric data in breast diagnostics]. Upravlenie bol'shimi sistemami: sbornik trudov – Management of large systems: collection of works. 2017. no. 70. pp. 113–135. (In Russ.).
34. Levshinskii V., Galazis C., Ovchinnikov L., Vesnin S., Goryanin I. Application of Data Mining and Machine Learning in Microwave Radiometry (MWR). Communications in Computer and Information Science (CCIS). 2020. vol. 1211. pp. 265–288. DOI: 10.1007/978-3-030-46970-2_13.
35. Germashev I.V., Dubovskaja V.I. [Application of fuzzy mathematics models for solving medical diagnostics problems]. Matematicheskaja fizika i kompyuternoe modelirovanie

- Mathematical Physics and Computer Modeling. 2021. vol. 24. no. 4. pp. 53–66. DOI: 10.15688/mpcm.jvolsu.2021.4.4.
- 36. Osmonov B., Ovchinnikov L., Galazis C., Emilov B., Karaibragimov M., Seitov M., Vesnin S., Losev A., Levshinskii V., Popov I., Mustafin C., Kasymbekov T., Goryanin I. Passive Microwave Radiometry for the Diagnosis of Coronavirus Disease 2019 Lung Complications in Kyrgyzstan. *Diagnostics*. 2021. vol. 11. no. 2. DOI: 10.3390/diagnostics11020259.
- 37. Losev A.G., Popov I.E., Gudkov A.G., Chizhikov S. V. [Intelligent analysis of microwave radiothermometry data in medical diagnostics]. *Nanotehnologii: razrabotka, primeneniye - XXI vek – Nanotechnology: development, application - 21st century*. 2023. vol. 15. no. 1. pp. 5–22. DOI: 10.18127/j22250980-202301-01. (In Russ.).
- 38. Ahmad H.F., Mukhtar H., Alaqail H., Seliaman M., Alhumam A. Investigating Health-Related Features and Their Impact on the Prediction of Diabetes Using Machine Learning. *Applied Sciences*. 2021. vol. 11. no. 3. DOI: 10.3390/app11031173.

Losev Alexandr — Ph.D., Dr.Sci., Professor, Head of the institute, Institute of mathematics and information technology, Volgograd State University. Research interests: artificial intelligence, data mining, mathematical modeling, elliptic partial differential equations on Riemannian manifolds. The number of publications — 191. alexander.losev@volsu.ru; 100, Universitetsky Av., 400062, Volgograd, Russia; office phone: +7(904)774-6821.

Popov Illarion — Postgraduate student, Department of mathematical analysis and theory of functions, Volgograd State University. Research interests: artificial intelligence, data mining, mathematical modeling. The number of publications — 28. popov.ilarion@volsu.ru; 100, Universitetsky Av., 400062, Volgograd, Russia; office phone: +7(937)088-8938.

Reznikova Anastasia — Master's student, Department of mathematical analysis and theory of functions, Volgograd State University. Research interests: artificial intelligence, data mining, mathematical modeling. The number of publications — 11. a.reznikova@volsu.ru; 100, Universitetsky Av., 400062, Volgograd, Russia; office phone: +7(917)845-2535.

textbfAcknowledgements. This work is supported by the Russian Science Foundation No. 25-21-00330, <https://rscf.ru/project/25-21-00330/>.

А.В. ЕРПАЛОВ, В.В. СИНИЦИН, А.Л. ШЕСТАКОВ
**EWT-CGAN АУГМЕНТАЦИЯ ДАННЫХ ИЗМЕРИТЕЛЬНЫХ
СИСТЕМ**

Ерпалов А.В., Синицин В.В., Шестаков А.Л. EWT-CGAN аугментация данных измерительных систем.

Аннотация. В статье представлен новый метод аугментации данных измерительных систем, разработанный для задач мониторинга состояния промышленного оборудования. Актуальность исследования обусловлена существенными ограничениями традиционных методов генерации синтетических данных, которые не способны адекватно воспроизводить сложные нестационарные сигналы с характерными переходными процессами, трендами и сезонными вариациями, наблюдаемыми в реальных промышленных условиях. Предложенный метод основан на интеграции двух современных методов: эмпирического вейвлет-преобразования (EWT) и условных генеративных состязательных сетей (Conditional GAN). Метод реализуется в три этапа: (1) адаптивная декомпозиция исходных сигналов на моды с помощью EWT, (2) категоризация мод с присвоением меток, (3) генерация синтетических данных с использованием Conditional GAN. Для комплексной оценки качества синтезированных сигналов применялся набор статистических метрик, включая расстояние Wasserstein (WS), коэффициент корреляции Пирсона (PCC) и среднеквадратическую ошибку (RMSE). Экспериментальные исследования проводились на реальных данных температурного датчика, работающего в условиях нестационарных режимов промышленного оборудования. Результаты демонстрируют значительное преимущество предложенного метода по сравнению с традиционным подходом timeGAN: снижение расстояния Wasserstein на 17%, увеличение коэффициента корреляции Пирсона на 57% и уменьшение среднеквадратической ошибки на 21%. Полученные данные свидетельствуют об эффективности метода в воспроизведении ключевых характеристик исходных сигналов. Разработанный метод позволяет создавать набор синтетических данных, необходимых для обучения современных нейросетевых моделей диагностики промышленного оборудования. Его практическое применение позволяет существенно сократить затраты на сбор экспериментальных данных, обеспечивая высокое качество синтезированных сигналов, что подтверждено статистическими метриками.

Ключевые слова: диагностика оборудования, сигналы с датчиков, аугментация, синтетические данные, эмпирическое вейвлет преобразование, генеративные состязательные сети.

1. Введение. В последние годы мониторинг состояния промышленного оборудования становится критически важным для предотвращения аварий и оптимизации производственных процессов. Однако сбор достаточного объема экспериментальных данных для обучения нейросетевых моделей технического состояния оборудования часто затруднен из-за высокой стоимости и длительности таких экспериментов. В этой связи актуальной задачей является разработка методов аугментации данных, способных генерировать синтетические сигналы, близкие к реальным, но не требующих дополнительных затрат на их сбор.

Традиционные методы генерации синтетических данных, такие как простое добавление шума или интерполяция, часто оказываются недостаточно эффективными для воспроизведения сложных нестационарных сигналов, характерных для промышленного оборудования. Эти сигналы могут содержать переходные процессы, тренды и сезонные вариации, которые трудно смоделировать стандартными подходами.

Целью данного исследования является разработка метода аугментации данных измерительных систем, способного решить проблему генерации нестационарных сигналов с сохранением их ключевых характеристик, таких как переходные процессы, тренды и сезонные изменения.

Одним из решением данной проблемы является применение моделей на основе глубоких нейронных сетей, в особенности генеративные сети, которые все больше используются в разных сферах деятельности, включая анализ временных рядов средств измерений в задачах мониторинга состояния промышленного оборудования. Опубликовано большое количество работ, в том числе обзоры по существующим вариациям генеративных состязательных нейронных сетей для разных областей применения [1 – 3].

Модели глубокого обучения типа GAN состоят из двух нейронных сетей: генератора и дискриминатора. Генератор G принимает случайный шум и пытается сгенерировать синтетические данные, распределение которых схоже с обучающим набором данных. Дискриминатор D пытается определить, являются ли сгенерированные данные реальными или «поддельными». Генератор стремится максимизировать частоту отказов дискриминатора, тогда как дискриминатор стремится минимизировать ее. На рисунке 1 показан общий вид архитектуры GAN.

GAN относится к семейству генеративных моделей и является альтернативным методом генерации синтетических данных, который не требует экспертных знаний в предметной области. Генеративные состязательные нейронные сети впервые предложены в работе [4] в 2014 году, где многослойный персептрон использовался как для дискриминатора, так и для генератора. В 2015 году авторы работы [5] предложили глубокую сверточную генеративную состязательную сеть (DCGAN) для генерации синтетических изображений. В дальнейшем исследователи постоянно совершенствовали ранние архитектуры GAN, функции потерь и метрики оценки, одновременно внедряя их в прикладные задачи.

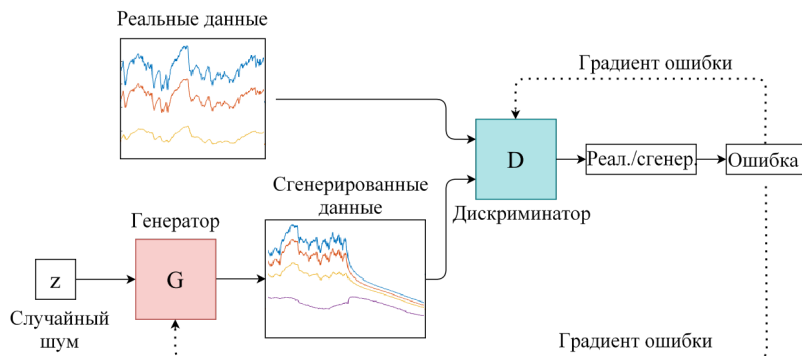


Рис. 1. Общий вид архитектуры генеративной состязательной нейронной сети

В настоящее время разработано большое количество вариаций GAN, однако каждая вариация направлена на решение задач в определенной сфере деятельности. Объясняется это тем, что традиционный GAN имеет недостатки, связанные с нестабильностью обучения из-за исчезновения градиента и коллапса мод. Основные пути решения этих недостатков: теоретические подходы, таких как изменение функции потерь; подходы, основанные на изменениях структур нейронных сетей, включая внутреннюю структуру [6] или добавления условий для входных данных [7].

Для изменения функции потерь авторы работы [8] предложили применять GAN совместно с методом наименьших квадратов (LSGAN). Такой подход предполагает использование в функции потерь метода наименьших квадратов для улучшения стабильности и скорости сходимости. Для изменений внутренней структуры и достижения более стабильной структуры авторы работы [5] предложили применение неконтролируемой глубокой сверточной GAN (DCGAN), которая работает на основе интеграции сверточных нейронных сетей с GAN. Авторы другой работы [6] предложили прогрессивно растущую GAN (PGGAN), использующую подход постепенного роста для генерации данных с высоким разрешением.

Отдельно важно затронуть генеративные состязательные нейронные сети с условиями (Conditional GAN), разработанные авторами работы [7]. Авторы предложили генератор и дискриминатор дополнить вспомогательной информацией (категориальные метки, модальность). Другими словами, в генератор и дискриминатор добавили дополнительный входной слой. Применяя такой метод

генерации данных, можно с помощью одной модели создать несколько наборов данных, разделенных на классы (категории).

Одной из главных проблем генеративных состязательных нейронных сетей является определение метрики, по которой дискриминатор оценивает, походит ли сгенерированный сигнал на исходный. Обычно это делается на основе сравнения распределений двух сигналов (исходного и сгенерированного). Авторы работы [9] предложили в качестве метрики использовать расстояние Wasserstein первого порядка (w_1), которое рассчитывается по следующей зависимости

$$w_1(u, v) = \int_{-\infty}^{+\infty} |U_{cdf}(x) - V_{cdf}(x)| dx, \quad (1)$$

где U_{cdf} и V_{cdf} – функции распределений, соответствующие плотностям распределений u и v .

Разработанный метод Wasserstein GAN [9] позволил решить основные проблемы обучения традиционного GAN. В частности, в процессе обучения WGAN не требует строгого баланса между дискриминатором и генератором, что в свою очередь позволяет использовать более простые структуры нейронных сетей. Кроме того, WGAN решает вопрос с явлением отбрасывания мод, которое часто проявляется при использовании традиционного GAN.

На основе описанных вариаций генеративных состязательных нейронных сетей предложены гибридные модели, например Conditional Wasserstein GAN [10]. Модель позволяет генерировать данные по условиям с хорошей сходимостью за счет применения расстояния Wasserstein в дискриминаторе.

Вариации GAN, указанные выше, в основном направлены на обработку изображений, но предоставленные фундаментальные структуры применяются также к сигналам временных рядов. В таком случае генеративный состязательные нейронные сети принято обозначать как «timeGAN».

Описанные выше методы аугментации подходят для генерации большинства сигналов средств измерений. Однако, несмотря на это существует ряд сигналов, генерация которых существующими методами будет некорректна и может привести к сигналам сильно отличающимся от исходных. К таким сигналам могут относиться данные средств измерений с характерными нестационарными переходными участками, сезонными изменениями. Обычно это

сигналы с датчиков, установленных на промышленном оборудовании, которое подвержено внезапным остановкам и имеет нестационарные режимы работы.

К примеру, подобными качествами обладают сигналы с датчиков температуры, установленные на объекте с нестационарными режимами работы. Обычно на временной реализации температуры хорошо дифференцируются зоны включения оборудования и работа под нагрузкой (рост температуры), зоны снятия нагрузки (снижение температуры), зоны временного останова оборудования и зоны длительного простоя оборудования (резкое падение температуры на всех сенсорах).

В работе предложена *EWT-CGAN* аугментация временных данных средств измерений промышленного оборудования на основе комплексного применения эмпирического вейвлет преобразования и генеративной состязательной нейронной сети.

2. Теоретические основы.

Метрики оценки качества синтезированных сигналов. Для оценки качества синтезированных сигналов существует широкий спектр метрик [11, 12], что свидетельствует об отсутствии общепринятого критерия. Обычно метрики подразделяются на количественные и качественные. Наиболее представительной качественной метрикой является органолептический метод. Однако в отличие от оценки изображений, оценить таким методом качество временных рядов сложно. Поэтому обычно проводят анализы *t-SNE* [13] и *PCA* [14] для визуализации того, насколько хорошо сгенерированные распределения напоминают исходные распределения [15].

Количественная оценка включает статистические метрики, используемые для анализа временных рядов, которые сведены в таблице 1. Эти метрики являются одними из наиболее часто используемых для оценки временных рядов и используются для оценки производительности GAN. В первую очередь связано это с тем, что они могут отражать сходство между данными обучения и синтетическими данными.

Здесь x_i – фактическое значение временного ряда x в момент времени i ; y_i – сгенерированное значение временного ряда y в момент времени i ; $\tilde{x}$ и $\tilde{y}$ – средние значения x и y соответственно; f_i выбирается исходя из сезонности и несезонности данных и принимается равным $y_i - M$ и $y_i - 1$ соответственно, где M – период x за сезон.

Таблица 1. Статистические метрики между предсказанными и фактическими значениями

Номер	Наименование	Зависимость
1	Коэффициент корреляции Пирсона	$PCC = \frac{\sum_{i=1}^N (x_i - \tilde{x})(y_i - \tilde{y})}{\sqrt{\sum_{i=1}^N (x_i - \tilde{x})^2 \sum_{i=1}^N (y_i - \tilde{y})^2}}$
2	Процентная среднеквадратическая разница	$PRD = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i)^2}}$
3	Среднеквадратическая ошибка	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2}$
4	Среднее значение квадратов разностей	$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2$
5	Средняя относительная ошибка	$MRE = \frac{1}{N} \sum_{i=1}^N \left \frac{x_i - y_i}{x_i - f_i} \right $
6	Средняя абсолютная ошибка	$MAE = \frac{1}{N} \sum_{i=1}^N x_i - y_i $

Помимо рассмотренных существуют метрики оценки производительности GAN на основе сходства между двумя изображениями. Большинство метрик успешно применяются и для данных временных рядов. К таким метрикам относятся Inception score (IS) [16], Frechet Inception Distance (FID) [17] и индекс структурного сходства (SSIM) [17].

Кроме того, существуют универсальные метрики, которые подходят для разных типов данных, в том числе и для временных рядов. Например, максимальное среднее расхождение (MMD) [18] является мерой сходства/различия между двумя распределениями вероятностей или расстояние Wasserstein (1).

Эмпирическое вейвлет преобразование. С развитием цифровой индустрии и оснащением предприятий большим количеством разных по типу средств измерений, возникла потребность в разработке новых методов обработки реальных данных. Конкретно методов, позволяющих работать с нестационарными сложными сигналами без предварительной экспертной оценки и сегментации сигнала.

Одной из современных тенденций в обработке нестационарных сигналов является применение новых методов декомпозиции сигналов, которые в отличие, например, от традиционного преобразования Фурье, декомпозируют сигнал не на заранее определенные базисы (синусоиды), а на некоторые процессы (детерминированные/ случайные), подобранные эмпирическим путем. Впервые подобный метод представлен в 1998 году Н. Хуангом в работе [19] – эмпирическая модовая декомпозиция (EMD). Позднее на основе этой идеи разработаны методы вариационной модовой декомпозиции [20], эмпирического вейвлет преобразования [21], эмпирического Фурье преобразования [22], которые решали вопросы ограниченного применения метода и неточной декомпозиции на моды.

В последние годы широкое распространение получил метод эмпирического вейвлет преобразования (EWT), предложенный Ж. Жилем в 2013 году [23]. Метод EWT заключается в построении семейства вейвлетов, адаптированных к обрабатываемому сигналу. С точки зрения преобразования Фурье – это эквивалентно построению набора полосовых фильтров. EWT сочетает в себе концепцию адаптивной декомпозиции EMD и обширную теорию вейвлет-преобразования.

Метод EWT – широко используемый метод сегментации сигналов. Принцип сегментации включает: адаптивное сегментирование спектра Фурье путем выделения локальных максимумов и построение набора полосовых фильтров пригодных для обработки сигналов и извлечения амплитудно-модулированных (*AM*) и частотно-модулированных (*FM*) компонентов сигнала.

Нормализованная ось Фурье, которая имеет периодичность 2π , разбита на N последовательных частей, т.е. $A_n = [\omega_{n-1}, \omega_n]$ ($\omega_0 = 0, \omega_N = \pi$), где ω_n – граничная точка между двумя частями и соответствующие значение это минимум между двумя соседними максимальными значениями в спектре Фурье. Очевидно, что $\bigcup_{n=1}^N A_n = [0, \pi]$. На рисунке 2 показана сегментация спектра с помощью EWT. Точка ω_n определяется как центральная точка A_n . Затем определяется ширина переходной области $T_n = 2\tau_n$.

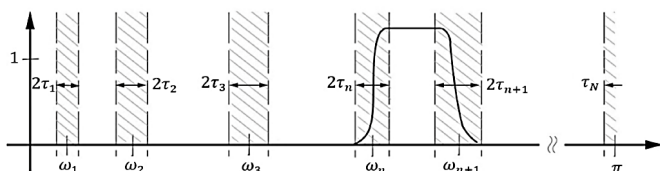


Рис. 2. Сегментация спектра с помощью EWT

Используя метод построения вейвлетов Литтлвуда-Пейли и Мейера, строится эмпирическая вейвлет функция. После определения A_n эмпирический вейвлет используется как полосовой фильтр. Эмпирическая вейвлет функция $\widehat{\psi}_n(\omega)$ и эмпирическая масштабная функция $\widehat{\phi}_n(\omega)$ определяются по формулам

$$\widehat{\phi}_n(\omega) = \begin{cases} 1, (|\omega| \leq (1-\gamma)\omega_n) \\ \cos \left\{ \frac{\pi}{2} \beta \left[\frac{1}{2\gamma\omega_n} (|\omega| - (1-\gamma)\omega_n) \right] \right\}, (1-\gamma)\omega_n \leq \omega \leq (1+\gamma)\omega_n, \\ 0, \text{others} \end{cases} \quad (2)$$

$$\widehat{\psi}_n(\omega) = \begin{cases} 1, (1+\gamma)\omega_n \leq |\omega| \leq (1-\gamma)\omega_{n+1} \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\omega_{n+1}} (|\omega| - (1-\gamma)\omega_{n+1}) \right) \right]; (1-\gamma)\omega_n \leq |\omega| \leq (1+\gamma)\omega_n \\ \sin \left[\frac{\pi}{2} \beta \left(\frac{1}{2\omega_n} (|\omega| - (1-\gamma)\omega_{n+1}) \right) \right]; (1-\gamma)\omega_n \leq |\omega| \leq (1+\gamma)\omega_n \\ 0, \text{others} \end{cases}, \quad (3)$$

где переходная функция $B(x)$, коэффициент γ и переходная фаза τ_n определяются из зависимостей (4).

$$\begin{aligned} \beta(x) &= x^4 (35 - 84x + 70x^2 - 20x^3), \\ \gamma &< \min_n \left(\frac{\omega_{n-1} - \omega_n}{\omega_{n-1} + \omega_n} \right), \quad \tau_n = \gamma\omega_n. \end{aligned} \quad (4)$$

После EWT коэффициент аппроксимации и коэффициент детализации могут быть выражены следующим образом

$$W_{f1}(0, t) = \langle x, \phi_1 \rangle = \int x(\tau) \overline{\phi_1(\tau - t)} d\tau = F^{-1} [x(\omega) \widehat{\phi}_1(\omega)], \quad (5)$$

$$W_{f2}(n, t) = \langle x, \psi_n \rangle = \int x(\tau) \overline{\psi_n(\tau - t)} d\tau = F^{-1} [x(\omega) \widehat{\psi}_n(\omega)]. \quad (6)$$

Тогда функциональное выражение восстановленного сигнала выглядит следующим образом

$$\begin{aligned} x(t) &= W_{f1}(0, t) * \varphi_1(t) + \sum_{n=1}^N W_{f2}(n, t) * \psi_n(t) = \\ &= F^{-1} \left[\widehat{W}_f(n, \omega) \varphi_1(\omega) + \sum_{n=1}^N \widehat{W}_f(n, \omega) \widehat{\psi}_n(\omega) \right], \end{aligned} \quad (7)$$

где «*» – оператор свертки, $W_{f1}(0, \omega)$ и $W_{f2}(n, \omega)$ – преобразование Фурье коэффициентов аппроксимации $W_{f1}(0, t)$ и $W_{f2}(n, t)$ соответственно. Сигнал f раскладывается в сумму нескольких однокомпонентных сигналов

$$x(t) = \sum_{k=0}^{N-1} x_k(t). \quad (8)$$

В общем виде блок-схема алгоритма метода EWT представлена на рисунке 3.

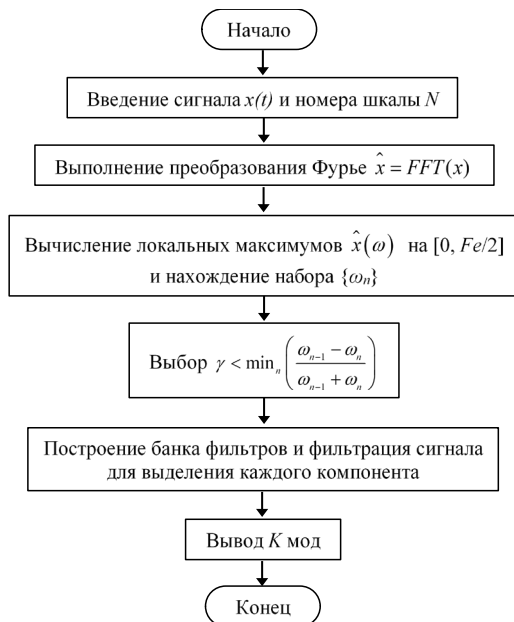


Рис. 3. Блок-схема алгоритма метода EWT

3. EWT-CGAN аугментация.

Обоснование выбора EWT и Conditional GAN. Объединение методов EWT и Conditional GAN обусловлено необходимостью решения ключевой проблемы генерации нестационарных сигналов промышленного оборудования – сохранения характерных особенностей переходных процессов, трендов и сезонных вариаций, которые традиционные методы типа timeGAN воспроизводят недостаточно точно.

Метод EWT выбран в качестве метода декомпозиции сигналов по нескольким причинам. Во-первых, в отличие от классического вейвлет-анализа или преобразования Фурье, EWT адаптируется к спектральным характеристикам конкретного сигнала, что критично для обработки нестационарных данных с изменяющимися во времени частотами. Во-вторых, EWT обеспечивает четкое разделение мод без явления «смешивания», характерных для других методов эмпирических декомпозиций. Это позволяет выделить физически значимые компоненты сигнала: высокочастотные шумы, «средние» частоты, связанные с изменениями режимов работы оборудования, и низкочастотные тренды. Наконец, EWT менее чувствителен к краевым эффектам по сравнению с другими адаптивными методами, что важно для работы с относительно короткими реализациями сигналов.

Использование Conditional GAN вместо стандартного timeGAN обусловлено тремя факторами. Первый – возможность управления процессом генерации через категориальные метки мод, что позволяет сохранять их индивидуальные характеристики. Второй – устойчивость Conditional GAN к «коллапсу мод», который часто наблюдается при генерации разнородных временных рядов. Третий фактор – гибкость: одна обученная модель может генерировать сигналы для разных режимов работы оборудования, изменяя входные метки.

Комбинация EWT и Conditional GAN создает синергетический эффект. EWT обеспечивает содержательную декомпозицию сигнала на физически интерпретируемые компоненты, а Conditional GAN точно воспроизводит распределение каждой моды. Такой подход особенно эффективен для сигналов с резкими переходами между режимами (например, пуск/останов оборудования), где традиционные методы генерации часто дают «размытые» или статистически несостоятельные результаты.

Описание архитектуры метода. Предложенная EWT-CGAN аугментация основана на комплексном применении эмпирического вейвлет преобразования (EWT) и генеративной состязательной нейронной сети с условными переменными (Conditional GAN).

Подобная связка позволяет корректно генерировать нестационарные сигналы средств измерений, близкие к реальным, в отличие от простого применения генеративных состязательных нейронных сетей. Архитектура *EWT-CGAN* аугментации представлена на рисунке 4.

Основой *EWT-CGAN* аугментации является генеративная состязательная сеть с условными переменными на входе. Однако исходный сигнал не сразу подается на дискриминатор, а предварительно декомпозируется на моды. Каждой моде присваивается категориальная метка (*label*). Таким образом на вход в дискриминатор (*D*) подаются одновременно и моды исходного сигнала и их обозначения в виде меток. Генератор (*G*) в этом случае учится синтезировать сигналы одновременно для всех мод исходного сигнала. Другими словами, в одной генеративной состязательной нейронной сети происходит обучение генерации одновременно нескольких разных процессов, соответствующих разным модам исходного сигнала. Последним этапом работы является сложение полученных синтезированных мод сигнала. При этом на стадии сложения мод есть возможность подкорректировать сигнал: сгладить низкочастотную составляющую, добавить/убрать тренд.

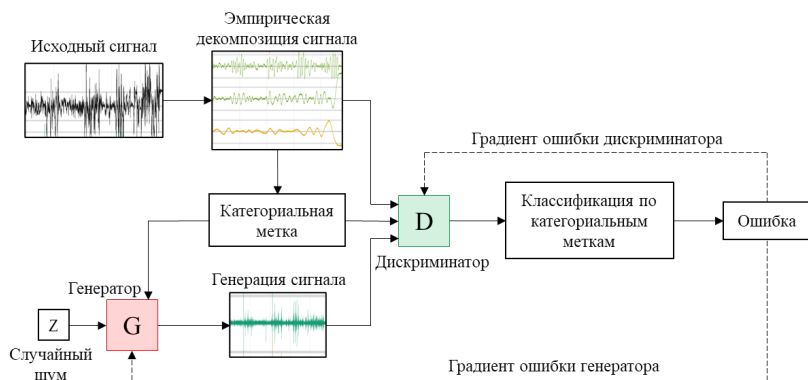


Рис. 4. Общий вид архитектуры EWT-CGAN аугментации

Настройки декомпозиции сигнала. Настройка процесса декомпозиции сигналов методом EWT в данном исследовании основывалась на оптимальном выборе количества мод, что является ключевым параметром для качественного анализа промышленных сигналов. Для сигналов от температурных датчиков установлено, что разложение на три характерные моды обеспечивает наиболее полное и физически обоснованное представление данных:

1. Высокочастотная мода – отражает шумовые составляющие сигнала.
2. Среднечастотная мода – соответствует рабочим режимам и переходным процессам.
3. Низкочастотная мода – описывает температурные тренды и долгосрочные изменения.

Такое разделение выбрано на основе анализа спектральной плотности мощности для 50 различных сигналов от температурных датчиков. Эксперименты показали, что увеличение числа мод свыше трех приводит к появлению искусственных компонент, а уменьшение – к потере значимых физических характеристик.

Для верификации качества декомпозиции использовались следующие критерии:

- энергетическая сбалансированность (разница энергий исходного и восстановленного сигналов не превышала 1%);
- минимальная взаимная корреляция между модами ($< 0,15$);
- физическая интерпретируемость каждой компоненты;
- сохранение ключевых особенностей исходного сигнала (переходных процессов, экстремумов, точек перегиба).

Такие параметры декомпозиции обеспечивают оптимальное представление сигналов для последующей обработки методом Conditional GAN, сохраняя при этом все существенные особенности исходных данных. Важно отметить, что предложенная конфигурация (три моды) показала свою эффективность именно для сигналов температурных датчиков промышленного оборудования, работающего в нестационарных режимах. Для сигналов другой природы (например, вибрационных или акустических) параметры декомпозиции могут потребовать корректировки.

Создание генератора и дискриминатора. Генератор представляет собой двухвходовую нейронную сеть, которая создает сигнал на основе массива случайных чисел и связанной с ним категориальной метки.

Алгоритм создания генератора:

- Создается массив случайного шума заданной размерностью.
- Категориальные метки преобразуются во входные векторы.
- Объединяются результаты из двух входов в один массив.
- Увеличивается частота дискретизации полученных массивов с помощью серии одномерных транспонированных сверточных слоев с нормализацией данных и слоями активации ReLU.

Количество транспонированных сверточных слоев и их настройки выбираются экспериментально на основе анализа длины входного сигнала и его частотных характеристик. Для исследуемого в работе процесса (1201 точка) выбрано 5 слоев. Для сигналов большей длины или с более сложными переходными процессами количество слоев может быть увеличено до 6–7.

В общем виде структуру генератора можно представить в виде последовательности транспонированных сверточных слоев, слоев нормализации и слоев активации (рисунок 5).

В качестве дискриминатора задается двухвходовая нейронная сеть, которая классифицирует «реальный» / «сгенерированный» сигнал и определяет их категориальные метки.

Алгоритм создания дискриминатора:

- На вход подается сигнал с генератора.
- Категориальные метки преобразуются в векторы.
- Результаты двух входов объединяются в один массив.
- Уменьшается размерность массивов, используя 1-D

сверточные слои с функций активации leakyReLU и масштабным коэффициентом 0,2. Коэффициент выбран как компромисс между устойчивостью к «затуханию» градиентов (при меньших значениях) и сохранением информации (при больших). Количество сверточных слоев и их настройки выбираются экспериментально.

В общем виде структуру дискриминатора можно представить в виде последовательности сверточных слоев и слоев активации (рисунок 5).

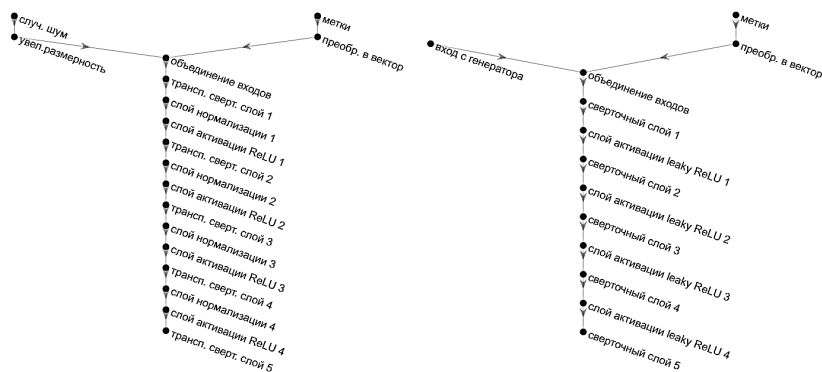


Рис. 5. Общая структура генератора (слева) и дискриминатора (справа)

Настройки процесса обучения нейросетевой модели. Ниже приведены рекомендации по настройке обучения модели:

- Для корректного обучения модели рекомендуется задавать количество эпох не менее 1000. Точное значение определяется по критерию сходимости метрик WS и PCC, когда дальнейшее обучение не улучшает их значения более чем на 1%.

- Размер пакета (Batch Size) выбирается исходя из размера исходного сигнала и вычислительных мощностей. Обычно принимается одно из следующих значений: 64, 128 или 256.

- Скорость обучения выбирается на основе grid-поиска в диапазоне [0,0001; 0,001] для обеспечения стабильного снижения потерь без колебаний.

- Рекомендуется процесс обучения проводить с использованием графического процесса (GPU) с целью сокращения времени вычисления.

- В качестве оптимизатора рекомендуется выбирать Adams, что обосновано адаптивностью скорости обучения и эффективностью при работе с зашумленными градиентами. Кроме того, оптимизатор Adams устойчив к различным начальным условиям благодаря адаптивному расчету моментов градиентов.

- Скорость затухания скользящих среднего градиента (gradientDecayFactor) в общем случае рекомендуется задавать равным 0,5, что обеспечивает баланс между учетом истории градиентов и быстрой адаптацией. Кроме того, слишком высокие значения ($> 0,9$) приводят к инерционности при смене режимов, а слишком низкие ($< 0,3$) вызывают излишние колебания процесса обучения.

- Скорость затухания скользящих квадратичного градиента (squaredGradientDecayFactor) в общем случае рекомендуется задавать равной 0,99 для длительного усреднения квадратов градиентов и стабилизации обучения при редких, но значительных выбросах. При наличии выраженных выбросов допускается уменьшение squaredGradientDecayFactor до 0,95.

4. Применение EWT-CGAN аугментации для реальных сигналов промышленного оборудования. EWT-CGAN аугментация апробирована на реальных данных средств измерений, представляющие собой сигналы с температурного датчика, установленного на агрегате с характерными нестационарными режимами работы.

Для аугментации использовано 495 временных реализаций процесса изменения нормированной температуры. Каждая реализация

состоит из 1201 отчета по времени. Фрагмент одной реализации показан на рисунке 6.

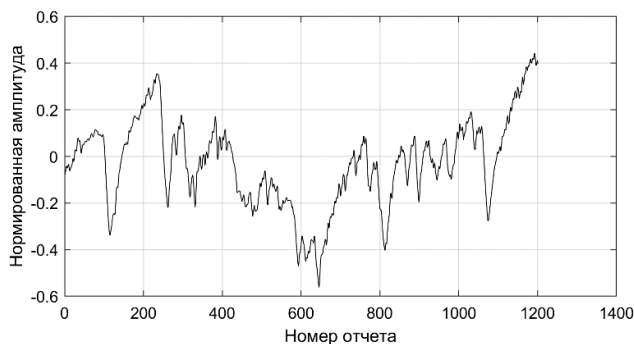


Рис. 6. Фрагмент временной реализации с температурного датчика

Выполнена аугментация данных температуры двумя методами: традиционным методом *timeGAN* и предложенным *EWT-CGAN*.

EWT-CGAN аугментация предполагает использование эмпирического вейвлет преобразования, для которого задались настройки, при которых каждая реализация декомпозировалась на три моды. В дальнейшем каждой моде присваивалась категориальная метка (1, 2 или 3, соответствующая номеру моды).

Структура генератора и структура дискриминатора выбиралась одинаковыми как для *timeGAN*, так и для *EWT-CGAN*. Гиперпараметры генератора и дискриминатора задавались также одинаковыми (таблица 2 и таблица 3).

Количество транспонированных сверточных слоев генератора (5 слоев, таблица 2) и размер ядер (5×1 , 10×1 и т.д.) выбраны экспериментально на основе анализа длины входного сигнала (1201 точка) и его частотных характеристик.

Количество сверточных слоев и размер ядер дискриминатора (5 слоев, таблица 3) выбраны на основе анализа качества генерации: добавление слоев сверх 5 не улучшало показатели Wasserstein и PCC, но увеличивало время обучения.

На основе выполненного grid-поиска величина скорости обучения составила 0,0002.

Проведено несколько запусков модели с разными начальными весами (инициализация *Xe* и *Glorot* – нормальное распределение). Результаты показали, что вариативность метрик (WS, PCC, RMSE) не превышает 5%, что подтверждает стабильность обучения.

Таблица 2. Гиперпараметры генератора

Операция	Размер ядра	Шаг	Число фильтров	Нормализация	Функция активации
Объединение входов	–	–	–	Нет	
Свертка	5×1	1	512	Да	ReLU
Свертка	10×1	4	256	Да	ReLU
Свертка	12×1	4	128	Да	ReLU
Свертка	5×1	4	64	Да	ReLU
Свертка	7×1	2	1	Нет	–

Таблица 3. Гиперпараметры дискриминатора

Операция	Размер ядра	Шаг	Число фильтров	Нормализация	Функция активации
Объединение входов	–	–	–	Нет	
Свертка	17×1	2	512	Нет	leakyReLU
Свертка	16×1	4	256	Нет	leakyReLU
Свертка	16×1	4	128	Нет	leakyReLU
Свертка	8×1	4	64	Нет	leakyReLU
Свертка	8×1	1	1	Нет	–

Характерный вид реализации температурного сигнала после аугментации разными методами показан на рисунке 7.

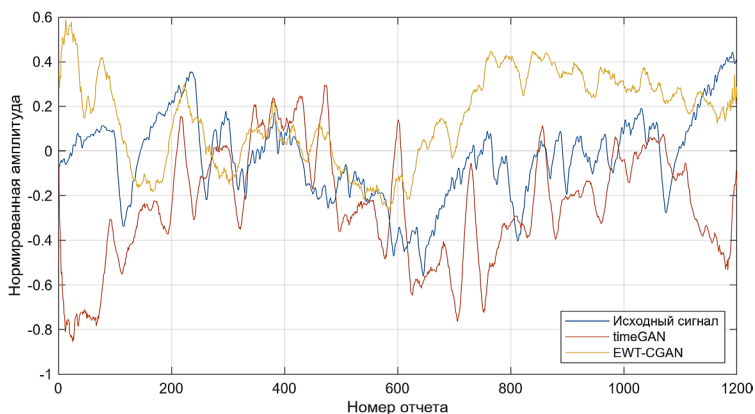


Рис. 7. Графики исходного и аугментированного разными методами сигнала температуры

Выполнена оценка качества аугментации сигнала с применение статистических метрик, описанных ранее. Вычисленные значения метрик для сигналов, полученных методами *timeGAN* и *EWT-CGAN* по отношению к исходному сигналу, сведены в таблицу 4.

Таблица 4. Метрики оценки качества аугментации

Метрики	MAE	MSE	RMSE	PRD	MRE	PCC	WS
<i>timeGAN</i>	0,281	0,132	0,364	1,898	0,218	0,133	0,212
<i>EWT-CGAN</i>	0,236	0,082	0,286	1,494	0,377	0,313	0,176
Сравнение	16%	38%	21%	21%	–	–	17%

Из таблицы 4 видно, что сигналы, сгенерированные методом *EWT-CGAN*, статистически более подобны исходному сигналу. Практически все метрики, в том числе наиболее значимая – расстояние Wasserstein, для метода *EWT-CGAN* на 20% ниже, чем для традиционного метода *timeGAN*. Коэффициент корреляции Пирсона (PCC) так же показал лучшую сходимость сигналов, сгенерированных методом *EWT-CGAN*, с исходным, в отличие от сигналов, сгенерированных методом *timeGAN*. Разница коэффициента достигает 57%. Однако рассчитанная средняя относительная ошибка (MRE) показала обратный эффект. Вероятно, это связано с особенностями вычислений коэффициента, в котором одним из основных показателей выступает сезонность процесса (п. 4 таблицы 1).

Другим методом оценки качества аугментации сигнала является непосредственное сравнение плотностей распределений исходного сигнала и сгенерированного. На рисунке 8 показаны сравнения плотностей распределений исходного сигнала и сгенерированного методами *EWT-CGAN* и *timeGAN* соответственно с рассчитанными значениями коэффициентов эксцесса и асимметрии.

Анализируя плотности распределений, важно отметить, что при генерации сигнала методом *timeGAN* плотность распределения сдвигается влево, что свидетельствует о преобладании более низких амплитудных значений. Для метода *EWT-CGAN* плотность распределения, наоборот, смещена вправо – в область с большими по величине амплитудами.

Анализ коэффициента эксцесса показал, что исходный сигнал соответствует нормальному закону распределения, а сгенерированные не соответствуют. При этом для сигнала, сгенерированного методом *timeGAN*, коэффициент эксцесса равен 2,43, а для метода *EWT-CGAN* – равен 2.

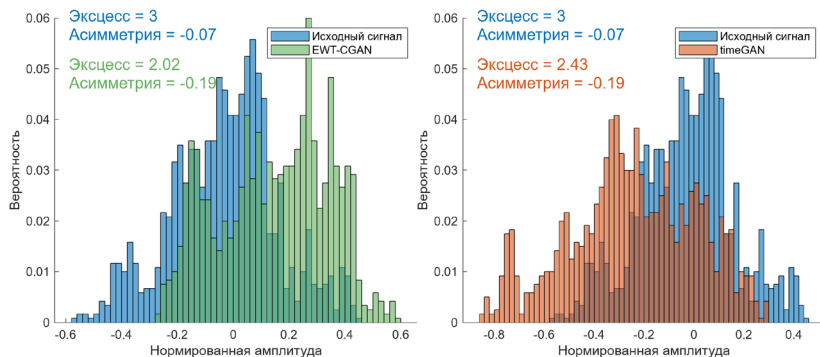


Рис. 8. Плотности распределений исходного сигнала и сгенерированного методом timeGAN (слева), исходного и сгенерированного методом EWT-CGAN (справа)

Аналогичная картина наблюдается при рассмотрении исходного сигнала и сгенерированного по двум методам в виде диаграммы размахов (рисунок 9).

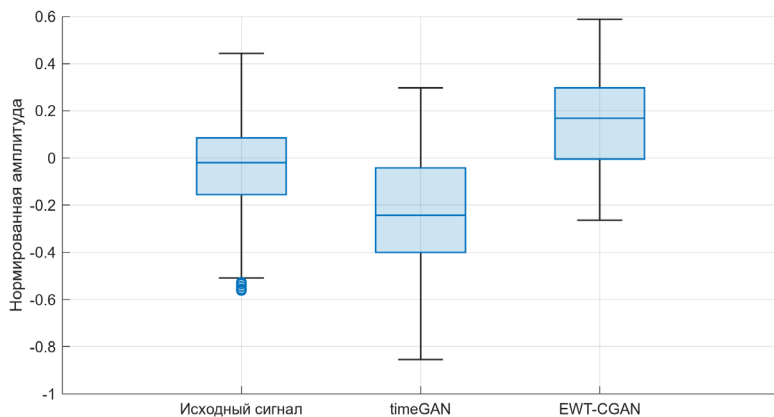


Рис. 9. Диаграмма размахов нормированных амплитуд для исходного и сгенерированных сигналов

Видно, что среднее значение для сгенерированных сигналов смещено относительно исходного, а общий вид плотности распределений для сгенерированных сигналов несколько отличается от исходного. Вероятно, такое расхождение связано с относительно небольшим по длине фрагментом реализации (1201 точка). Большие по размеру фрагменты данных требуют заметно больших

вычислительных ресурсов и не рассматривались в работе. Кроме того, важно отметить, что для аугментации с применением GAN (timeGAN и *EWT-CGAN*) необходимо обучать нейронную сеть не для одной реализации, а для набора реализаций одной и той же длины. В работе использовалось 495 реализаций по 1201 точек каждая. Другими словами, рассмотрена реализация длиной 594 495 точек, фрагментированная на 495 реализаций.

Оценка вычислительных ресурсов и масштабируемость.

Адаптивное разложение сигнала на моды методом EWT зависит от длины сигнала. Для реализации из 1201 точки время обработки составляет порядка 0,1–0,5 секунды на CPU (Intel Core i7). Для 495 реализаций общее время декомпозиции не превышает 5 минут, что демонстрирует низкие требования к ресурсам на этапе декомпозиции.

Основная вычислительная нагрузка приходится на обучение генератора и дискриминатора. В сравнении с классическим timeGAN, *EWT-CGAN* требует на 15–20% больше времени из-за обработки нескольких мод. Однако ключевое преимущество – возможность обучения одной модели для всех мод (благодаря категориальным меткам), что сокращает общие затраты по сравнению с обучением отдельных GAN для каждой моды. На GPU (NVIDIA Tesla V100) обучение модели для 495 реализаций (1201 точка каждая) занимает ~1,5 часа (1000 эпох), что сопоставимо с timeGAN, но обеспечивает более высокое качество генерации.

Для крупных производственных систем с тысячами датчиков возможна распределенная обработка путем параллельной декомпозиции сигналов на узлах CPU и обучения *EWT-CGAN* на кластерах GPU.

5. Заключение. Разработан метод *EWT-CGAN* аугментации временных рядов, основанный на комбинированном использовании эмпирического вейвлет-преобразования сигналов и генеративной состязательной нейронной сети. Предложенный метод отличается от существующих своей адаптивностью и способностью учитывать в сигналах тренды, выбросы и переходные процессы, что позволяет синтезировать сложные сигналы, близкие к реальным эксплуатационным условиям, без необходимости проведения дорогостоящих экспериментов или разработки математической модели объекта с виртуальными испытаниями. Предложенный метод *EWT-CGAN* аугментации данных обеспечивает формирование релевантного обучающего набора для нейросетевых моделей, используемых для оценки технического состояния промышленного оборудования, с меньшим уровнем ошибок.

Метод апробирован на реальных данных, полученных с температурного датчика. Количественная оценка эффективности метода проводилась с использованием статистических метрик. Результаты показали, что сигналы, сгенерированные методом *EWT-CGAN*, статистически более близки к исходным данным. Все метрики, включая ключевую – расстояние Wasserstein, для метода *EWT-CGAN* оказались в среднем на 20% ниже. Коэффициент корреляции Пирсона для метода *EWT-CGAN* более чем на 50% выше, что подтверждает высокую степень сходства сгенерированных сигналов с исходными. Однако при анализе плотностей распределений амплитуд наблюдаются некоторые различия в форме распределения и, соответственно, в коэффициенте эксцесса, как для метода *EWT-CGAN*, так и для метода *timeGAN*. Вероятно, это связано с относительно короткими фрагментами реализаций, что обусловлено ограниченными вычислительными ресурсами.

Литература

1. Musa N., Gital A.Y., Aljojo N., et al. A systematic review and Meta-data analysis on the applications of Deep Learning in Electrocardiogram. *J. Ambient Intell. Humaniz. Comput.* 2023. vol. 14. no. 7. pp. 9677–9750. DOI: 10.1007/s12652-022-03868-z.
2. Zhang D., Ma M., Xia L. A comprehensive review on GANs for time-series signals. *Neural Comput. Appl.* 2022. vol. 34. no. 5. pp. 3551–3571. DOI: 10.1007/s00521-022-06888-0.
3. Brophy E., Wang Z., She Q., Ward T. Generative Adversarial Networks in Time Series: A Systematic Literature Review. *ACM Comput. Surv.* 2023. vol. 55. no. 10. DOI: 10.1145/3559540.
4. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 2014. vol. 27. pp. 2672–2680.
5. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. 4th Int. Conf. Learn. Represent. (ICLR 2016). 2016. pp. 1–16.
6. Karras T., Aila T., Laine S., Lehtinen J. Progressive growing of GANs for improved quality, stability, and variation. 6th Int. Conf. Learn. Represent. (ICLR 2018). 2018. pp. 1–26.
7. Mirza M., Osindero S. Conditional Generative Adversarial Nets. *arXiv preprint arXiv:1411.1784*. 2014. pp. 1–7.
8. Mao X., Li Q., Xie H., Lau R.Y.K., Wang Z., Smolley S.P. Least Squares Generative Adversarial Networks. *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*. 2017. vol. 2017. pp. 2813–2821. DOI: 10.1109/ICCV.2017.304.
9. Arjovsky M., Chintala S., Bottou L. Wasserstein GAN. *Proceedings of the 34 th International Conference on Machine Learning*. 2017. vol. 70. pp. 214–223.
10. Fabbri C. Conditional Wasserstein Generative Adversarial Networks. 2017.
11. Akcay S., Atapour-Abarghouei A., Breckon T.P. GANomaly: Semi-supervised Anomaly Detection via Adversarial Training. *Lect. Notes Comput. Sci.* 2019. vol. 11363. pp. 622–637. DOI: 10.1007/978-3-030-20893-6_39.
12. Zenati H., Foo C.-S., Lecouat B., Manek G., Chandrasekhar V.R. Efficient GAN-Based Anomaly Detection. 2018.

13. Maaten L. van der Hinton G. Visualizing Data using t-SNE. *Journal of Machine Learning Research*. 2008. vol. 9. pp. 2579–2605.
14. Bryant F.B., Yarnold P.R. Principal-components analysis and exploratory and confirmatory factor analysis. *Reading and understanding multivariate statistics*. Washington, DC, US: American Psychological Association, 1995. pp. 99–136.
15. Yoon J., Jarrett D., van der Schaar M. Time-series generative adversarial networks. *Adv. Neural Inf. Process. Syst.* 2019. vol. 32. pp. 1–11.
16. Salimans T., Goodfellow I., Zaremba W., Cheung V., Randford A., Chen X. Improved techniques for training GANs. *30th Conference on Neural Information Processing Systems (NIPS 2016)*. Barcelona, 2016.
17. Heusel M., Ramsauer H., Unterthiner T., Nessler B., Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. *Adv. Neural Inf. Process. Syst.* 2017. vol. 2017. pp. 6626–6637.
18. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* 2004. vol. 13. no. 4. pp. 600–612. DOI: 10.1109/TIP.2003.819861.
19. Huang N.E., Shen Z., Long S.R., Wu M.C., Shih H.H., Zheng Q., Yen N.-C., Tung C.C., Liu H.H. The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 1998. vol. 454. pp. 903–995. DOI: 10.1098/rspa.1998.0193.
20. Dragomiretskiy K., Zosso D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* 2014. vol. 62. no. 3. pp. 531–544. DOI: 10.1109/TSP.2013.2288675.
21. Gilles J. Empirical wavelet transform. *IEEE Trans. Signal Process.* 2013. vol. 61. no. 16. pp. 3999–4010. DOI: 10.1109/TSP.2013.2265222.
22. Singh P., Joshi S.D., Patney R.K., Saha K. The Fourier decomposition method for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 2017. vol. 473. no. 2199. DOI: 10.1098/rspa.2016.0871.
23. Gilles J. Empirical Wavelet Transform. *IEEE Trans. Signal Process.* 2013. vol. 61. no. 16. pp. 3999–4010. DOI: 10.1109/TSP.2013.2265222.

Ерпалов Алексей Викторович — канд. техн. наук, заместитель директора, центр виброиспытаний и мониторинга состояния конструкций, НИИ опытного машиностроения, ФГАУО ВО «ЮУрГУ (НИУ)». Область научных интересов: вибропрочность конструкций, применение нейросетевых технологий в задачах мониторинга технического состояния, аугментация данных средств измерений. Число научных публикаций — 42. erpalovav@susu.ru; проспект Ленина, 76, 454080, Челябинск, Россия; р.т.: +7(351)267-9900.

Синицин Владимир Владимирович — канд. техн. наук, заместитель заведующего, НИИ технической самодиагностики и самоконтроля приборов и систем, ФГАУО ВО «ЮУрГУ (НИУ)». Область научных интересов: техническая диагностика, прогнозирование остаточного ресурса промышленного оборудования, верификация математических моделей. Число научных публикаций — 33. sinitscinvv@susu.ru; проспект Ленина, 76, 454080, Челябинск, Россия; р.т.: +7(351)267-9900.

Шестаков Александр Леонидович — д-р техн. наук, профессор, председатель совета ректоров вузов урфо, заслуженный работник высшей школы, президент, ФГАУО ВО «ЮУрГУ (НИУ)». Область научных интересов: динамика измерений, самодиагностика средств измерений, самодиагностика промышленных систем. Число научных публикаций — 235. president@susu.ru; проспект Ленина, 76, 454080, Челябинск, Россия; р.т.: +7(351)267-9900.

Поддержка исследований. Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации (государственное задание на выполнение фундаментальных научных исследований № FENU-2023-0010 (2023010ГЗ)).

A. ERPALOV, V. SINITSIN, A. SHESTAKOV
**EWT-CGAN DATA AUGMENTATION FOR MEASUREMENT
SYSTEMS**

Erpalov A., Sinitsin V., Shestakov A. **EWT-CGAN Data Augmentation for Measurement Systems.**

Abstract. The article presents a new data augmentation method for measurement systems, designed for industrial equipment condition monitoring tasks. The relevance of the study stems from the significant limitations of traditional synthetic data generation methods, which fail to adequately reproduce complex non-stationary signals with characteristic transient processes, trends, and seasonal variations observed in real industrial environments. The proposed method integrates two advanced techniques: empirical wavelet transform (EWT) and conditional generative adversarial networks (Conditional GAN). The method is implemented in three stages: (1) adaptive decomposition of raw signals into modes using EWT, (2) mode categorization with label assignment, and (3) synthetic data generation using Conditional GAN. A set of statistical metrics was used to comprehensively assess the quality of synthesized signals, including Wasserstein distance (WS), Pearson correlation coefficient (PCC), and root mean square error (RMSE). Experimental studies were conducted on real-world temperature sensor data obtained under non-stationary industrial equipment conditions. The results demonstrate a significant advantage of the proposed method over the traditional TimeGAN approach: a 17% reduction in Wasserstein distance, a 57% increase in Pearson correlation coefficient, and a 21% decrease in RMSE. These findings confirm the method's effectiveness in reproducing key characteristics of the original signals. The developed method enables the creation of synthetic datasets required for training modern neural network models in industrial equipment diagnostics. Its practical application significantly reduces the costs associated with experimental data collection while ensuring high-quality synthesized signals, as validated by statistical metrics.

Keywords: equipment diagnostics, sensor signals, data augmentation, synthetic data, empirical wavelet transform, generative adversarial networks.

References

1. Musa N., Gital A.Y., Aljojo N., et al. A systematic review and Meta-data analysis on the applications of Deep Learning in Electrocardiogram. *J. Ambient Intell. Humaniz. Comput.* 2023. vol. 14. no. 7. pp. 9677–9750. DOI: 10.1007/s12652-022-03868-z.
2. Zhang D., Ma M., Xia L. A comprehensive review on GANs for time-series signals. *Neural Comput. Appl.* 2022. vol. 34. no. 5. pp. 3551–3571. DOI: 10.1007/s00521-022-06888-0.
3. Brophy E., Wang Z., She Q., Ward T. Generative Adversarial Networks in Time Series: A Systematic Literature Review. *ACM Comput. Surv.* 2023. vol. 55. no. 10. DOI: 10.1145/3559540.
4. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 2014. vol. 27. pp. 2672–2680.
5. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. 4th Int. Conf. Learn. Represent. (ICLR 2016). 2016. pp. 1–16.

6. Karras T., Aila T., Laine S., Lehtinen J. Progressive growing of GANs for improved quality, stability, and variation. 6th Int. Conf. Learn. Represent. (ICLR 2018). 2018. pp. 1–26.
7. Mirza M., Osindero S. Conditional Generative Adversarial Nets. arXiv preprint arXiv:1411.1784. 2014. pp. 1–7.
8. Mao X., Li Q., Xie H., Lau R.Y.K., Wang Z., Smolley S.P. Least Squares Generative Adversarial Networks. Proc. IEEE Int. Conf. Comput. Vis. (ICCV). 2017. vol. 2017. pp. 2813–2821. DOI: 10.1109/ICCV.2017.304.
9. Arjovsky M., Chintala S., Bottou L. Wasserstein GAN. Proceedings of the 34 th International Conference on Machine Learning. 2017. vol. 70. pp. 214–223.
10. Fabbri C. Conditional Wasserstein Generative Adversarial Networks. 2017.
11. Akcay S., Atapour-Abarghouei A., Breckon T.P. GANomaly: Semi-supervised Anomaly Detection via Adversarial Training. Lect. Notes Comput. Sci. 2019. vol. 11363. pp. 622–637. DOI: 10.1007/978-3-030-20893-6_39.
12. Zenati H., Foo C.-S., Lecouat B., Manek G., Chandrasekhar V.R. Efficient GAN-Based Anomaly Detection. 2018.
13. Maaten L. van der Hinton G. Visualizing Data using t-SNE. Journal of Machine Learning Research. 2008. vol. 9. pp. 2579–2605.
14. Bryant F.B., Yarnold P.R. Principal-components analysis and exploratory and confirmatory factor analysis. Reading and understanding multivariate statistics. Washington, DC, US: American Psychological Association, 1995. pp. 99–136.
15. Yoon J., Jarrett D., van der Schaar M. Time-series generative adversarial networks. Adv. Neural Inf. Process. Syst. 2019. vol. 32. pp. 1–11.
16. Salimans T., Goodfellow I., Zaremba W., Cheung V., Randford A., Chen X. Improved techniques for training GANs. 30th Conference on Neural Information Processing Systems (NIPS 2016). Barcelona, 2016.
17. Heusel M., Ramsauer H., Unterthiner T., Nessler B., Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. Adv. Neural Inf. Process. Syst. 2017. vol. 2017. pp. 6626–6637.
18. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 2004. vol. 13. no. 4. pp. 600–612. DOI: 10.1109/TIP.2003.819861.
19. Huang N.E., Shen Z., Long S.R., Wu M.C., Shih H.H., Zheng Q., Yen N.-C., Tung C.C., Liu H.H. The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. A Math. Phys. Eng. Sci. 1998. vol. 454. pp. 903–995. DOI: 10.1098/rspa.1998.0193.
20. Dragomiretskiy K., Zosso D. Variational Mode Decomposition. IEEE Trans. Signal Process. 2014. vol. 62. no. 3. pp. 531–544. DOI: 10.1109/TSP.2013.2288675.
21. Gilles J. Empirical wavelet transform. IEEE Trans. Signal Process. 2013. vol. 61. no. 16. pp. 3999–4010. DOI: 10.1109/TSP.2013.2265222.
22. Singh P., Joshi S.D., Patney R.K., Saha K. The Fourier decomposition method for nonlinear and non-stationary time series analysis. Proc. R. Soc. A Math. Phys. Eng. Sci. 2017. vol. 473. no. 2199. DOI: 10.1098/rspa.2016.0871.
23. Gilles J. Empirical Wavelet Transform. IEEE Trans. Signal Process. 2013. vol. 61. no. 16. pp. 3999–4010. DOI: 10.1109/TSP.2013.2265222.

Erpalov Aleksey — Ph.D., Deputy director, Center for vibration testing and structural condition monitoring, research institute of experimental mechanical engineering, South Ural State University (National Research University). Research interests: structural vibration resistance, application of neural network technologies in technical condition monitoring tasks, measurement data augmentation. The number of publications — 42. erpalovav@susu.ru; 76, Lenin Ave., 454080, Chelyabinsk, Russia; office phone: +7(351)267-9900.

Sinitstin Vladimir — Ph.D., Deputy head, Research laboratory of technical self-diagnostics and self-monitoring of instruments and systems, South Ural State University (National Research University). Research interests: technical diagnostics, residual life prediction of industrial equipment, verification of mathematical models. The number of publications — 33. sinitstinvv@susu.ru; 76, Lenin Ave., 454080, Chelyabinsk, Russia; office phone: +7(351)267-9900.

Shestakov Aleksander — Ph.D., Dr.Sci., Professor, Chairman of the council of rectors of universities of the ural federal district, honored worker of higher education, President, South Ural State University (National Research University). Research interests: measurement dynamics, self-diagnostics of measuring instruments, self-diagnostics of industrial systems. The number of publications — 235. president@susu.ru; 76, Lenin Ave., 454080, Chelyabinsk, Russia; office phone: +7(351)267-9900.

Acknowledgements. This work was financially supported by Ministry of Science and Higher Education of the Russian Federation (FENU-2023-0010).

С.В. ДВОРНИКОВ, А.А. МАСЛОВА, Д.В. ВАСИЛЬЕВА
**ВЕРОЯТНОСТНАЯ ОЦЕНКА ПОМЕХОЗАЩИЩЕННОСТИ
СОСТАВНЫХ РАДИОЛИНИЙ ПЕРЕДАЧИ ВИДЕО
В УСЛОВИЯХ ВЗАИМНЫХ ПОМЕХ**

Дворников С.В., Маслова А.А., Васильева Д.В. **Вероятностная оценка помехозащищенности составных радиолиний передачи видео в условиях взаимных помех.**

Аннотация. Переход к автоматизированным системам регулирования на неохраемых железнодорожных переездах обусловил необходимость решения вопросов их безопасности. Наиболее рациональным решением указанной проблемы является применение систем видеонаблюдения, обеспечивающих трансляцию изображений, как на железнодорожные станции, в зоне ответственности которых находятся железнодорожные переезды, так и на локомотивы подвижных составов. С этой целью активно используют системы передачи информации, организованные на основе сетей широкополосного доступа и LTE. Но поскольку работа таких сетей организуется вдоль железнодорожных путей, в том числе в условиях активного применения различных излучающих средств, то для них характерно является нарушение условий электромагнитной совместимости, приводящее к сбою работы в результате непреднамеренного блокирования отдельных каналов. Поэтому анализ условий, при которых наступает сбой работы сети передачи видео, а также прогнозный расчет параметров радиолиний, при которых обеспечивается заданный уровень устойчивости работы сети, является актуальными. Рассмотрены технологии и особенности работы сетей стандарта LTE. Обоснованы показатели и критерии оценки функционирования линий передачи видео в рамках технических возможностей стандарта. Проведена математическая постановка задачи исследования. Определены исходные данные для разработки аналитической модели вероятностной оценки функционирования сети передачи видео. Разработан аналитический аппарат расчета вероятности блокирования каналов, учитывающий взаимную интенсивность использования частотного трафика конфликтующими средствами. Проанализированы требования ГОСТ Р 53111-2008, определяющего условия, при которых обеспечивается устойчивость функционирования сети связи общего пользования. Получено выражение вероятностной оценки, характеризующее вероятность нарушения (блокирования) работы сети, определяемой как канальными шумами и замираниями, так и помехами непреднамеренного характера от сторонних источников радиоизлучений при нарушении условий их электромагнитной совместимости. Представлены результаты аналитического моделирования, раскрывающие условия, при которых обеспечивается успешное функционирование сети передачи результатов видеонаблюдения. Обосновано, что работа сетей передачи видео в условиях взаимных помех при нарушении требований по электромагнитной совместимости, более чувствительна к изменению отношения значений интенсивностей потоков сети и источников сторонних излучений, работающих в режиме с программной перестройкой рабочей частоты, чем к энергетическим соотношениям полезного и мешающего сигналов.

Ключевые слова: взаимные помехи, системы передачи видео, сети LTE, вероятность блокирования работы сети, помехозащищенность радиолиний.

1. Введение. Переход к автоматизированным системам регулирования на неохраемых железнодорожных переездах

обусловил необходимость решения вопросов их безопасности [1, 2]. В таких условиях наиболее рациональным решением видится применение систем видеонаблюдения с трансляцией изображений [3], как на железнодорожные станции, в зоне ответственности которых находятся железнодорожные переезды, так и на локомотивы подвижных составов. С этой целью активно используют системы передачи информации, организованные на основе сетей широкополосного доступа и LTE (англ.: *Long-Term Evolution*) [4, 5].

Учитывая уровень ответственности, возлагаемый на охранные системы видеонаблюдения, к ним предъявляются достаточно жесткие требования по своевременности и достоверности предоставляемой информации. Но поскольку работа таких сетей организуется вдоль железнодорожных путей, над которыми протянуты высоковольтные провода, то в таких условиях неизбежно возникнут непреднамеренные помехи, обусловленные сложной электромагнитной обстановкой, которая усугубится наличием подвижного состава, как правило, состоящего из металлических вагонов.

Таким образом, работа сетей передачи информации от систем видеонаблюдения, будет проходить в сложной сигнально-помеховой обстановке, характеризующейся наличием, как взаимных помех [6], так и помех, обусловленных спецификой распространения радиоволн диапазона частот работы сетей широкополосного доступа и LTE [7]. В таких условиях велика вероятность того, что назначенный канал может оказаться непригодным для работы, поэтому прагматичным решением видится использования режима с программной перестройкой рабочей частоты (англ.: *Frequency-Hopping Spread Spectrum* – FHSS) [8], аналогично тому, как это реализовано в сетях GSM (англ.: *Global System for Mobile Communications*). Оценке обоснованности такого выбора посвящена настоящая работа. В статье рассмотрены условия, при которых наступает сбой работы сети передачи видео, а также рассчитаны параметры радиолиний, при которых обеспечивается заданный уровень устойчивости работы сети.

При этом в качестве исследуемой конфигурации сетей передачи видео рассмотрена система, состоящая из камеры установленной на переезде, подключенной к видеотерминалу, расположенному в кабине локомотива посредством составной радиолинии.

2. Общая постановка задачи и обоснование показателей оценки эффективности. Удобство выбора сетей стандарта LTE для организации передачи видеоинформации, в первую очередь, обусловлено их доступностью для применения, а также наличием широкого ассортимента соответствующего оборудования. Поскольку в

основе большинства оборудования LTE лежит SDR (англ.: *Software-Defined Radio*) платформа [9, 10], то открывается возможность для корпоративных сетей осуществлять индивидуальные настройки, в частности, организовывать режим FHSS [11].

В указанном режиме за счет псевдослучайной перестройки рабочей частоты (ППРЧ) минимизируется вероятность долговременного выбора непригодного канала, что позволяет повысить общую доступность частотного ресурса для большего числа абонентов, по отношению к режиму жесткого закрепления каналов. Если при этом обеспечить скорость перестройки FHSS такой, что при совместном выборе рабочей частоты различными абонентами не будет происходить взаимное нарушение их работы, то такое решение можно считать обоснованным [12].

Следует отметить, что в США разрешено использование систем FHSS в нелицензионных диапазонах частот 902–928 МГц, 2400–2483,5 МГц и 5725–5850 МГц [13]. При этом в системах FHSS допускается применение передач с излучаемой мощностью в 1 Ватт, в то время как для систем без расширенного спектра, значение мощности излучения ограничено 1 мВт.

Прагматичность реализации режима FHSS определяется широкими возможностями по организации одновременной работы большого числа абонентов в ограниченном частотном диапазоне. Так, например, в технологии eXRS (англ.: *eXtreme Radio Service*) при реализации режима FHSS используется 50 рабочих частот из 700 возможных [14, 15]. Но даже такая организация позволяет формировать частотный план на основе порядка десяти миллиардов частотно-временных матриц, определяющих сеансовый алгоритм работы абонентов в режиме FHSS. Указанный выбор минимизирует вероятность случайного совпадения частотного плана для двух групп независимых пользователей, что открывает возможность одновременной работы большого числа сетей в ограниченном диапазоне частот [16]. Как уже отмечалось, важным аспектом реализации режима FHSS является возможность минимизации потери информации при наличии каналов с многолучевым распространением в разрешенном списке рабочих частот.

Данный подход успешно использован в стандарте Bluetooth (адаптации на английский язык датского слова «Blåtand»), (англ.: *Wireless Personal Area Network – WPAN*) [17]. В указанном стандарте диапазон частот разбивается на 79 рабочих каналов, перестройка в пределах которых осуществляется со скоростью 1600 раз в секунду. При этом на длительности такого скачка передается один пакет

данных. Кроме того, режим FHSS реализован и в сетях стандарта GSM, где он определен как *slow frequency hopping* [18]. В указанном режиме связь между мобильным устройством и базовой станцией осуществляется пакетами данных в пределах выделенного временного интервала. Каждый пакет данных имеет длительность около 577 мкс. Тогда при длительности кадра $T_{\text{кад}} = 4,615$ мс, скорость перестройки рабочей частоты для стандарта GSM составляет $V_{\text{FHSS}} = 1 / T_{\text{кад}} = 217$ скачков в секунду (1/с).

Вместе с тем, анализ работ [11, 12, 16, 19] показал, что даже применение режима FHSS, полностью не исключает проблем с обеспечением электромагнитной совместимости (ЭМС) при совместной работе радиоэлектронных средств (РЭС) различных сетей, совместно использующих один и тот же диапазон рабочих частот.

Поэтому вопросы оценки эффективности режима FHSS при организации работы сетей LTE при передаче изображений с систем видеонаблюдения, установленных на не охраняемых железнодорожных переездах, требует дополнительного исследования.

При этом следует отметить, что LTE использует режим FHSS только для восходящей линии связи. Учитывая, что в данном стандарте минимальная длительность информационной посылки определяется длительностью слота, которая для стандарта LTE составляет 0,5 мс, то скорость перестройки выбора рабочих частот будет составлять $V_{\text{FHSS}} = 2000$ 1/с [20].

В основу оценки эффективности функционирования системы связи стандарта LTE, используемую для передачи изображения от камер видеонаблюдения, были положены общие положения теории систем [4]. При этом в качестве показателей определены достоверность, безопасность и своевременность передачи данных видеоизображения [21]. При этом в качестве показателя по своевременности определен временной интервал необходимый для передачи сообщений заданного объема [22], с учетом времени реакции машиниста и тормозного пути подвижного состава. Конкретное значение зависит от категории скорости подвижного состава [2].

Поскольку для формирования видеоизображений используются цифровые видеокамеры, то для оценки достоверности связи целесообразно определить показатели, используемые при оценке качества дискретных видов передач. Например, такой показатель как вероятность битовой ошибки P_b , или вероятность ошибки приема элемента сигнала $P_{\text{ош}}$ или битовой ошибки [23]. В этом случае критерием достоверности следует рассматривать величину допустимой вероятности ошибки $P_{\text{ош доп}}$ с учетом вида модуляции и

используемых мер по обеспечению заданной помехоустойчивости приема (помехоустойчивое кодирование, различные методы разнесенного приема, перемежение и т.д.) [11].

Требования по безопасности связи, как правило, определяются на уровне вторичных сетей. На практике они могут ограничиваться процедурами технического маскирования и разграниченного доступа. На данном этапе их целесообразно вывести в рамки ограничений.

С учетом введенных допущений и ограничений, определим способность сети передачи видеоизображения обеспечивать выполнение предъявляемых к ней требований по достоверности и своевременности, как ее эффективность Q , которую предлагается оценивать по следующим показателям:

- достоверности передачи с вероятностью ошибки не выше допустимой $P(P_{\text{ош}} \leq P_{\text{ош доп}})$;
- и своевременности передачи, т.е. вероятности того, что время, затраченное на передачу видеоизображения, не превысит допустимого значения $P(T_{\text{пер}} \leq T_{\text{пер доп}})$ при условии выполнения требований по достоверности.

Поскольку рассматриваемые значения вероятностей $P(P_{\text{ош}} \leq P_{\text{ош доп}})$ и $P(T_{\text{пер}} \leq T_{\text{пер доп}})$ определяются характеристиками радиоканала, то и эффективность такой сети также предлагается рассматривать с позиций выполнения указанных требований.

Тогда математическую задачу сформулируем следующей редакцией

$$Q \rightarrow \min_{T_{\text{пер}} \leq T_{\text{пер доп}}} \max_{P_{\text{ош}} \leq P_{\text{ош доп}}} Q(P(T_{\text{пер}} \leq T_{\text{пер доп}}), P(P_{\text{ош}} \leq P_{\text{ош доп}})). \quad (1)$$

Учитывая, что теоретический предел дальности связи для скорости в 1 Мбит/с для диапазона 2600 МГц составляет 3,2 км, то для обеспечения требований по своевременности и достоверности передачи видеоизображения в сети LTE, необходимо создавать составную линию радиосвязи.

Поэтому, в интересах решения задачи необходимо разработать модель функционирования сети в режиме FHSS, состоящей из составных линий радиосвязи, с учетом возможных замираний в канале, обусловленных многолучевым характером распространения радиоволн и взаимных помех, обусловленных совпадением выбора рабочих частот при реализации режима FHSS сторонних РЭС.

В качестве основных ограничений будем полагать следующее:

1. Разрабатываемая модель не связана со спецификой сжатия видеоданных.

2. В модели не рассматриваются форматы конкретных кодеков.

3. Архитектура составной линии радиосвязи стандарта LTE на железнодорожном переезде. Для реализации составной линии радиосвязи, спецификация стандарта LTE предлагает ряд вариантов.

В первую очередь, следует обратить внимание на то, что, в части радиодоступа (подсистема E-UTRAN, англ.: *Evolved Universal Terrestrial Access Network*), сеть беспроводной связи стандарта LTE предполагает однородную и неоднородную архитектуру.

Основу однородной архитектуры сети составляют множество макросот, обладающих одинаковым радиусом покрытия территорий и имеющих одинаковую мощность базовых станций. Такая сеть, как правило, используется для организации связи в крупных населенных пунктах. В свою очередь, неоднородная архитектура сети макросот представляет собой совокупность маломощных радиостанций, выполняющих либо функции ретрансляционного оборудования, либо самостоятельной базовой станции малого, по сравнению с покрытием макросоты, радиуса действия.

Реализация неоднородной архитектуры предусматривает выбор оптимального с точки зрения требований, предъявляемых к сети, и условий ее функционирования элементов. Такими элементами являются фемтосоты (англ.: Home eNB), транзитные узлы (англ.: *Relay Node – RN*) или удаленные радиоголовки (англ.: *Remote Radio Head – RRH*). Каждый из этих элементов предполагает использование различных протоколов взаимодействия с ядром пакетной сети и пользовательским оборудованием. Применение фемтосот характерно для покрытия зон доступа внутри зданий с большим числом потенциальных абонентов. Они применяются с целью снижения нагрузки на макросоту. Наиболее интересными, в контексте данного исследования, являются транзитные узлы и удаленные радиоголовки, различие между которыми отражено в таблице 1.

Таблица 1. Характеристики элементов неоднородной архитектуры E-UTRAN

Параметр	Транзитный узел	Удаленная радиоголовка
Тип присоединения к макросоте	Радиоинтерфейс	Волоконно-оптическая сеть связи
Отображение с точки зрения макросоты	Как пользовательское оборудование	Как составная часть базовой станции
Архитектура управления сетью	Децентрализованное	Централизованное

Ключевой особенностью железнодорожного транспорта является его территориальная протяженность, вдоль которой на большей части, находится малонаселенная или даже не заселенная местность, не имеющая инфраструктуры, позволяющей обеспечить непрерывное техническое обслуживание оборудования и прокладку множества кабелей связи.

Именно большие расстояния и слабо развитая инфраструктура, а также участки местности с различными климатическими условиями настаивают на выбор беспроводной архитектуры сети передачи данных и голосового трафика. С учетом указанных обстоятельств, на железнодорожных переездах, для организации радиолинии предлагается использовать транзитные узлы.

Взаимодействие с транзитным узлом, в общем виде, представляется как первичная обработка пакета данных от пользовательского оборудования транзитным узлом (транзитный узел отображается как базовая станция), и последующая передача его на базовую сеть для обработки (транзитный узел отображается как пользовательское оборудование). Такое техническое решение посредством радиодоступа абонентских терминалов к сети LTE предложено в патенте RU(11) № 2 496 263, опубликованного 20.10.2013. Бюл. № 29 МПК H04W 36/02 (2009.01).

В указанном техническом решении для организации передачи видеотрансляции в кабину локомотива подвижного состава введен мобильный терминал. Структура такой сети радиопередачи с применением транзитных узлов отражена на рисунке 1.

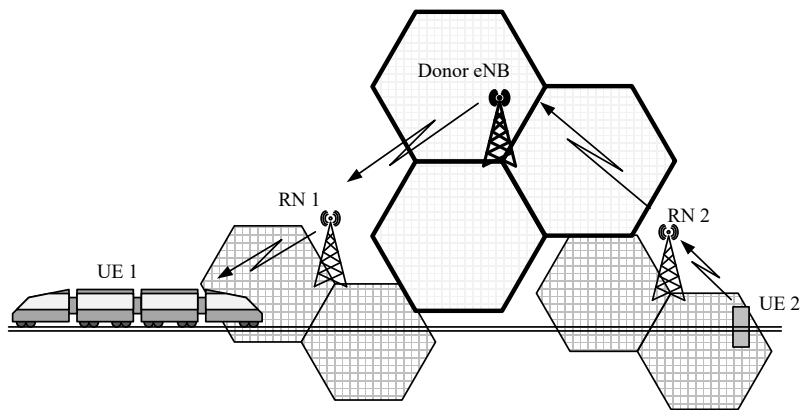


Рис. 1. Реализация сети передачи видеотрансляции в кабину локомотива подвижного состава с железнодорожного переезда посредством транзитного узла

На рисунке 1 условными обозначениями представлены следующие элементы: UE 1 – абонентский терминал, представляющий собой приемо-передатчик стандарта LTE размещенный в кабине локомотива подвижного состава; UE 2 абонентский терминал, представляющий собой приемо-передатчик стандарта LTE размещенный на неохраняемом железнодорожном переезде, получающий данные от видеокамер; RN 1, 2 – транзитные узлы, в зоне действия которых находятся подвижной состав и железнодорожный переезд соответственно; Donor eNB – базовая станция с функцией взаимодействия с транзитными узлами (донорская).

Еще одним вариантом построения сети беспроводной передачи видеоданных с железнодорожного переезда на монитор, размещенный в кабине машиниста локомотива подвижного состава, является использование отдельных элементов ретрансляторов LTE.

Проведенный анализ показал, что как отечественная, так и зарубежная научная общественность не уделяет этому варианту такого обширного внимания, как вышеупомянутым видам архитектуры неоднородной сети LTE. Однако это направление является не менее перспективным для его внедрения на железнодорожном транспорте.

К основным причинам недостаточного внимания к этому направлению следует отнести то, что архитектуре сети E-UTRAN с применением транзитных узлов присущ недостаток, связанный с необходимостью использования специализированной базовой станции, в которую должна быть включена функция взаимодействия с транзитными узлами. Следовательно, для участков железной дороги уже использующих сети 4G LTE, построенных без учета указанной возможности, дополнительное развертывание транзитных узлов не имеет смысла.

Таким образом, для обеспечения дополнительного видеоконтроля критически важных объектов железной дороги, расположенных на ее неохраняемых участках, необходим поиск альтернативных путей решения проблемы, обусловленной недостаточной дальностью зон обслуживания типовых базовых станций стандарта 4G LTE.

Широкое применение ретрансляционные узлы (англ.: *Power Amplifier* – PA) нашли в сетях коммерческой связи, как простое решение, обеспечивающее доступ в сеть Интернет на объектах, расположенных на значительном удалении от базовых станций сотовых операторов. При этом данное решение не требует специальных функций, в отличие от вышеописанного варианта, так

как логическое взаимодействие между ретрансляторами и базовыми станциями сведено к минимуму.

Суть ретрансляции заключается в приеме ослабленного сигнала от базовой станции, усилении его по мощности и передаче либо к пользовательскому терминалу, либо к следующему ретранслятору в цепи. Усиление по мощности сигнала, с одной стороны, повышает вероятность гарантированной доставки данных, но, с другой, вместе с сигналом усиливается и шум. Эта проблема дополнительно актуализирует и настоящее исследование.

Схема с ретрансляторами LTE в системе видеонаблюдения на железнодорожном переезде представлена на рисунке 2.

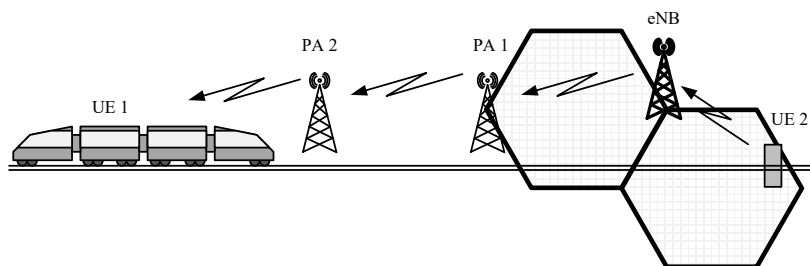


Рис. 2. Реализация сети передачи видеотрансляции в кабину локомотива подвижного состава с железнодорожного переезда посредством ретрансляторов

На рисунке 2 условными обозначениями представлены следующие элементы: PA 1, PA 2, – последовательная цепь ретрансляторов; eNB – базовая станция.

В случае, если внешние воздействия на канал связи будут создавать повышенные шумы, то может быть предложена комбинированная схема, с транзитным узлом, как одним из звеньев в цепи.

Таким образом, зная сигнально-помеховую обстановку на железнодорожном переезде, у проектировщиков есть несколько альтернативных путей построения сети беспроводной связи.

Первая схема актуальна для вновь развертываемой сети LTE. В этом случае, заранее выбирается тип базовой станции, в зависимости от предполагаемой архитектуры сети.

Вторая схема позволяет модернизировать уже существующую сеть связи, и не требует капитального перестроения существующей архитектуры сети.

Как было замечено ранее, помимо вышеописанных систем, спецификация стандарта предполагает еще один вариант построения сети – посредством удаленных радиоголовок.

Работы по этой тематике, как правило, ориентированы на архитектуру сетей пятого поколения. Это объясняется тем, что, несмотря на повышенную емкость сети 5G, ее базовые станции не обеспечивают и половины дальности связи, которая предусмотрена в сетях на базе стандарта 4G. Очевидно, что потенциал системы 5G и системы LTE-Advanced-PRO может быть реализован именно по такой структуре.

Как уже отмечалось, для железной дороги на протяженных участках и перегонах, в сложных погодных условиях и пересеченной местности, вариант построения беспроводной сети связи более экономичен. Однако на крупных сортировочных станциях и узлах, где остро стоит необходимость усовершенствования технических процессов, архитектура сети с использованием удаленных радиоголовок достойна повышенного внимания, в том числе и в составе составной радиолинии.

На первый взгляд такая техническая сложность в перспективе оправдывает себя, особенно если вопрос касается повышения безопасности на железнодорожном транспорте в условиях внедрения все более и более высокоскоростных подвижных составов. Схема применения архитектуры сети с удаленными радиоголовками представлена на рисунке 3.

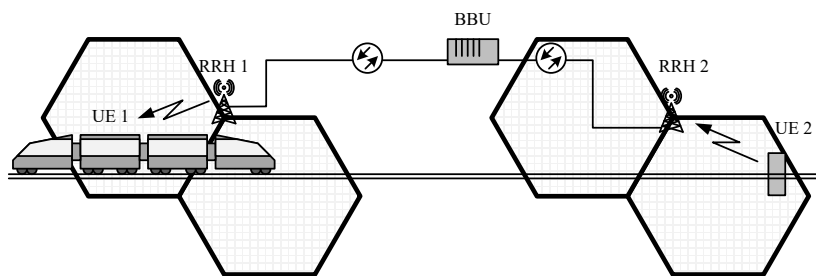


Рис. 3. Архитектура системы видеонаблюдения на железнодорожном переезде посредством внедрения удаленных радиоголовок

На рисунке 3 условными обозначениями показаны удаленные радиоголовки RRH 1 (Remote Radio Head) и RRH 2 в пределах приближения подвижного состава на расстояние не менее двойной максимальной длины тормозного пути при экстренном торможении и

зоне размещения железнодорожного переезда соответственно, и ВВU (Base Band Unit) – блока базовой полосы.

Таким образом, по результатам анализа различных вариантов построения сети беспроводной связи, можно сделать вывод о том, что возможности сетей стандарта LTE позволяют организовать систему видеотрансляции на железнодорожных переездах в любом регионе и при любых уже существующих архитектурах сети. Но в любом случае выбор варианта сети должен быть обоснован, в том числе с учетом возможных радиоэлектронных конфликтов с уже развернутым оборудованием, в том числе и LTE. В интересах других ведомств.

4. Разработка аналитической модели вероятностной оценки функционирования сети передачи изображений от систем видеонаблюдения. С учетом ограниченной дальности связи для передач диапазона 2600 МГц, искомую модель сети передачи информации от систем видеонаблюдения предлагается разрабатывать в виде составной линии радиосвязи. При этом особый акцент в разрабатываемой модели будет сделан на учет излучений сторонних источников (ИСИ) в восходящих линиях связи, работающих в режиме FHSS со скоростью $V_{\text{FHSS}} = 2000$ 1/с, и приводящих к ухудшению условий ЭМС, снижающих эффективность функционирования сети передачи изображения.

В качестве основы для разработки модели предлагается использовать результаты, полученные авторами в [24].

При обосновании исходных данных будем полагать, что искомая сеть также работает в режиме FHSS с аналогичной скоростью для минимизации вероятности долговременного назначения канала с многолучевым распространением радиоволн.

Таким образом, в качестве негативных факторов, приводящих к блокированию частотных каналов сети передачи изображений, будут выступать ИСИ и замирания в многолучевых каналах.

Тогда с позиций независимости источников деструктивных воздействий [25], вероятность блокирования каналов сети передачи информации, наступающей от систем видеонаблюдения, будет определяться следующим выражением

$$P_{\text{бл}} = P_3 + P_{\text{вп}} - P_3 \cdot P_{\text{вп}}, \quad (2)$$

где P_3 – вероятность блокирования, обусловленная замираниями и шумами в канале; $P_{\text{вп}}$ – вероятность блокирования канала, обусловленная помехами ИСИ при нарушении условий ЭМС.

В свою очередь величина $P_{\text{вп}}$ определяется соотношением рабочих частот n подверженных взаимным помехам к общему числу рабочих каналов m , задействованных при организации режима с FHSS, что позволяет записать:

$$P_{\text{вп}} = n / m . \quad (3)$$

При этом значение n можно учитывать следующим образом:

$$n = \sum_{i=1}^N i P_i ,$$

где N – общее количество радиосредств, совместно использующих группу из m частот, работа которых может привести к проблемам ЭМС; P_i – вероятность одновременной работы на излучение i радиостанций.

Величину P_i рассчитаем как

$$P_i = \frac{C_N^i (\alpha/\mu)^i}{\sum_{j=0}^N C_N^j (\alpha/\mu)^j} . \quad (4)$$

В выражении (4) использованы следующие значения:

λ – параметр, представляющий собой интенсивность заявок, определяемых удельной среднестатистической нагрузкой пользователя в сети;

$\mu = 1/m_{\text{об}}$ – параметр (интенсивность), характеризующий длительность сеанса связи, с использованием одной частоты (одного рабочего канала);

$m_{\text{об}} = \mathbf{M}[T]$ – среднее значение (математическое ожидание) длительности сеанса T ;

j – количество радиостанций, не использующих данную частоту, т.е. $j = N - i$.

С учетом значений, составляющих выражения (4), формулу (3) можно переписать к виду:

$$n = \frac{1}{p} \sum_{i=1}^N i C_N^i (\lambda/\mu)^i , \quad (5)$$

где $\rho = \sum_{j=0}^N C_N^j (\lambda/\mu)^j$ – суммарное значение всех возможных комбинаций отношения λ/μ , приходящихся на каждую j -ю радиостанцию.

Далее проведем разложение величины ρ в соответствии с биномом Ньютона [26]:

$$\sum_{j=0}^N C_N^j (\lambda/\mu)^j = (1 + \lambda/\mu)^N. \quad (6)$$

И с учетом выражения (6) перепишем формулу (4) к следующему виду:

$$P_i = C_N^i \frac{(1 + \lambda/\mu)^i}{(1 + \lambda/\mu)^N} = C_N^i \left(\frac{\lambda\mu^{-1}}{1 + \lambda\mu^{-1}} \right)^i \left(\frac{1}{1 + \lambda\mu^{-1}} \right)^{N-i}. \quad (7)$$

Полученное выражение (7) представляет собой распределение Бернулли, которое позволяет получить не только оценку вероятности занятости каждого из рабочих каналов $\lambda\mu^{-1} / (1 + \lambda\mu^{-1})$, при одновременной работе всех радиостанций, использующих рассматриваемый частотный ресурс, но и рассчитать вероятность противоположного события $1 / (1 + \lambda\mu^{-1})$ [27].

Рассмотренный подход позволяет интерпретировать рабочие каналы, подверженные взаимным помехам, как функцию, характеризующую количество радиосредств, одновременно работающих на излучение в одном частотном диапазоне с различной стратегией смены частотных позиций.

В этом случае выражение (3), характеризующее вероятность среднего числа рабочих частот подверженных взаимным помехам, в серии из N проведенных биномиальных испытаний, определяемых выражением (7), можно представить следующим образом [28]:

$$P_{\text{вп}} = \frac{N}{m} \left(\frac{\lambda\mu^{-1}}{1 + \lambda\mu^{-1}} \right), \quad (8)$$

где m – общее число рабочих каналов, используемых в режиме FHSS.

Блокировка канала, определяемая выражением (8), подразумевает собой вероятность того, что на частоте, используемой в режиме FHSS, в момент передачи на ней сигнала будет осуществляться работа ИСИ. При этом отношение N/m можно рассматривать как частоту одновременного выбора рабочей частоты различными источниками.

Однако следует заметить, что одновременное использование рабочей частоты несколькими ИСИ не приводит к однозначному блокированию канала, т.е. нарушению связи. Такое заключение обусловлено тем, что для блокировки канала, мощность сигналов ИСИ, в точке приема должна превышать допустимое значение энергетического соотношения между полезным и мешающим сигналом. Только в этом случае произойдет нарушение связи. Для оценки последствий деструктивных воздействий со стороны ИСИ воспользуемся требованиями, значения вероятности наступления которых $P_{\text{ВДФ}}$ приведены в таблице 2, согласно ГОСТ Р 53111-2008 «Устойчивость функционирования сети связи общего пользования», в котором определены последствия воздействия внешних дестабилизирующих факторов (ВДФ).

Таблица 2. Зависимость вероятности сохранения работоспособности элементов сети электросвязи от ущерба, причиняемого им воздействием ВДФ

Уровень ВДФ	Ущерб сети связи наносимый ВДФ, %	Вероятность сохранения работоспособности сети	Уровень нарушения безопасности сети
Высокий	до 50	0,5	Высокий
Средний	до 30	0,7	Умеренный
Низкий	до 10	0,9	Низкий

Тогда с учетом (8), выражение (2), характеризующее вероятность нарушения (блокирования) работы сети, определяемой помехами ИСИ при нарушении условий ЭМС, с учетом $P_{\text{ВДФ}}$ и примет следующий вид:

$$P_{\text{бл}} = P_3 + P_{\text{ВДФ}} \frac{N}{m} \left(\frac{\lambda \mu^{-1}}{1 + \lambda \mu^{-1}} \right) - P_3 \frac{N}{m} \left(\frac{\lambda \mu^{-1}}{1 + \lambda \mu^{-1}} \right) P_{\text{ВДФ}}. \quad (9)$$

Значение P_3 определяются текущим отношением мощности полезного сигнала к мощности канальных шумов и помех. Поскольку в сети стандарта LTE планируется, то требования к допустимой

вероятности битовой ошибки в канале, определяющие значение $P_{\text{бл}}$, могут быть определены в соответствии с [29] (таблица 3).

Таблица 3. Требования к вероятности битовой ошибки в канале для заданного качества видео контента

Уровень качества изображения	Допустимый диапазон вероятности ошибок
Высокое качество	менее 1×10^{-4}
Хорошее качество	от 1×10^{-4} до 4×10^{-4}
Удовлетворительное качество	от 4×10^{-4} до 8×10^{-4}
Плохое качество	от 8×10^{-4} до 1×10^{-3}
Очень плохое качество	более 1×10^{-3}

Следовательно, учитывая, что изначальная настройка сети обеспечивает высокое качество изображения, то в расчетах будем полагать $P_3 = 1 \times 10^{-5}$. В свою очередь выбор значений $P_{\text{ВДФ}}$ будет осуществляться с учетом требований таблицы 2.

5. Анализ результатов моделирования. Полученное выражение (9) открывает возможности по проведению моделирования, с целью раскрытия допустимых условий, характеризующих успешное функционирование сети передачи результатов видеонаблюдения.

С этой целью построена зависимость $P_{\text{бл}}$ как функции аргумента (λ/μ) при различных значениях $P_{\text{ВДФ}}$, определяемых требованиями таблицы 1, и значениях $N/m = 0,001$ и $N/m = 0,01$ (при ОСШ в канале, обеспечивающем вероятность битовой ошибки $P_3 = 1 \times 10^{-5}$).

Графики приведены на рисунках 4 и 5.

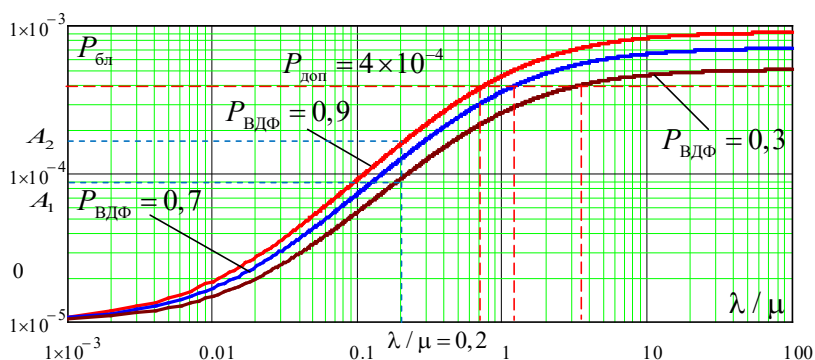


Рис. 4. Зависимость вероятности битовой ошибки в канале от соотношения интенсивностей работы сети, при различных требованиях к качеству передаваемого изображения (при $N/m = 0,001$)

Полученные результаты моделирования позволяют сделать следующее заключение. Величина $P_{\text{ВДФ}}$, несмотря на существенный разброс качественных показателей, определяемых ущербом, наносимым системе связи, вариативность вероятности битовой ошибки весьма незначительна.

В качестве примера, на рисунках 4 и 5 мелким пунктиром показаны значения вероятностей блокирования (точки A_1 и A_2), для отношения значений интенсивностей $\lambda/\mu = 0,2$ при различных величинах вероятности $P_{\text{ВДФ}}$, определяемых уровнем внешних деструктивных факторов.

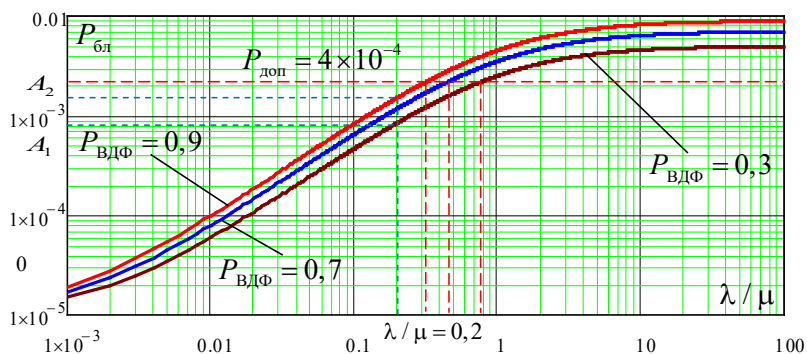


Рис. 5. Зависимость вероятности битовой ошибки в канале от соотношения интенсивностей работы сети, при различных требованиях к качеству передаваемого изображения (при $N/m = 0,01$)

Отметим, что различия $P_{\text{бл}}$ для уровней нарушения безопасности связи от высокого до низкого изменяются в диапазоне от 8×10^{-4} до от 2×10^{-3} . Вместе с тем различия в отношениях интенсивностей λ/μ при уровне $P_{\text{доп}} = 4 \times 10^{-4}$ в тех же условиях (отмечено крупным пунктиром), составляют 2,7 раза (при $N/m = 0,01$) и более чем в 7,1 раза (при $N/m = 0,001$).

Интересным видится зависимость величины $P_{\text{бл}}$ как функции отношения N/m , характеризующей одновременный выбор рабочей частоты (рисунок 6).

Согласно данным, представленным на рисунке 6, допустимая вероятность ошибки в канале, приводящая к блокированию его работы, наступает при отношении N/m порядка 0,001. При этом величина $P_{\text{бл}}$ достаточно слабо зависит от значения вероятности $P_{\text{ВДФ}}$.

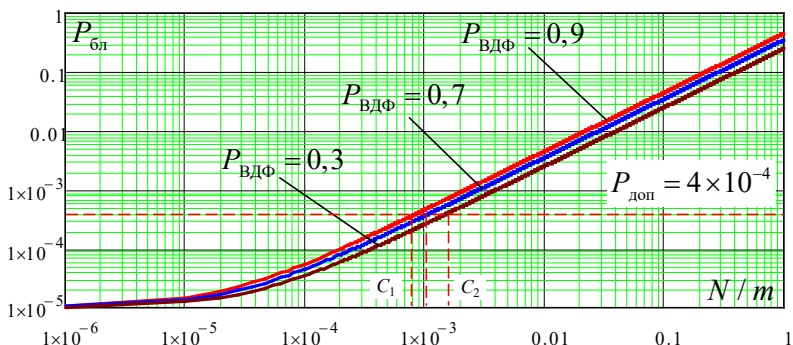


Рис. 6. Зависимость вероятности битовой ошибки в канале от соотношения N/m , при различных требованиях к качеству передаваемого изображения

Следовательно, можно сделать вывод, что работа сетей передачи видео в условиях взаимных помех со стороны ИСИ при нарушении требований по ЭМС, более чувствительна к изменению отношения значений интенсивностей потоков сети и ИСИ, работающих в режиме FHSS, чем к энергетическим соотношениям полезного и мешающего сигналов. Но это при условии, что минимальное значение начинается от допустимого уровня вероятности достоверного приема.

6. Заключение. Полученные результаты исследования работы сетей передачи информации от систем видеонаблюдения показали, что несмотря на применение технологий LTE, обладающих достаточно высокой помехозащищенностью приема, составные линии связи оказались уязвимы к воздействию сигналов сторонних РЭС, выступающих в качестве источников деструктивных воздействий непреднамеренного характера (при условии того, что средняя вероятность одновременного использования канала превысит вероятность 0,001).

При этом проведенные расчеты показали, что наиболее значимым фактором, определяющим помехозащищенность сетей, является соотношение интенсивностей смены частотных позиций конфликтующими сторонами, определяющих занятость рабочего канала. То есть, чем меньше задействуется пораженный канал при организации режима FHSS, тем выше эффективность функционирования сети. Причем фактор, характеризующий пригодность самого канала, не является столь определяющим, как продолжительность времени его использования.

С другой стороны, увеличение скорости смены рабочих каналов в режиме FHSS, ведет к уменьшению энергетического потенциала, и как результат – к уменьшению дистанции связи и возрастанию потребного числа ретрансляторов.

Авторы отмечают, что в проведенном исследовании не рассмотрено влияния скорости движения подвижного состава на характер многолучевого распространения и вероятность блокирования.

Дальнейшее направление исследования авторы связывают с повышением достоверности приема за счет применения методов совместной частотно-временной обработки [30, 31].

Литература

1. Синецкий А.С. Перспективы развития технических средств ЖАТ на переездах // Автоматика, связь, информатика. 2023. № 11. С. 10–11.
2. Ахмедзянов Г.Г., Дремин В.В., Литвинов А.В. Совершенствование учета параметров объектов железнодорожной инфраструктуры для повышения эффективности эксплуатации железнодорожного переездного комплекса // Современные наукоемкие технологии. 2022. № 1. С. 9–13. DOI: 10.17513/snt.39002.
3. Трунаев А.М., Иваницкая И.Л. Анализ средств обеспечения безопасности на железнодорожных переездах // Сборник научных трудов Донецкого института железнодорожного транспорта. 2021. № 63. С. 4–12.
4. Имарова О.Б. Следящая система управления торможением поезда в зоне сближения с неохраняемым и необслуживаемым железнодорожным переездом // Вестник Ростовского государственного университета путей сообщения. 2020. № 3(79). С. 43–51. DOI: 10.46973/0201-727X_2020_3_43.
5. Лукьянов А.С., Толстых Д. С., Буравцова А.Н. Технологии мобильного широкополосного доступа в сетях LTE и применение в ведомственных структурах // Вестник Воронежского института высоких технологий. 2019. № 13(1). С. 116–119.
6. Singh U., Dua A., Kumar N., Guizani M. QoS Aware Uplink Scheduling for M2M Communication in LTE/LTE-A Network: A Game Theoretic Approach // IEEE Transactions on Vehicular Technology. 2022. vol. 71. no. 4. pp. 4156–4170. DOI: 10.1109/TVT.2021.3132535.
7. Дворников С. В., Степанин Д. В., Дворников А. С., Букарева А. П. Формирование векторов признаков сигналов из вейвлет-коэффициентов их фреймовых преобразований // Информационные технологии. 2013. № 5. С. 46–49.
8. Фокин Г.А., Волгушев Д.Б. Использование SDR-технологии для задач сетевого позиционирования. Процедуры приема и обработки опорных сигналов LTE // Вестник СибГУТИ. 2023. Т. 17. № 1. С. 52–65. DOI: 10.55648/1998-6920-2023-17-1-52-65.
9. Dvornikov S.S., Zheglov K.D., Dvornikov S.V. SSB signals with controlled pilot level // T-Comm. 2023. vol. 17. no. 3. pp. 41–47. DOI: 10.36724/2072-8735-2023-17-3-41-47.
10. Abusalma A. The effect of implementing artificial intelligence on job performance in commercial banks of Jordan // Management Science Letters. 2021. vol. 11. pp. 2061–2070. DOI: 10.5267/j.msl.2021.3.003.

11. Артюшенко В.М., Воловач В.И. Оценка влияния гауссовского шума, помех гармонического типа и комбинированных помех на беспроводные устройства ближнего радиуса действия // Радиотехника. 2023. Т. 87. № 9. С. 124–137. DOI: 10.18127/j00338486-202309-11.
12. Алексеев В. Сравнительные характеристики широкополосных и узкополосных сетей LPWAN нелицензируемого диапазона для приложений M2M и IoT // Беспроводные технологии. 2019. № 3(56). С. 10–20.
13. Li J., Li X., Li H., et al. Fabrication of the Cu/AgCuTi/Nb composite for superconducting radio-frequency material under extreme service conditions based on electroplating additive manufacturing // Materials & Design. 2024. vol. 244. pp. 113220. DOI: 10.1016/j.matdes.2024.113220.
14. Flores Soriano M. Solar radio bursts impact on the International GNSS Service Network during Solar Cycle 24 // Journal of Space Weather and Space Climate. 2024. vol. 14. DOI: 10.1051/swsc/2024034.
15. Иванова Д.В., Маркова Е.В., Шоргин С.Я., Гайдамака Ю.В. Модели совместного обслуживания трафика EMBB и URLLC на основе приоритетов в промышленных развертываниях 5G NR // Информатика и ее применения. 2023. Т. 17. № 4. С. 64–70. DOI: 10.14357/19922264230409.
16. Журавлев Д.В., Кузьменко Р.В., Чепелев М.Ю., Минаков Д.С. Разработка и реализация алгоритма анализа сигналов стандарта Bluetooth // Вестник Воронежского института ФЦИН России. 2021. № 3. С. 14–22.
17. Vasilyeva D.V., Dvornikov S.V., Yakushenko S.A., Dvornikov S.S. Automation of detection procedures based on the results of processing images from video surveillance systems. Collection of abstracts of reports of the III International Forum “Mathematical methods and models in high-tech production”. 2023. pp. 157–158.
18. Vasilyeva D.V., Dvornikov S.V., Yakushenko S.A., Dvornikov S.S. Automation of forest fire detection procedures based on the results of video processing // Scientific and analytical journal “Bulletin of the St. Petersburg University of the State Fire Fighting Service of the Ministry of Emergency Situations of Russia”. 2024. vol. 2023. no. 4. pp. 47–58. DOI: 10.61260/2218-130X-2024-2023-4-47-58.
19. Дворников С.В., Дворников С.С., Коноплев М.А. Алгоритм распознавания сигналов радиосвязи на основе симметрических матриц. Информационные Технологии. 2010. № 9. С. 75–77.
20. Abbasov I.B. Image Recognition in Agriculture and Landscape Protection // International Journal of Science and Research. 2020. vol. 9. no. 12. pp. 757–763. DOI: 10.21275/SR201212144831.
21. Себекин Г.В., Шурков А.О., Маслов А.А., Андреев Ю.В. Построение мультисервисной платформы спутниковой связи на основе решений для сетей передачи данных стандарта LTE (3GPP) // Успехи современной радиоэлектроники. 2024. № 2. С. 66–75. DOI: 10.18127/j20700784-202402-06.
22. Батенков К.А., Королев А.В., Миронов А.Е. Эффективность использования канального ресурса при объединении и кластеризации передаточных возможностей звена мультисервисной сети связи // Информационные системы и технологии. 2020. № 2(118). С. 101–105.
23. Макаренко С.И. Интероперабельность организационно-технических систем. Санкт-Петербург: Научное издание, 2024. 313 с.
24. Дворников С.С., Дворников С.В., Леонов Д.М., Махфуд М.Г. Эффективность функционирования локальных радиосетей в сложной радиоэлектронной обстановке // Информация и космос. 2023. № 1. С. 29–34.
25. Дюк В.А., Малыгин И.Г. Сравнение алгоритмов распознавания типов транспортных средств по параметрам их силуэтов // Морские интеллектуальные технологии. 2018. № 4-4(42). С. 197–201.

26. Чубарико В.Н. Обобщённая формула бинома Ньютона и формулы суммирования. Чебышевский сборник. 2020. Т. 21. № 4. С. 270–301. DOI: 10.22405/2226-8383-2020-21-4-270-301.
27. Ананьевский С.М., Невзоров В.Б. О некоторых вероятностных распределениях, связанных с классической схемой Бернулли // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2022. Т. 9. № 2. С. 201–208. DOI: 10.21638/spbu01.2022.202.
28. Васильева Д.В., Дворников С.С., Толстуха Ю.Е., Обрезков П.С., Дворников С.В. Формирование Векторов признаков для систем видеонаблюдения // Вопросы Радиоэлектроники. Серия: Техника Телевидения. 2023. № 4. С. 62–68.
29. Шелухин О.И., Ахметшин Р.Р., Руднев А.Н., Иванов Ю.А. Влияние различных типов ошибок в каналах беспроводного доступа на качество систем видеонаблюдения // T-Comm: Телекоммуникации и транспорт. 2012. № 11. С. 62–66.
30. Sheikh H.R., Bovik A.C. Image information and visual quality // IEEE Transactions on Image Processing. 2006. vol. 15. no. 2. pp. 430–444. DOI: 10.1109/tip.2005.859378.
31. Ключко В.К., Ву Б.Х. Исследование частотно-временной обработки сигналов в системе доплеровских приемников // Цифровая обработка сигналов. 2024. № 4. С. 33–39.

Дворников Сергей Викторович — д-р техн. наук, профессор, кафедра радиотехнических и оптоэлектронных комплексов, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» (ГУАП); профессор, кафедра радиосвязи, Военная академия связи им. С.М. Буденного. Область научных интересов: построение помехозащищенных систем радиосвязи, формирование и обработка сигналов сложных структур. Число научных публикаций — 609. practcdsv@yandex.ru; Тихорецкий проспект, 67, 194064, Санкт-Петербург, Россия; р.т.: +7(812)247-9400.

Маслова Анна Андреевна — аспирант, кафедра электрическая связь, Федеральное государственное бюджетное образовательное учреждение высшего образования «Петербургский государственный университет путей сообщения Императора Александра I». Область научных интересов: разработка, и совершенствование методов исследования, моделирования и проектирования сетей, систем и устройств телекоммуникаций. Число научных публикаций — 10. Bloodyelis@yandex.ru; Московский проспект, 9, 190031, Санкт-Петербург, Россия; р.т.: +7(909)580-1511.

Васильева Дина Владимировна — аспирант, старший преподаватель, кафедры радиотехнических систем института радиотехники и инфокоммуникационных технологий, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» (ГУАП). Область научных интересов: цифровая обработка сигналов. Число научных публикаций — 45. dolli.dina@mail.ru; улица Большая Морская, 67, 190000, Санкт-Петербург, Россия; р.т.: +7(931)385-9181.

S. DVORNIKOV, A. MASLOVA, D. VASILIEVA
**PROBABILISTIC ESTIMATION OF NOISE IMMUNITY
OF COMPOSITE VIDEO TRANSMISSION RADIO LINES
IN CONDITIONS OF MUTUAL INTERFERENCE**

Dvornikov S., Maslova A., Vasilieva D. Probabilistic Estimation of Noise Immunity of Composite Video Transmission Radio Lines in Conditions of Mutual Interference.

Abstract. The transition to automated regulation systems at unguarded railroad crossings has necessitated the solution of their safety issues. The most rational solution to this problem is the use of video surveillance systems that provide the transmission of images both to the railroad stations, in the area of responsibility of which the railroad crossings are located, and to the locomotives of rolling stock. For this purpose, information transmission systems organized on the basis of broadband and LTE networks are actively used. But since the operation of such networks is organized along the railway tracks, including in conditions of active application of various radiating devices, they are characterized by violation of electromagnetic compatibility conditions, leading to failure of operation as a result of inadvertent blocking of separate channels. Therefore, the analysis of conditions under which the failure of the video transmission network occurs, as well as the predictive calculation of the parameters of radio lines, which provides a given level of stability of the network, is relevant. Technologies and peculiarities of LTE standard networks operation are considered. Indicators and criteria for evaluating the functioning of video transmission lines within the technical capabilities of the standard are substantiated. The mathematical formulation of the research problem is carried out. The initial data for the development of an analytical model of probabilistic assessment of video transmission network functioning are determined. The analytical apparatus for calculating the probability of channel blocking, taking into account the mutual intensity of frequency traffic usage by conflicting devices, is developed. The requirements of GOST R 53111-2008, defining the conditions under which the stability of public communication network operation is ensured, are analyzed. The expression of probabilistic estimation characterizing the probability of disruption (blocking) of network operation determined by both channel noise and fading, and unintentional interference from third-party sources of radio emissions due to violations of electromagnetic compatibility is obtained. The results of analytical modeling are presented, revealing the conditions under which the successful functioning of the video surveillance transmission network is ensured. It is substantiated that the operation of video transmission networks under conditions of mutual interference in violation of electromagnetic compatibility requirements is more sensitive to changes in the ratio of intensity values of network streams and sources of third-party radiations operating in the mode with programmed tuning of the operating frequency than to energy ratios of useful and interfering signals.

Keywords: mutual interference, video transmission systems, LTE networks, probability of network blocking, interference immunity of radio lines.

References

1. Sinetsky A.S. [Prospects for the development of technical means of Railway Automated Traffic Control at crossings]. *Avtomatika, svjaz', informatika – Automation, Communications, Informatics*. 2023. no. 11. pp. 10–11. (In Russ.).
2. Akhmedzyanov G.G., Dremine V.V., Litvinov V.V., Litvinov A.V. [Improvement of the railway infrastructure objects parameters accounting for the railway crossing complex operation efficiency increase]. *Sovremennye naukoemkie tehnologii –*

- Modern Science-Intensive Technologies. 2022. no. 1. pp. 9–13. DOI: 10.17513/snt.39002. (In Russ.).
3. Трунаев А.М., Иванецкая И.Л. Analiz sredstv obespecheniya bezopasnosti na zheleznodorozhnyh pereezdah [Analysis of the means to ensure safety at the railway crossings]. Sbornik nauchnykh trudov Doneckogo instituta zheleznodorozhnogo transporta [Collection of scientific papers of Donetsk Institute of Railway Transport]. 2021. no. 63. pp. 4–12. (In Russ.).
4. Imarova O.B. [Tracking system of train braking control in the zone of approaching an unguarded and unattended railroad crossing]. Vestnik Rostovskogo gosudarstvennogo universiteta putej soobshheniya – Vestnik of Rostov State University of Railway Communication. 2020. no. 3(79). pp. 43–51. DOI: 10.46973/0201-727X_2020_3_43. (In Russ.).
5. Lukyanov A.S., Tolstykh D.S., Buravtsova A.N. [Technologies of mobile broadband access in LTE networks and application in departmental structures]. Vestnik Voronezhskogo instituta vysokikh tehnologij – Bulletin of the Voronezh Institute of High Technologies. 2019. no. 13(1). pp. 116–119. (In Russ.).
6. Singh U., Dua A., Kumar N., Guizani M. QoS Aware Uplink Scheduling for M2M Communication in LTE/LTE-A Network: A Game Theoretic Approach. IEEE Transactions on Vehicular Technology. 2022. vol. 71. no. 4. pp. 4156–4170. DOI: 10.1109/TVT.2021.3132535.
7. Dvornikov S.V., Stepynin D.V., Dvornikov A.S., Bukareva A.P. [Formation of signal feature vectors from wavelet coefficients of their frame transformations]. Informacionnye tehnologii – Information Technologies. 2013. no. 5. pp. 46–49. (In Russ.).
8. Fokin G.A., Volgushev D.B. [Using SDR-technology for network positioning tasks. Procedures of reception and processing of LTE reference signals]. Vestnik SibGUTI – Bulletin of SibGUTI. 2023. vol. 17. no. 1. pp. 52–65. DOI: 10.55648/1998-6920-2023-17-1-52-65. (In Russ.).
9. Dvornikov S.S., Zheglov K.D., Dvornikov S.V. SSB signals with controlled pilot level. T-Comm. 2023. vol. 17. no. 3. pp. 41–47. DOI: 10.36724/2072-8735-2023-17-3-41-47.
10. Abusalma A. The effect of implementing artificial intelligence on job performance in commercial banks of Jordan. Management Science Letters. 2021. vol. 11. pp. 2061–2070. DOI: 10.5267/j.msl.2021.3.003.
11. Artyushenko V.M., Volovach V.I. [Evaluation of the influence of the Gaussian noise, harmonic type interference and combined interference on the short-range wireless devices]. Radiotekhnika – Radio engineering. 2023. vol. 87. no. 9. pp. 124–137. DOI: 10.18127/j00338486-202309-11. (In Russ.).
12. Alekseev V. [Comparative characteristics of broadband and narrowband unlicensed LPWANs for M2M and IoT applications]. Besprovodnye tehnologii – Wireless Technologies. 2019. no. 3(56). pp. 10–20. (In Russ.).
13. Li J., Li X., Li H., et al. Fabrication of the Cu/AgCuTi/Nb composite for superconducting radio-frequency material under extreme service conditions based on electroplating additive manufacturing. Materials & Design. 2024. vol. 244. pp. 113220. DOI: 10.1016/j.matdes.2024.113220.
14. Flores Soriano M. Solar radio bursts impact on the International GNSS Service Network during Solar Cycle 24. Journal of Space Weather and Space Climate. 2024. vol. 14. DOI: 10.1051/swsc/2024034.
15. Ivanova D.V., Markova E.V., Shorgin S.Ya., Gaidamaka Y.V. [Models of EMBB and URLLC traffic co-servicing based on priorities in industrial 5G NR deployments]. Computer science and its applications – Informatics and its applications. 2023. vol. 17. no. 4. pp. 64–70. DOI: 10.14357/19922264230409. (In Russ.).

16. Zhuravlev D.V., Kuzmenko R.V.V., Chepelev M.Yu, Minakov D.S. [Development and implementation of the algorithm for analyzing signals of Bluetooth standard]. Vestnik Voronezhskogo instituta FSIN Rossii – Vestnik Voronezh Institute of FSIN Russia. 2021. no. 3. pp. 14–22. (In Russ.).
17. Vasilyeva D.V., Dvornikov S.V., Yakushenko S.A.I., Dvornikov S.S. Automation of detection procedures based on the results of processing images from video surveillance systems. Collection of abstracts of reports of the III International Forum “Mathematical methods and models in high-tech production”. 2023. pp. 157–158.
18. Vasilyeva D.V., Dvornikov S.V., Yakushenko S.A., Dvornikov S.S. Automation of forest fire detection procedures based on the results of video processing. Scientific and analytical journal “Bulletin of the St. Petersburg University of the State Fire Fighting Service of the Ministry of Emergency Situations of Russia”. 2024. vol. 2023. no. 4. pp. 47–58. DOI: 10.61260/2218-130X-2024-2023-4-47-58.
19. Dvornikov S.V., Dvornikov S.S., Konoplev M.A. [Algorithm of radio communication signals recognition on the basis of symmetric matrices]. Informacionnye Tehnologii – Information technologies. 2010. no. 9. pp. 75–77. (In Russ.).
20. Abbasov I.B. Image Recognition in Agriculture and Landscape Protection. International Journal of Science and Research. 2020. vol. 9. no. 12. pp. 757–763. DOI: 10.21275/SR201212144831.
21. Sebekin G.V., Shchurkov A.O., Maslov A.A., Andreev Yu. V. Building a multiservice platform of satellite communication on the basis of solutions for LTE (3GPP) data networks. Uspehi sovremennoj radioelektroniki – The successes of modern radio electronics. 2024. no. 2. pp. 66–75. DOI: 10.18127/j20700784-202402-06. (In Russ.).
22. Batenkov K.A., Korolev A.V., Mironov A.E. [Efficiency of channel resource utilization at unification and clustering of transmission capabilities of a link of a multiservice communication network]. Informacionnye sistemy i tehnologii – Information systems and technologies. 2020. no. 2(118). pp. 101–105. (In Russ.).
23. Makarenko S.I. Interoperabel'nost' organizacionno-tehnicheskikh sistem [Interoperability of organizational and technical systems]. Sankt-Peterburg: Naukoemkie tehnologii, 2024. 313 p. (In Russ.).
24. Dvornikov S.S., Dvornikov S.V., Leonov D.M., Makhfud M.G. [Efficiency of using local radio networks in a digital radio-electronic system]. Informacija i kosmos – Information and space. 2023. no. 1. pp. 29–34. (In Russ.).
25. Duke V.A., Malygin I.G. [Comparison of algorithms for recognizing vehicle types based on the parameters of their silhouettes]. Morskie intellektual'nye tehnologii – Marine Intelligent Technologies. 2018. no. 4-4(42). pp. 197–201. (In Russ.).
26. Chubariko V.N. [Generalized formula of Newton's binomial and summation formulas]. Chebyshevskij sbornik – Chebyshev Collection. 2020. vol. 21. no. 4. pp. 270–301. DOI: 10.22405/2226-8383-2020-21-4-270-301. (In Russ.).
27. Ananievsky S.M., Nevzorov V.B. [On some probability distributions related to the classical Bernoulli scheme]. Vestnik Sankt-Peterburgskogo universiteta. Matematika. Mehanika. Astronomija – Vestnik of St. Petersburg University. Mathematics. Mechanics. Astronomy. 2022. vol. 9. no. 2. pp. 201–208. DOI: 10.21638/spbu01.2022.202. (In Russ.).
28. Vasilyeva D.V., Dvornikov S.S., Tolstukha Yu.E., Obrezkov P.S., Dvornikov S.V. [Formation of feature vectors for video surveillance systems]. Voprosy Radioelektroniki. Serija: Tehnika Televidenija – Radio Electronics issues. Series: Television Technology. 2023. no. 4. pp. 62–68. (In Russ.).
29. Shelukhin O.I., Akhmetshin R.R., Rudnev A.N., Ivanov Yu.A. [Influence of different types of errors in wireless access channels on the quality of video surveillance systems]. T-Comm: Telecommunications and Transportation. 2012. no. 11. pp. 62–66. (In Russ.).

30. Sheikh H.R., Bovik A.C. Image information and visual quality. IEEE Transactions on Image Processing. 2006. vol. 15. no. 2. pp. 430–444. DOI: 10.1109/tip.2005.859378.
31. Klochko V.K., Wu B.H. [Investigation of the frequency-time signal processing in the system of Doppler receivers]. Cifrovaja obrabotka signalov – Digital Signal Processing. 2024. no. 4. pp. 33–39. (In Russ.).

Dvornikov Sergey — Ph.D., Dr.Sci., Professor, Department of radio engineering and optoelectronic complexes, Federal State Autonomous Educational Institution of Higher Education «St. Petersburg State University of Aerospace Instrumentation» (SUAI); Professor, Department of radio communications, Military Academy of Communications. Research interests: construction of noise-protected radio communication systems, formation and processing of signals of complex structures. The number of publications — 609. practicsv@yandex.ru; 67, Tikhoretsky Av., 194064, St. Petersburg, Russia; office phone: +7(812)247-9400.

Maslova Anna — Postgraduate student, Department of electrical communications, Department of Electrical Communications, Federal State Budgetary Educational Institution of Higher Education «St. Petersburg State University of Railway Transport of the Emperor Alexander I». Research interests: development and improvement of methods of research, modeling and design of networks, systems and devices of telecommunications. The number of publications — 10. Bloodyelis@yandex.ru; 9, Moskovsky Av., 190031, St. Petersburg, Russia; office phone: +7(909)580-1511.

Vasilieva Dina — Postgraduate student, senior lecturer, Department of radio engineering systems, institute of radio engineering and info-communication technologies, Federal State Autonomous Educational Institution of Higher Education «St. Petersburg State University of Aerospace Instrumentation» (SUAI). Research interests: digital signal processing. The number of publications — 45. dolli.dina@mail.ru; 67, Bolshaya Morskaya St., 190000, St. Petersburg, Russia; office phone: +7(931)385-9181.

А.М. МАЕВСКИЙ, В.А. РЫЖОВ, Т.А. ФЕДОРОВА, И.А. ПЕЧАЙКО,
Н.М. БУРОВ

СТОХАСТИЧЕСКАЯ ДИНАМИЧЕСКАЯ МОДЕЛЬ ПОДВОДНОЙ БЕСПРОВОДНОЙ СЕНСОРНОЙ СЕТИ, ОСНОВАННАЯ НА МОДИФИЦИРОВАННОМ ЛУВЕНСКОМ АЛГОРИТМЕ КЛАСТЕРИЗАЦИИ

Маевский А.М., Рыжов В.А., Федорова Т.А., Печайко И.А., Буров Н.М. Стохастическая динамическая модель подводной беспроводной сенсорной сети, основанная на модифицированном лувенском алгоритме кластеризации.

Аннотация. В данной работе представлен инновационный подход к кластеризации и маршрутизации в подводных беспроводных сенсорных сетях (ПБСС), основанный на модифицированном Лувенском алгоритме учитывающим расстояния между узлами, вероятности успешной передачи сообщений и текущий уровень энергии сенсоров. Класс задач, для которых используется указанный подход – гибридные подводные акустические беспроводные сенсорные сети с мобильным (надводным или подводным) шлюзом. Предложенный метод реализует механизм динамической рекластеризации сети на основе мониторинга энергетических ресурсов узлов, что позволяет перераспределять роли сенсоров в сети (обыкновенных и референсных) и обеспечивать адаптацию сети к изменениям внешней среды и состоянию сенсоров. Разработанный алгоритм нацелен на повышение энергоэффективности, минимизацию потерь данных и сокращение числа ретрансляций при передаче сообщений в условиях ограниченной пропускной способности гидроакустического канала связи. MAC-уровень базируется на TDMA протоколе, исключая возможность коллизий за счёт использования независимых временных слотов для сенсоров. Динамическая модель решает ключевые задачи управления ресурсами ПБСС: снижает энергозатраты, повышает надёжность передачи данных, уменьшает длительность передачи сообщений и продлевает срок автономной работы сети. Модель учитывает пространственную трёхмерную структуру размещения сенсоров и оптимизирует размещение референсных узлов, предотвращая перегрузку и появление «узких мест» в сети. Целью работы является построение топологии, минимизирующей энергопотребление и потери сообщений при эффективной маршрутизации данных к мобильному шлюзу. Гибкость и адаптивность предложенного подхода делает его применимым в реальных подводных задачах мониторинга окружающей среды и исследованиях мирового океана. Модифицированный Лувенский алгоритм с введенным порогом кластеризации оказался эффективным инструментом, обеспечивающим детальный анализ влияния проектных параметров на основные функциональные характеристики ПБСС. Сочетание модифицированного Лувенского алгоритма и модифицированного метода Дейкстры в рамках предложенной имитационной стохастической модели позволяет создавать устойчивую, адаптивную и энергоэффективную кластерную архитектуру сети.

Ключевые слова: подводные беспроводные акустические сенсорные сети, алгоритмы кластеризации, алгоритмы маршрутизации, Лувенский алгоритм, модифицированный алгоритм Дейкстры, TDMA протокол, энергоэффективность и надежность подводной беспроводной акустической сенсорной сети.

1. Введение. В данной работе предложен инновационный подход к кластеризации и маршрутизации для ПБСС, который

использует модифицированный Лувенский алгоритм. Оригинальность предлагаемого метода заключается в том, что он учитывает как расстояние между сенсорами, и, как следствие, вероятность успешной доставки сообщения, так и текущий уровень их энергии. Такой подход позволяет адаптировать топологию сети в реальном времени в зависимости от изменений в подводной среде и энергетического баланса узлов и проводить динамическую рекластеризацию сети. Это обеспечивает равномерное распределение нагрузки между сенсорами, что в свою очередь увеличивает срок автономной работы сети и снижает вероятность перегрузки отдельных узлов. Основная цель разработанного алгоритма – обеспечить энергоэффективность сети, минимизировать потери данных и количество ретрансляций в условиях ограниченной пропускной способности гидроакустических каналов связи и энергетического ресурса батарей сенсоров.

Главным новшеством является механизм динамической рекластеризации, основанный на мониторинге уровня энергии сенсоров. При снижении энергетических ресурсов узлов их роли перераспределяются, что минимизирует число ретрансляций и поддерживает высокую эффективность сети. Также в подходе реализован MAC протокол на основе TDMA, который исключает коллизии, так как предписывает сенсорам работать в независимых временных слотах, исключая взаимные помехи в гидроакустических каналах связи.

Представленный метод маршрутизации решает несколько ключевых проблем управления ресурсами в ПБСС. Во-первых, благодаря учёту вероятности успешной передачи данных и оптимизации маршрутов, снижается количество потерянных сообщений. Во-вторых, энергопотребление сети уменьшается за счёт балансировки нагрузки и рационального использования узлов-ретрансляторов. Одновременная передача сообщений сенсоров из разных кластеров в одних и тех же временных слотах сокращает общую продолжительность передачи сообщений. Наконец, динамическая рекластеризация обеспечивает надёжность сети, увеличивает её устойчивость к отказам и продлевает срок автономной работы, что делает метод применимым для реальных подводных задач, таких как мониторинг экосистем или разведка ресурсов.

Подводная сеть сенсоров рассматривается как пространственная 3D структура, где сенсоры размещены в толще воды на различных глубинах с шагом развертывания, значительно меньшим размеров области акватории. Сенсоры передают собранные данные о состоянии окружающей среды через центральные референсные узлы в каждом

кластере на мобильный шлюз. Передача данных на референсный узел осуществляется либо напрямую, либо через последовательность ретрансляционных узлов.

Важной задачей при оптимизации работы ПБСС является задача кластеризации. Разбиение сенсорной сети на кластеры определяется прежде всего оптимальным расположением референсных узлов, так как именно к ним стекается вся информация от обыкновенных узлов сети. Очевидно, что пропускная способность референсного узла ограничена и при большой размерности сети и неудачном расположении возможно появление «узких мест», в которых резко возрастают энергозатраты, приводящие к быстрой разрядке аккумуляторов. Очевидно, кластеризация будет зависеть от топологии самой сети.

Общую задачу маршрутизации, которая будет решаться ниже, можно сформулировать следующим образом – нахождение такого разбиения множества сенсоров на кластеры и поиск в кластерах оптимальных маршрутов доставки сообщений, при котором:

- Минимизируются энергетические затраты на передачу сообщений внутри каждого кластера;
- Максимизируется вероятность успешной передачи данных к референсным узлам;
- Балансируется нагрузка на сенсоры, чтобы предотвратить преждевременную потерю энергии отдельными узлами.

Основная цель – это оптимизация топологии сети, т.е. формирование такой структуры сети, которая минимизирует затраты энергии и потери сообщений при эффективной маршрутизации с уменьшением числа ретрансляций и оптимизации маршрутов для передачи данных к референсным узлам и далее к мобильному шлюзу.

В модели заложена возможность динамической адаптации: при изменении энергетического состояния или топологии (например, из-за перемещения сенсоров или изменения условий среды) проводится перераспределение ролей (основных и референсных узлов) и пересмотр кластерной структуры. Этот формализм обеспечивает основу для разработки подхода к кластеризации и маршрутизации, ориентированного на энергоэффективность и надёжность работы сети в условиях динамической подводной среды.

2. Обзор литературы. Проектирование эффективных и надежных ПБСС предполагает решение проблем, связанных с серьезными особенностями подводного акустического канала, определяемыми его динамическими свойствами – большой задержкой распространения сигнала, ограниченной полосой пропускания, низкой

скоростью передачи данных, высокой вероятностью битовых ошибок [1, 2]. К тому же, если принять во внимание то, что ПБСС потребляют гораздо больше энергии, чем наземные сенсорные сети [3], а технические компоненты ПБСС имеют ограниченные энергоресурсы, пополнение которых проблематично [4], то управление энергопотреблением в ПБСС, определяющее эксплуатационную жизнь сети, является большой проблемой.

Помимо указанного, нужно учитывать, что различные практические приложения, такие как сбор океанографических данных, мониторинг загрязнений, морская разведка, навигационная безопасность, распределенное тактическое наблюдение [5, 6], накладывают свои специфические ограничения и требования на архитектуру и проектные параметры разрабатываемых ПБСС.

Одной из главных целей проектирования эффективной ПБСС, независимо от сценариев приложений, является достижение высокого коэффициента доставки пакетов при ограничении сквозной задержки и энергопотребления компонентов сети [7], т.е. обеспечение высокой надежности сети.

Очевидно, что проектирование ПБСС, отвечающей критериям надежности и энергоэффективности, является комплексной и достаточно сложной проблемой. Рассматриваемая оптимизационная задача является многопараметрической и многокритериальной, требует всестороннего анализа влияния на основные функциональные характеристики ПБСС совокупности параметров различной природы: как технических, так и физических.

Подобному анализу посвящено большое количество исследований, связанных с разработкой энергоэффективных и надежных ПБСС.

Энергоэффективность определяет время жизни сети (при ограниченной емкости батарей узлов-сенсоров). В процессе эксплуатации в сети могут появляться энергетические дыры/пустые области, связанные с прекращением работы некоторых узлов и приводящие к снижению производительности сети. Причинами появления пустых областей являются: выход из строя узлов из-за ограниченного срока службы источников питания; проблемы, обусловленные повреждением сенсоров; разреженная топология сети; изменчивая характеристика гидроакустического канала; воздействие технических объектов и морских животных. При выходе из строя некоторого числа узлов сеть теряет работоспособность полностью или частично (могут оказаться недоступны группы узлов, маршруты доставки сообщений от которых проходили через вышедшие из строя

узлы). Энергоэффективность сети зависит от большого числа физических и технических параметров, связанных со средой функционирования, топологией сети, техническими параметрами приемо-передающих устройств. Корректный учет/выбор этих параметров позволяет снизить негативный эффект от появления «пустых областей» в сети [8 – 11].

Существует большое количество работ, посвященных улучшению энергоэффективности ПБСС. Эти работы были связаны с исследованиями по следующим направлениям:

- улучшением технических параметров источников энергии [12 – 16],
- контролем / управлением топологией сети [17 – 19];
- использованием мобильных узлов/шлюзов [17, 19 – 21];
- разработкой энергоэффективных протоколов маршрутизации [22 – 28];
- разработкой эффективных протоколов доступа к среде (MAC-протоколов) [29 – 32];
- модернизацией протоколов физического уровня (методов формирования и обработки сигнала) / улучшением технических параметров модемов [33 – 35].

Надежность сети определяется коэффициентом доставки пакетов (Packet Delivery Ratio – PDR) / коэффициентом потери пакетов (Packet Loss Ratio – PLR), а также сквозной задержкой передачи пакета. Указанные характеристики определяют качество обслуживания с точки зрения пользователя конкретного приложения.

Вероятность доставки пакетов определяется отношением количества пакетов, успешно достигших шлюза, к количеству пакетов, переданных узлом-источником. Сквозная задержка доставки пакета определяется средней задержкой передачи между узлом-источником и узлом-приемником (шлюзом).

Вероятность доставки пакетов связана с топологией сети – количеством сенсорных узлов и плотностью их распределения в акватории, с используемым протоколом маршрутизации, с нагрузкой сети, с использованием в составе сети мобильных шлюзов, коэффициентом затухания сигнала, мощностью излучателя, чувствительностью приемника, окружающими шумами.

В исследованиях показано [36], что значение PDR увеличивается по мере увеличения плотности узлов сети для большинства протоколов маршрутизации. Это связано с тем, что чем больше количество узлов в сети, тем выше вероятность корректного выбора следующего узла пересылки. Таким образом, снижается

вероятность потери пакетов из-за недоступности узла в пределах зоны пересылки. При этом средняя сквозная задержка доставки пакета уменьшается.

Эффективные алгоритмы протоколов маршрутизации разрабатываются таким образом, чтобы обеспечить лучшее значение PDR при ограничении на энергопотребление и сквозную задержку [36].

Задача обеспечения высоких значений PDR для эффективного протокола маршрутизации, связана с решением проблемы уменьшения негативного влияния энергетических дыр/пустых областей. Результаты исследований влияния на PDR различного количества пустых областей в акватории с фиксированным числом узлов показывают, что с увеличением плотности узлов в акватории при одинаковом количестве пустых областей для эффективных протоколов маршрутизации PDR возрастает. PDR также связан с нагрузкой сети. Как правило, с увеличением нагрузки сети (количеством пересылаемых пакетов в единицу времени) возрастает количество потерянных пакетов, что приводит к снижению PDR. Поэтому разрабатываемые протоколы маршрутизации нацелены на улучшение производительности, связанной с увеличением пропускной способности сети и уменьшением длительных сквозных задержек.

Задачей, решение которой влияет и на эффективность, и на надёжность функционирования ПБСС является управление топологией сети.

Контроль / управление топологией сети с помощью моделей кластеризации [37] является одним из важных инструментов, позволяющим сбалансировать энергопотребление, продлить срок службы сети и повысить надёжность. При этом модель кластеризации должна быть динамической и учитывать текущее состояние системы (как энергетическое, так и связанное с изменением внешней среды). Ранее разрабатывались модели, поддерживающие процессы на отдельных уровнях, например, протоколы для транспортного или сетевого уровней [38]. Они не использовали информацию, имеющуюся на других уровнях, например, на физическом, и совместной оптимизации решения не проводилось. Поэтому общая производительность моделируемых сетей была относительно невысокой [40 – 42]. Перспективным является развитие моделей/протоколов, использующие кроссуровневую структуру [39, 43], которая позволяет эффективно использовать данные, полученные на разных уровнях системной модели OSI.

Перечисленные факты, связанные с задачами повышения энергоэффективности и надежности сети, определяют направления исследований по разработке моделей/протоколов/алгоритмов функционирования ПБСС.

Основываясь на выполненном анализе существующих исследований, в рамках настоящей работы авторы предлагают новую имитационную стохастическую динамическую кроссуровневую модель ПБСС, использующую на сетевом и MAC уровнях – модифицированный Лувенский алгоритм, алгоритм Дейкстры, TDMA протокол, и обеспечивающую построение оптимальной топологии, минимизирующей энергопотребление и потери сообщений при маршрутизации данных к мобильному шлюзу.

3. Моделирование сложных динамических сетей. Рассмотрим общие свойства и универсальные закономерности, характерные для всех сетей, несмотря на их разнообразие. Сети можно определить как мультиагентные системы, состоящие из определённого числа узлов, соединённых между собой связями (рёбрами). Архитектура сети определяется её топологией – способом соединения узлов. Подводная беспроводная сенсорная сеть включает множество из N сенсоров $B = \{b_1, b_2, \dots, b_N\}$ где каждый сенсор b_i характеризуется определенным набором параметров.

Особенностью сетей является их сложность. Сложные сети могут включать множество узлов, с разным количеством связей между ними, и эти связи могут изменяться со временем. Такое поведение сетей делает их изучение важным для понимания структурных и динамических свойств, таких как появление, исчезновение узлов или изменение связей. Примеры таких сетей включают биологические, социальные и технические системы.

Традиционная задача изучения сложных сетей (мультиагентных систем) заключается в анализе их структуры и описании основных характеристик. Однако, простого понимания структуры недостаточно. Для изучения сложных сетей необходимо учитывать их динамическую природу, так как изменения в связях и узлах определяют функциональность и поведение сети в целом.

Важен переход к более высокому уровню анализа сетей, где учитываются динамические процессы, происходящие в узлах и связях. Для таких систем определение только структурных свойств недостаточно. Необходимо учитывать состояние сети – значения динамических переменных, описывающих поведение узлов и связей. Основной задачей становится изучение динамики сети, то есть изменений, происходящих во времени в узлах и связях. Если

изменения состояния сети происходят значительно быстрее, чем изменения её топологии, то можно считать топологию статической (фиксированной во времени) и анализировать влияние структурных и динамических параметров на активность сети. Однако в более сложных случаях, когда узлы и связи изменяются в течение времени, требуется учитывать их эволюцию.

В полной мере все сказанное относится и к подводным беспроводным сенсорным сетям, представляющим собой яркий пример динамически изменяющихся мультиагентных систем.

В теории сложных сетей основными объектами изучения являются ансамбли, состоящие из конечного числа N узлов (вершин), которые соединены между собой связями (рёбрами). Топология связей задаётся квадратной матрицей смежности (или матрицей связей), обозначаемой как $A = \{a_{ij}\}, i, j = 1, \dots, N$. Элементы этой матрицы указывают на существование связи между узлами i и j и принимают значения:

$a_{ij} = 1$ означает, что между узлами i и j существует связь;

$a_{ij} = 0$ означает, что связи между узлами i и j нет.

Связи между узлами могут быть направленными (ориентированными) или ненаправленными. Для ориентированных сетей в матрице смежности $a_{ij} \neq a_{ji}$, так как направления связи $i \rightarrow j$ и $j \rightarrow i$ различны. В ненаправленных сетях матрица симметрична ($a_{ij} = a_{ji}$).

Более сложные сети могут быть описаны с помощью взвешенных связей, где каждой связи присваивается числовой вес w_{ij} , что позволяет учитывать её интенсивность или значимость. Такие связи описываются весовой матрицей W .

В случае взвешенных подводных беспроводных сенсорных сетей матрица W может обозначать, например, вероятность доставки сообщения. Такое представление дает целостный взгляд на сеть, позволяя эффективно определять прямые и косвенные пути связи. Весовые матрицы позволяют анализировать и оптимизировать маршруты передачи данных благодаря поиску кратчайших путей – с помощью алгоритмов графов (например, алгоритм Дейкстры или Флойда-Уоршелла) можно быстро находить кратчайшие пути между узлами, снижая задержки и энергопотребление. Веса в матрице (например, пропускная способность или емкость узлов) позволяют равномерно распределять трафик. Приоритет коротких и надежных

маршрутов помогает снижать энергозатраты в ресурсозависимых сетях, таких как сенсорные.

В нашем случае весовые матрицы содержат вероятности успешной передачи сообщений $w_{ij} = p(r_{ij})$ между сенсорами i и j .

4. Модифицированный Лувенский метод. Основной целью использования модифицированного Лувенского метода является решение задачи разбиения сети на k кластеров $(C_1, C_2, \dots, C_k)$. Поиск оптимального числа кластеров является отдельной задачей, которая решается в зависимости от выбираемой топологии сети [17 – 19], от наличия в гибридной ПБСС одного или нескольких мобильных шлюзов, частоты съема информации в сети и скорости движения шлюза. В данной работе оптимальным для подводных сетей было принято разбиение, когда число кластеров составляет примерно 5% от общего числа сенсоров в акватории [44]. Это позволит исключить коллизии между различными кластерами и построить маршрут мобильного шлюза таким образом, чтобы данные успевали передаваться на него даже при высокой частоте съема информации.

В качестве базового алгоритма кластеризации в нашем случае был выбран достаточно популярный Лувенский метод (Louvain), который показывает большую точность и превосходит существующие аналоги по вычислительной скорости на сетях с известной структурой [45]. Он используется для анализа сетей, таких как социальные, коммуникационные или сенсорные, и применяется для поиска групп узлов, которые тесно связаны друг с другом.

В основе алгоритма лежит оптимизация коэффициента модулярности M (меры качества разбиения сети на кластеры) кластеров вершин. Она используется для оценки того, насколько хорошо сеть разделена на группы узлов, тесно связанных друг с другом, с минимальным числом связей между различными группами. Коэффициент модулярности M в оригинальном алгоритме представляет собой плотность ребер внутри кластера в диапазоне $[-1, 1]$ в сравнении с количеством межкластерных ребер. Значения, близкие к 1, означают высокую плотность связей внутри кластеров и малое число межкластерных связей. Значение $M = 0$ говорит о том, что разбиение сравнимо со случайным распределением связей. Фактически модулярность – это сумма весов ребер, инцидентных вершинам внутри кластера, за минусом суммы весов ребер вершин, связанных с вершинами других кластеров. Она служит основой для многих алгоритмов кластеризации, включая Лувенский алгоритм.

Лувенский метод нацелен на разбиение сети на кластеры таким образом, чтобы модулярность M , характеризующая плотность связей внутри кластеров по сравнению со связями между кластерами, была максимальной. Высокая модулярность означает, что узлы внутри одного кластера сильно взаимосвязаны, а между кластерами связи минимальны. Использовать понятие модулярности можно как для взвешенных, так и для невзвешенных графов.

В контексте ПБСС модулярность M используется для создания энергоэффективной топологии: кластеризация снижает количество дальних передач, что уменьшает расход энергии. Увеличивается вероятность успешной передачи данных внутри кластеров. Динамическое обновление кластеров на основе модулярности позволяет учитывать изменения в энергосостоянии узлов или потери узлов.

Предложенный модифицированный алгоритм Лувенского метода заключается в использовании некоторого порога δ , ограничивающего число кластеров в акватории.

Изначально, на первой итерации, будем считать, что если вероятность доставки сообщения между двумя узлами меньше 0.5, то эти узлы несвязаны, т.е. значение элемента матрицы смежности для них равно нулю.

Таким образом, в таблице маршрутизации прописаны начальная матрица смежности A и весовая матрица W :

$$a_{ij}(it=0) = 0, \text{ если } p(r_{ij}) < 0.5 ;$$

$$a_{ij}(it=0) = 1, \text{ если } p(r_{ij}) \geq 0.5 ;$$

$$w_{ij} = p(r_{ij}), \text{ для всех узлов сети.}$$

Далее, в модифицированном Лувенском методе при определении коэффициента модулярности предлагается учитывать только те ребра, веса которых больше заданного порогового значения δ . Ребра с меньшими весами, чем пороговое значение δ , не учитываются при кластеризации. В результате эта процедура повторяется итерационно, в зависимости от размеров получаемых кластеров. Пусть первоначально заданная величина порога равна $\delta=0.95$. Это значит, что если вес ребра меньше 0.95, то считается, что эти узлы не связаны, то есть соответствующий коэффициент в матрице

связности становится равен нулю. Если заданный порог δ не приводит к разбиению сети на оптимальное число кластеров (принятое в данной работе равным 5% от общего числа сенсоров в сети), то на следующей итерации он снижается на 0,05.

В данной работе будет проведено сравнение модифицированного Лувенского алгоритма с другими методами и исследование влияния выбранного порога δ на время жизни сети.

Если задать слишком высокий порог (в пределе $\delta=1.00$) то все сенсорами станут вырожденными кластерами и сеть окажется несвязанной. Если задать слишком низкий порог (менее 0.5) то почти все сенсоры в сети окажутся связанными и при равномерном покрытии акватории вся сеть будет одним большим кластером. Пример образования кластеров внутри сети при разных значениях порога δ приведен на рисунке 1.

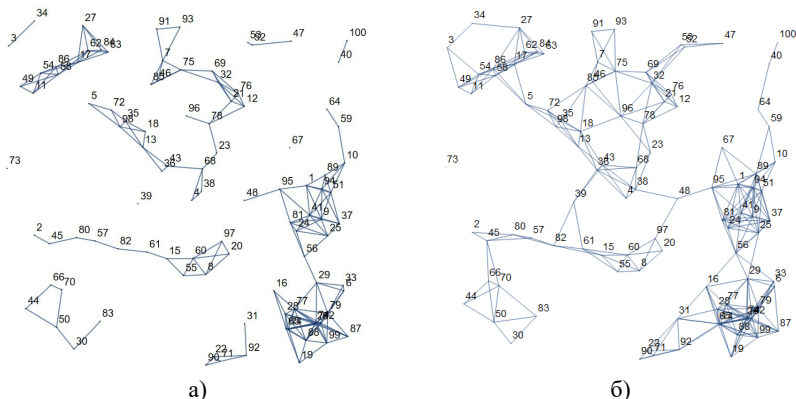


Рис. 1. Построение графа сети с учетом порога δ : а) порог $\delta=0.95$, б) порог $\delta=0.85$

В результате, если число соединенных кластеров начинается с одной вершины, то на каждом шаге алгоритма порог необходимо уменьшать на 0,05. Алгоритм заканчивает работу при достижении оптимального числа невырожденных кластеров.

Значение коэффициента модулярности M рассчитывается как:

$$M = \frac{1}{2m} \sum_{i,j} \left[a_{ij} w_{i,j} - \frac{w_i w_j}{2m} \right] \delta(c_i, c_j), \quad (1)$$

где a_{ij} – элемент матрицы смежности равный 1 или 0 на данной итерации, а $w_{i,j}$ – вес ребра между вершинами i и j , которые могут объединиться в кластер ($w_{i,j} > \delta$).

Величины w_i и w_j – это сумма весов ребер i и j , инцидентных вершинам i и j ($w_i > \delta, w_j > \delta$), на которых элементы матрицы смежности равны 1 на данной итерации, так что имеют место формулы:

$$\sum_k w_{jk} a_{ij} = w_j, \sum_l w_{il} a_{ij} = w_i, \quad (2)$$

где вершины k и l – это все узлы сети, связанные с i и j соответственно. Число m – это общее количество рёбер в сети, которые имеют веса больше заданного порога δ , так что отношение $\frac{w_i w_j}{2m}$ – это средневзвешенная вероятность передачи сигнала от узлов i и j по всей сети. Разность между величиной $a_{ij} w_{i,j}$ и этим отношением в формуле (1) показывает, что если вероятность доставки между узлами i и j выше, чем средневзвешенная вероятность передачи сигнала от узлов i и j по всей сети, то она положительна, и узлы имеет смысл сгруппировать в один кластер. Если эта разность отрицательна, то скорее всего узлы должны принадлежать разным кластерам.

$\delta(c_i, c_j)$ – это индикатор принадлежности узлов i и j к одному кластеру ($\delta(c_i, c_j) = 1$, если узлы принадлежат одному кластеру, и $\delta(c_i, c_j) = 0$ иначе). Данный индикатор учитывает только те рёбра, которые соединяют узлы внутри одного и того же кластера.

Коэффициент модулярности отдельного кластера определяется в модифицированном Лувенском методе как:

$$M_c = \frac{\sum_{i=1}^n w_i^c}{2m} \{c_i = c_j = c\} - \left(\frac{\sum_{i=1}^k (w_i^c + w_i^{c'})}{2m} \{c_i = c\} \right)^2, \quad (3)$$

$$M_c = \frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m} \right)^2,$$

где M_c – модулярность кластера, w_i^c – вес внутрикластерного ребра, больший выбранного порога ($w_i^c > \delta$), а $w_i^{c'}$ – вес межкластерного ребра, больший порога ($w_i^{c'} > \delta$).

Тогда сумма $\sum_{i=1}^n w_i^c = \Sigma_{in}$ – это сумма весов ребер между узлами внутри сообщества с (каждое ребро учитывается дважды). Увеличение величины Σ_{in} приводит к повышению модулярности, так как узлы становятся более "связанными" в рамках кластера. Сумма $\sum_{i=1}^k (w_i^c + w_i^{c'}) = \Sigma_{tot}$ – это сумма всех весов ребер для узлов внутри сообщества (включая ребра, соединяющие узлы сообщества с узлами за его пределами). Чем меньше величина Σ_{tot} , тем меньше "утечек" связей за пределы кластера, что также повышает модулярность.

Суммирование выполняется соответственно до значения n – числа внутрикластерных ребер и до k – общего числа ребер, инцидентных всем вершинам кластера.

Так как узлы в разных сообществах не вносят вклад в модулярность, общее значение модулярности можно записать как:

$$M = \sum_c M_c. \quad (4)$$

В алгоритме можно выделить основные стадии, которые группируются в два этапа:

На первой стадии выполняется поиск локальных кластеров минимальных размеров с оптимальным значением функции модулярности. Каждой вершине присваивается номер своего кластера и рассчитывается модулярность кластеров.

Выбирается произвольный кластер, из которого вершина переставляется по очереди в каждый соседний кластер. После перестановки подсчитывается изменение коэффициента модулярности ΔM до и после перестановки вершины и выбирается кластер с наибольшим изменением модулярности так, что $\Delta M > 0$, в котором закрепляется переставленная вершина. Если ни одна из перестановок не увеличивает модулярность, то вершина остается в том же кластере.

Данный пункт повторяется, пока изменения коэффициента не будут неизменными. Формула ΔM обеспечивает сравнение текущего вклада узла i в модулярность с его вкладом после переноса.

$$\Delta M = \left[\frac{\Sigma_{in} + w_i}{2m} - \left(\frac{\Sigma_{tot} + w_i}{2m} \right)^2 \right] - \left[\frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m} \right)^2 \right]. \quad (5)$$

Для каждого узла вычисляется, как его перемещение из текущего кластера в другой (или создание нового кластера) повлияет на значение модулярности ΔM . Узел переносится в тот кластер, который обеспечивает наибольшее увеличение модулярности. Если перемещение узла не улучшает модулярность (или ухудшает её), узел остаётся в своём текущем кластере.

Если узел i имеет сильные связи с другими узлами внутри кластера, перемещение узла в другой кластер снижает Σ_{in} и, как следствие, уменьшает модулярность. Если узел имеет более сильные связи с узлами другого кластера, перемещение увеличивает Σ_{in} в новом кластере и уменьшает связи Σ_{tot} с узлами текущего кластера, что повышает модулярность. Алгоритм находит оптимальный кластер для каждого узла путём итеративных проверок всех возможных перемещений.

Сжатие сети – это второй этап работы Лувенского алгоритма, следующий за локальной оптимизацией модулярности. Этот этап направлен на уменьшение размерности исходной сети путём формирования новой, более компактной сети. На этом этапе узлы, принадлежащие одному кластеру, объединяются в суперузлы.

Общая идея шагов на втором этапе следующая. Граф с кластерами преобразуется в мультиграф. Вершины кластеров объединяются в одну вершину с преобразованием ребер между ними в петли, а кратные ребра между полученными вершинами заменяются одним. Новые веса петель и ребер пересчитываются как среднее арифметическое из весов старых ребер. Эта процедура повторяется, пока не будет достигнута оптимальная величина коэффициента модулярности (когда изменений модулярности больше не будет происходить).

После сжатия сети уменьшается количество узлов и рёбер, что ускоряет последующие итерации алгоритма. Сложные связи между множеством узлов сводятся к простым связям между суперузлами. Это

помогает выявлять кластеры более высокого уровня, если они существуют.

После сжатия алгоритм снова начинает работу с первого этапа – локальной оптимизации, уже для новой, упрощённой сети. Итерации продолжаются до тех пор, пока модулярность не перестанет существенно изменяться. Это завершает работу Лувенского алгоритма, предоставляя финальную кластеризацию сети, рисунок 2.

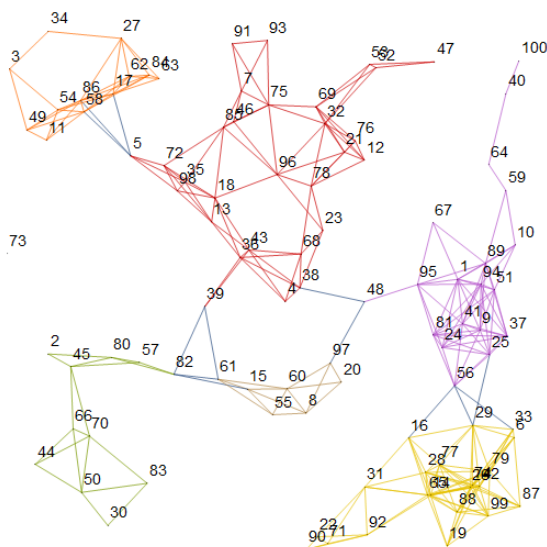


Рис. 2. Результат работы модифицированного Лувенского алгоритма.
Разбиение сети из 100 сенсоров на 5 кластеров

Выбор референсного узла. В каждом кластере нужно выбрать референсный узел. Для этого из общего графа сети выделим подграф, содержащий только узлы, принадлежащие одному кластеру. Для выбора референсного узла при первой кластеризации сети используется средневзвешенная вероятность передачи сигнала от узла i к узлу j , обозначаемая $w(b_j)$, которая показывает, сколько рёбер эффективно (с учетом вероятности доставки) соединяет узел i с другими узлами внутри кластера. Формула расчёта средневзвешенной вероятности передачи сигнала на узел i :

$$w(b_j \in C_l) = \frac{1}{2m_l} \sum_{b_i \in C_l} w_{ij}, \quad (6)$$

где w_{ij} – вероятность успешной передачи данных между узлами i и j , принадлежащими одному и тому же кластеру. Референсным узлом кластера при первой кластеризации назначается узел с наибольшей средневзвешенной вероятностью передачи сигнала:

$$b_{ref(C_l)} = \max_{b_j \in C_l} w(b_j \in C_l). \quad (7)$$

При разряде батарей сенсоров и появлении необходимости динамической рекластеризации сети, узлы с большим числом связей окажутся, очевидно, наиболее разряженными, поскольку через них будет проходить максимальный трафик. Это значит, что кроме связности в дальнейшем необходимо учитывать остаточную энергию сенсора. Обозначим ее как

$$\varepsilon_j = \frac{E_j}{E_0}, \quad (8)$$

E_0 (Дж) – начальный уровень заряда батареи сенсора, E_j (Дж) – текущее состояние заряда батареи, тогда формула расчёта средневзвешенной вероятности передачи сигнала на узел j при рекластеризации заменится на формулу средневзвешенной эффективности работы узла:

$$Eff_{node}(b_j \in C_l) = \frac{1}{2m_l} \sum_{b_i \in C_l} w_{ij} \varepsilon_j, \quad (9)$$

а референсным будет назначен узел

$$b_{ref(C_l)} = \max_{b_j \in C_l} Eff_{node}(b_j \in C_l). \quad (10)$$

Итак, в контексте подводных сенсорных сетей, Лувенский алгоритм используется для группировки сенсоров в кластеры, основываясь на следующих параметрах:

Узлы, расположенные близко друг к другу, объединяются в один кластер для минимизации энергозатрат на передачу данных.

Сенсоры с высокой остаточной энергией чаще выбираются в качестве референсных узлов (глав узлов).

Учитывается вероятность успешной передачи сообщений между узлами, зависящая от расстояния, мощности сигнала и условий среды.

5. Использование модифицированного алгоритма Дейкстры для нахождения оптимальных путей передачи сообщений в кластере. В условиях подводной среды требуется разработка маршрутов передачи сообщений внутри кластеров, которые обеспечат максимальную вероятность успешной доставки сообщений и минимизацию потерь.

Для достижения этой цели в данной модели предложено использование модифицированного алгоритма Дейкстры, адаптированного для поиска путей, которые максимизируют вероятность передачи сообщений, минимизируют энергозатраты и продлевают время жизни сети. Этот подход позволяет не только учитывать расстояние между узлами, но и адаптироваться к специфическим особенностям подводной среды.

В отличие от традиционных методов, фокусирующихся на минимизации расстояний или времени передачи, предложенный алгоритм оценивает вероятность доставки сообщения по каждому возможному маршруту с учетом остаточной энергии узлов. Для каждого сенсора рассчитывается не просто кратчайший путь, а маршрут, обеспечивающий максимальную вероятность передачи данных. Это позволяет значительно сократить количество потерянных сообщений и повысить общую производительность сети.

Алгоритм также учитывает динамические параметры сети, такие как расход энергии на узлах и ухудшение качества связи в определённых участках. Каждый маршрут между двумя сенсорами определяется через последовательность промежуточных узлов, причём выбор ретрансляторов происходит с учётом вероятностей передачи и уровня остаточной энергии. Такой подход обеспечивает гибкость и адаптивность сети, позволяя ей эффективно функционировать даже в условиях изменения окружающей среды и эксплуатационных параметров. В результате сеть становится более устойчивой, долговечной и способной к высококачественной передаче данных в подводных условиях.

В классическом варианте алгоритм Дейкстры [46] ищет кратчайший путь между узлами графа, минимизируя расстояние. В случае ПБСС ключевым фактором становится не минимизация

расстояния, а максимизация вероятности успешной передачи сообщений. Рассмотрим шаги модифицированного алгоритма.

Узлы сенсоров представляются вершинами графа $G = (B, L)$, где B – множество сенсоров, а L – рёбра, представляющие возможные пути связи между сенсорами.

Каждое ребро между узлами i и j имеет вес, определяемый вероятностью успешной передачи сообщения $w_{ij} = p(r_{ij})$, которая в общем случае зависит от расстояния между узлами, акустических характеристик канала и мощности передающего модема и уровнем остаточной энергии узла передатчика i и узла приемника j , так что формула расчёта средневзвешенной вероятности передачи сигнала заменится на формулу средневзвешенной эффективности передачи по ребру:

$$\varepsilon_i = \frac{E_i}{E_0}, \quad \varepsilon_j = \frac{E_j}{E_0}, \quad Eff_{route}(b_i, b_j) = w_{ij} \varepsilon_i \varepsilon_j, \quad (11)$$

где E_0 (Дж) – начальный уровень заряда батареи сенсора, E_i (Дж) и E_j (Дж) – текущее состояние заряда батареи.

Для каждого узла $b \in B$ вводится значение $D(b)$, отражающее текущую вероятность передачи сообщения от исходного узла до узла v . В начале $D(b) = 0$ для всех узлов, кроме исходного узла s , где $D(s) = 1$. С него и начинается работа алгоритма. Он добавляется в маршрут S . Для каждого узла u , соседнего с s , пересчитывается вероятность успешной передачи в соответствии с формулой:

$$D(u) = \max(D(u), D(s) \cdot Eff_{route}(b_s, b_u)), \quad (12)$$

где $Eff_{route}(b_s, b_u)$ – средневзвешенная эффективность передачи по ребру между узлами s и u . Среди всех узлов, не входящих в маршрут S , выбирается узел b с максимальным значением $D(b)$:

$$D(b) = \max_{u \notin S} D(u). \quad (13)$$

Этот узел добавляется в маршрут S . Алгоритм повторяет шаги до тех пор, пока не будут обработаны все узлы в графе или пока не будет найден путь до заданного узла. Построение маршрута возможно только внутри кластера.

В результате в каждом кластере $C_k \subset B$ оказываются построены оптимальные маршруты с наибольшей эффективностью передачи сообщений от референсного сенсора $b_{ref(C_k)}$ к остальным сенсорам в кластере. Определим функцию пути $\mathcal{R}(b_i, b_j)$ как путь между двумя сенсорами b_i и b_j через цепочку промежуточных сенсоров с вероятностью передачи:

$$Eff_{route}(b_i, b_j) = \prod_{k=1}^m Eff_{route}(b_k, b_{(k+1)}), \quad (14)$$

где m – количество промежуточных сенсоров на пути. Модифицированный алгоритм Дейкстры находит путь $\mathcal{R}(b_i, b_{ref(C_k)})$ с наибольшей эффективностью передачи сообщения для каждого ($b_i \in C_k$). Для каждого узла учитывается остаток энергии, что позволяет избегать маршрутов через узлы с критически низким уровнем заряда.

Вместо одного оптимального пути могут быть рассчитаны несколько маршрутов с высокой вероятностью передачи, что повышает надёжность сети. Алгоритм применяется локально для кластеров, что снижает вычислительные затраты и повышает масштабируемость сети.

После нахождения оптимальных путей передачи данных с помощью модифицированного алгоритма Дейкстры от каждого сенсора до референсного узла $b_{ref(C_k)}$ в кластере C_k , каждый сенсор получает одну из следующих ролей. Референсный узел $b_{ref(C_k)}$, это узел, который отвечает за сбор данных от всех остальных сенсоров в кластере. Сенсор ретранслятор b_i , это сенсор, который находится на пути передачи данных от других сенсоров к референсному. Сенсор b_i классифицируется как клиент, если он передает свои данные через ретранслятор, но не участвует в ретрансляции сообщений других узлов (крайний сенсор в цепочке передачи сообщений). Если путь от сенсора

b_i до референсного узла содержит только его самого и референсный узел, то он считается соло сенсором.

Ретрансляторы, клиенты и соло сенсоры являются обычными сенсорами, в отличие от референсных, которые собирают все сообщения и передают их на мобильный шлюз. Необходимо отметить, что сенсоры клиенты и соло обладают наибольшим запасом энергии, так как им требуется передать только одно свое сообщение.

Уровни ретрансляции определяют положение ретрансляторов на пути от обычных сенсоров до референсного узла. Для вычисления уровня ретрансляции необходимо рассчитать число промежуточных узлов между ретранслятором и референсным узлом. Присвоение уровня происходит в порядке от референсного узла к периферийным узлам.

6. Управление доступом к среде на основе расписания передач. Для управления доступом к среде в предлагаемой модели было использован один из наиболее распространенных методов управления доступом к среде в ПБСС – TDMA мультиплексирование с разделением по времени [47]. В TDMA каждому узлу назначается отдельный периодический временной интервал (слот), в течение которого он может передавать пакет без помех от других узлов в сети. Рассмотрим реализацию такого протокола. Итак, в активной фазе узлы бодрствуют в некоторых слотах и спят в течение оставшихся слотов, когда нет передачи или приема данных. Эта фаза разделена на несколько TDMA-кадров, каждый из которых состоит из нескольких слотов. Эти слоты зарезервированы путем присвоения каждому узлу своего номера. Другими словами, каждый номер представляет собой временное резервирование канала, оптимальное пространственное распределение номеров – это отдельная оптимизационная задача, которая решается за счет использования минимального количества номеров. Сенсоры с одинаковыми номерами могут одновременно передавать пакеты данных без каких-либо коллизий. Длина каждого слота фиксирована и равна задержке распространения сигнала в акватории плюс небольшое защитное время. Защитное время используется для обеспечения того, чтобы отдельные передачи не мешали друг другу.

На первом этапе временные слоты выделяются для сенсоров, работающих в режиме соло (без ретрансляторов). На следующем этапе временные слоты назначаются сенсорам-клиентам. Алгоритм принимает на вход список всех сенсоров. Затем происходит итеративный перебор сенсоров, пока все не будут распределены по

временным слотам. При распределении сенсора в слот проверяются следующие условия:

Отсутствие интерференции: сенсор может быть назначен в слот только в том случае, если он находится на таком расстоянии, чтобы не создавать помех сенсорам, уже распределённым в данный временной слот.

Уникальность адреса: в данном слоте не должно быть других сенсоров, отправляющих данные в тот же целевой сенсор.

Учет минимального номера слота: номер текущего слота должен быть больше минимального допустимого номера для рассматриваемого сенсора.

Если сенсор распределён во временной слот с номером i , то для сенсора, которому он отправляет данные, значение минимального номера слота обновляется. Новое значение будет на единицу больше текущего слота, если ранее заданное значение меньше.

В каждом TDMA-кадре каждый сенсор знает о своих собственных зарезервированных временных слотах в зависимости от его номера, а также о слотах, зарезервированных его соседними узлами. Таким образом, они могут запланировать активацию либо для передачи своего собственного пакета данных в течение зарезервированных слотов, либо для получения пакета данных от соседнего узла. Они спят в других оставшихся слотах, когда нет передачи или приема данных. Эта схема повторяется в течение каждого TDMA-кадра. Продолжительность этой фазы, т. е. количество TDMA-кадров, зависит от изменений топологии из-за смещения сенсора или истощения энергии. Для этой фазы следует рассматривать более короткую продолжительность в сценариях с быстрыми изменениями топологии и более длительный период в сценариях со стационарными или ограниченными мобильными узлами.

Длина каждого TDMA-кадра находится в обратной зависимости от скорости передачи данных, которая представлена в единицах пакетов в секунду. Чем выше скорость передачи данных, тем короче время цикла и, следовательно, меньше время ожидания.

Каждый TDMA-кадр разделен на количество N_{slot} слотов одинакового размера. Количество слотов можно рассчитать следующим образом:

$$N_{slot} = \frac{T_{frame}}{t_{slot}}, \quad (15)$$

T_{frame} – это максимальное время передачи всех сообщений в сети за один кадр, рассчитанное на основе наихудшего случая, когда каждый сенсор может иметь свой собственный слот для отправки сообщения (в таком случае количество слотов равно количеству сенсоров). В работе эта величина принята как константа, равная 7.8 сек.

Длина слота имеет вид:

$$t_{slot} = t_s + T_C + \tau \quad (16)$$

и складывается из времени, необходимого для отправки сообщения $t_s = 0.02$ с, задержки распространения сигнала $T_C = 2 \frac{r}{v_c}$ и τ – это некоторое защитное время, которое используется для обеспечения того, чтобы отдельные передачи не мешали друг другу. Каждому сенсору в течении одного кадра разрешено выходить в эфир только один раз (в одном из временных слотов). Время T_c пока сигнал распространяется в водной среде, сенсор не может ни отправлять другие сообщения, ни получать, а вынужден находиться в режиме ожидания.

7. Динамическая рекластеризация. Процесс динамической реконфигурации кластерной структуры в подводной беспроводной сенсорной сети играет решающую роль в поддержании её устойчивой и энергоэффективной работы в течение продолжительного времени. Сеть, состоящая из множества сенсорных узлов, сгруппированных в кластеры, сохраняет свою функциональность до тех пор, пока достаточное количество узлов обладает необходимым уровнем энергии для передачи сообщений. По мере того, как энергетические ресурсы отдельных узлов истощаются, становится необходимым пересмотреть структуру сети и перераспределить функциональные обязанности между оставшимися активными узлами.

На каждом этапе передачи данных осуществляется контроль текущего энергетического состояния узлов. Каждый сенсор b_i имеет уровень заряда E_i , и если он опускается ниже заданного порогового значения $E_i < E_{\min}$, узел считается неактивным и исключается из дальнейшей маршрутизации. Перестройка сети активируется в момент, когда хотя бы один узел достигает критического энергетического уровня $E_i < E_{\min}$.

В результате запускается механизм динамической адаптации сети, который включает в себя пересчёт вероятностей для новой кластеризации, обновление матриц связей между узлами, а также

формирование обновлённой топологии и пересмотр временного расписания кластеров.

Процесс передачи данных продолжается в новой конфигурации сети до тех пор, пока количество активных сенсоров B_{act} не упадет ниже определенного порогового значения, при котором сеть считается неработоспособной: $B_{act} < B_{min}^{act}$.

Перейдем теперь к этапу имитационного моделирования.

Опишем коммуникационную 3D архитектуру подводной беспроводной сенсорной сети, которая будет оптимизирована с использованием нового разработанного алгоритма.

8. Коммуникационная структура ПБСС в трёхмерном пространстве. В данной модели подводная беспроводная сенсорная сеть развёрнута в акватории с неоднородными физическими характеристиками среды и имеет трёхмерную пространственную организацию. Сеть включает в себя как обычные сенсоры, так и референсные узлы, размещённые в произвольной 3D-конфигурации. Для сбора и передачи данных используется мобильный шлюз, которым может выступать волновой глайдер или другое автономное устройство, способное перемещаться в водной среде, рисунок 3.

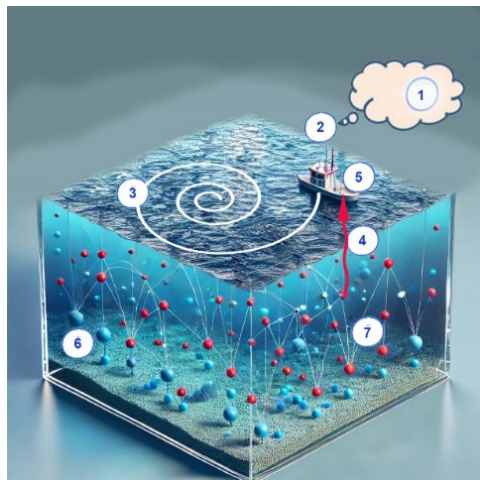


Рис. 3. Модель 3D подводной беспроводной сенсорной сети с одним мобильным шлюзом для произвольной 3D-топологии сети. 1 – центр обработки данных, 2 – линия беспроводной связи, 3 – траектория шлюза, 4 – мобильный шлюз, 5 – гидроакустический канал, 6 – обычные сенсоры, 7 – референсные сенсоры

Сеть организована в виде множества кластеров, каждый из которых содержит свой референсный узел, расположенный в определённой области и принимающий данные от обычных сенсоров, находящихся в пределах данного кластера. Обычные сенсорные узлы осуществляют измерения набора физических характеристик окружающей среды, проводят первичную обработку информации и передают её на референсный узел – либо напрямую, либо через промежуточных соседей по кластеру. После этого референсные узлы агрегируют полученные данные и отправляют их вверх по иерархии – на мобильный шлюз.

В рамках имитационного моделирования функционирования ПБСС используется трёхмерное представление подводной акватории. Пространство моделируется в виде объёма с площадью $S = n \times n$ в горизонтальной проекции и глубиной h . Сенсоры b_i размещаются в этом объёме либо равномерно, либо в соответствии с заданной схемой, зависящей от специфики наблюдаемой среды. Такое размещение может учитывать, например, температурные градиенты, а также особенности рельефа дна.

Каждый сенсор характеризуется координатами в пространстве $\vec{x}_i = (x_i, y_i, z_i)$, где x_i, y_i задают положение в горизонтальных слоях, а z_i – глубину размещения сенсора по вертикали. Эти координаты определяют начальную топологию сети. Распределение узлов может быть равномерным или следовать заданной модели. Предлагается использовать:

Случайное распределение, при котором в плоскости $x_i, y_i \sim U(0, n)$, а по вертикали $z_i = z_0$. Регулярную модель, которая предполагает, что сенсоры располагаются в узлах регулярной решетки с шагом, равномерно покрывающей всю область: $x_i = i \cdot s, y_j = j \cdot s, z = z_0$. Регулярную решётку с отклонением:

$$x_i = i \cdot s + \epsilon_x, y_j = j \cdot s + \epsilon_y, z = z_0, \quad (17)$$

$$\epsilon_x, \epsilon_y \sim U(-\Delta, \Delta),$$

где Δ задаёт амплитуду отклонений.

Псевдослучайное распределение, созданное на основе последовательности Халтона: $(x_i, y_i) = \text{Halton}(b_1, b_2, i)$, где b_1, b_2, i –

взаимно простые основания последовательности Халтона, i – индекс сенсора.

Четыре предлагаемые модели распределения предоставляют гибкость в выборе начальной топологии сети в зависимости от задач и требований к покрытию акватории. Данный выбор топологий обоснован в первую очередь учетом особенностей реальной эксплуатации сети, где объекты не могут располагаться в точно заданных позициях (процесс сброса сенсоров всегда осуществляется с некоторой неопределенностью, связанной с не одномерной скоростью хода судна обеспечения, морским волнением и человеческим фактором).

Все сенсоры разделяются на набор кластеров, каждый из которых состоит из группы узлов, связанных друг с другом. В рамках кластера информация передается к центральному референсному узлу и далее на мобильный шлюз.

Модель коммуникации может учитывать стохастическую природу подводной среды, где вероятность успешной передачи сообщений определяется как стационарными параметрами, заданными в модели среды, так и изменяющимися условиями, включая временные и пространственные колебания параметров.

9. Модель подводной среды. Коммуникационная модель описывает процесс обмена сообщениями между сенсорами, где вероятность успешной передачи сигнала $p(r_{ij})$ между двумя узлами b_i и b_j определяется условиями среды и расстоянием r_{ij} между ними. С увеличением этого расстояния, а также при неблагоприятных параметрах среды, вероятность успешной доставки уменьшается. На результат передачи влияют характеристики используемого модема, свойства среды и особенности распространения сигнала. Поэтому модель среды должна учитывать такие важные факторы, как затухание акустических волн, многолучевое распространение, шумы и прочие физические особенности океанической среды.

Одним из ключевых параметров, характеризующих качество подводной связи, является вероятность ошибки на бит, и служит важным показателем надёжности передачи информации в условиях воздействия аддитивного белого гауссовского шума [48].

Следуя [49], считаем, что температура, солёность и кислотность постоянны, тогда коэффициент затухания сигнала в среде при заданной частоте звуковых колебаний можно оценить по обобщенной формуле Торпа [49].

В подводной среде, где распространение сигнала сопровождается значительным затуханием, шумами и многолучевыми

эффектами, одним из ключевых факторов, влияющих на качество связи, является битовая ошибка (BER). Важно учитывать это при использовании двоичной фазовой манипуляции (BPSK), одного из наиболее надёжных методов модуляции в шумной среде. Если сигнал распространяется в канале с рэлеевским замиранием, то есть с эффектом случайных глубоких провалов амплитуды сигнала, вызванных многолучевыми отражениями, то битовая ошибка между передающим и принимающим устройствами может быть рассчитана, следуя [50].

Чем выше BER – тем ниже вероятность того, что пакет дойдёт до адресата без ошибок. Явная связь между физическими параметрами среды и вероятностью успешной доставки данных, и, следовательно, с надёжностью всей системы приведена авторами в работах [51 – 53].

10. Проектные параметры моделирования функциональных характеристик ПБСС. Технические параметры модема, такие как мощность модема P_{bi} , несущая частота f и полоса пропускания используются в модели среды, позволяя рассчитать вероятность успешной доставки сообщения вдоль заданного ребра, соединяющего два сенсора. Модель позволяет использовать и другие технические параметры, такие как характеристики (например, направленность) антенны. В предлагаемой работе антенна считается ненаправленной. В дальнейшем в имитационной модели использованы следующие проектные параметры. Частота работы модема принята равной $f = 60$ кГц с полосой частот $B = 30$ кГц, битовая скорость – $f_{bit} = 12,8$ кбит/с. Считается, что все передаваемые сообщения s имеют одинаковую длину $N_{bit} = 256$ бит. Суммарное время передачи сообщения и получения подтверждения о доставке определяется как $t_s = N_{bit} / f_{bit} = 0.02$ с.

Энергетическое состояние E_i сенсора, это уровень оставшейся энергии сенсора, уменьшающийся при каждой передаче, приёме сообщений и в режиме ожидания. Энергия сенсоров – ключевой параметр, определяющий срок службы сети. В частности, выполняя рекластеризацию и выбирая новые референсные узлы исходя из их энергетического состояния E_i , можно добиться существенного увеличения времени жизни сети. Энергетические характеристики модемов приняты следующими: $P_s = 25$ Вт – максимальная излучаемая мощность передающего модема; $E_s = P_s \cdot t_s = 0.5$ Дж – энергия, необходимая для отправки одного сообщения; $P_w = 0.3$ Вт – мощность,

затраченная на ожидание и прием сообщения, эмпирическая цифра для процессорной системы, ожидающей прием сообщения; $P_{inf} \approx P_w$ – мощность, затраченная на сбор информации. Это средняя мощность работающей процессорной системы с малым потреблением энергии; $E_0 = 864$ кДж – емкость батареи напряжением 12В.

Предполагается, что время работы сети между двумя измерениями равно T . За это время все результаты измерений должны быть доставлены в мобильный шлюз.

Для определения оптимальной архитектуры сети используется разработанная авторами математическая модель, основанная на вероятностном подходе и критериях оптимальности функционирования ПБСС с точки зрения связности. При этом учитывается ограничение на максимальное число перепосылок, необходимых для успешной доставки сообщений.

Не все попытки передать сообщение от сенсорного узла соседу могут оказаться успешными. Если допустить возможность многократных попыток передачи, осуществляемых через определенные интервалы времени, то доставка сообщения от сенсорного устройства к референсному узлу будет случайной величиной, зависящей от вероятности связности сети. Будем считать, что, если после фиксированного числа передач ближайшим соседям сообщение не доставлено, оно считается потерянным и сенсор переходит к пересылке следующего стоящего в очереди сообщения.

11. Вычислительные эксперименты. Имитационный комплекс моделирования разработан на основе языка программирования Python с использованием дополнительных библиотек. На первом этапе было проведено исследование зависимости исходного количества кластеров от выбранного порога кластеризации δ , частоты передачи сигнала и расстояний между сенсорами в акватории для 3D сети из 100 сенсоров.

Первый этап показал, что при регулярном или близком к регулярным расположениям сенсоров введение порога δ кластеризации незначительно меняет число кластеров в сети, и оно сохраняется в пределах близких к 5% от общего количества. При увеличении частоты сигнала при одинаковых расстояниях количество кластеров увеличивается, что обусловлено большей потерей данных на больших частотах.

На следующем этапе было проведено сравнительное моделирование разработанного алгоритма с алгоритмами LEACH, ICA, PPWURC и др. [51]. Входные параметры моделирования представлены в таблице 1.

Таблица 1. Входные параметры моделирования

Проектные параметры сети	Числовые характеристики
Размер акватории (м)	150×150×150
Начальная энергия (Дж)	40
Частота передачи сигнала (кГц)	40
Количество сенсоров (шт)	30

Рассматривался вариант случайного распределения сенсоров в акватории. Результаты согласно параметрам из таблицы 2 представлены на рисунке 4.

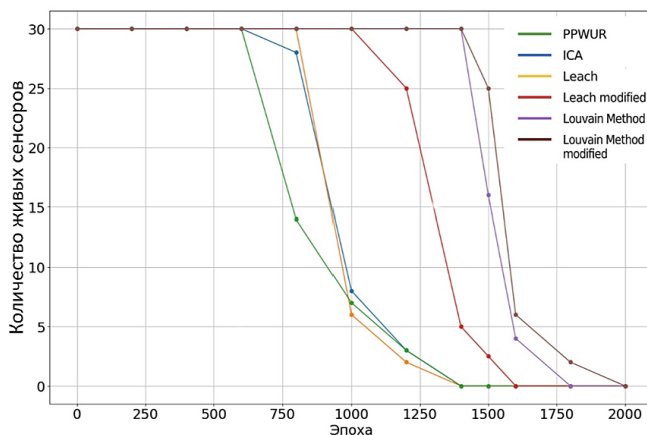


Рис. 4. Зависимость количества живых сенсоров от количества эпох для различных алгоритмов

Второй эксперимент показал значительное увеличение времени автономной работы сети при применении Лувенского алгоритма по сравнению с другими алгоритмами. Для конфигурации с 30 сенсорами начало отказа работы сети (первый выход из строя узла) наблюдается лишь после 1400-й эпохи, тогда как при использовании других методов кластеризации сенсоры начинают выходить из строя уже к 1000-й эпохе или ранее. На представленном графике также видно, что использование порогового значения 0.95 в модифицированной версии Лувенского алгоритма позволяет дополнительно замедлить темпы разряда сенсоров примерно на 3% по сравнению с базовой версией алгоритма, за счет небольшого увеличения общего числа кластеров и тем самым обеспечивая более плавное снижение количества рабочих узлов.

Следующий этап представляет собой анализ влияния различных значений порогового ограничения кластеризации в Лувенском алгоритме на время жизни подводной сенсорной сети. В рамках исследования рассматривались пороговые значения δ , равные 0.95, 0.8, 0.7, 0.5, а также вариант без порогового ограничения. В качестве тестовой области была выбрана акватория размером $10\,000 \times 10\,000$ метров. Сеть имела пространственную структуру: сенсоры были размещены на трёх фиксированных уровнях глубины – 100, 300 и 500 метров. В каждом слое располагалось по 100 сенсоров, размещённых на основе регулярной ортогональной решётки с заданным случайным отклонением координат. Критерием завершения работы сети считался выход из строя 200 сенсоров. Пример рассматриваемой топологии сенсорной сети приведен на рисунке 5.

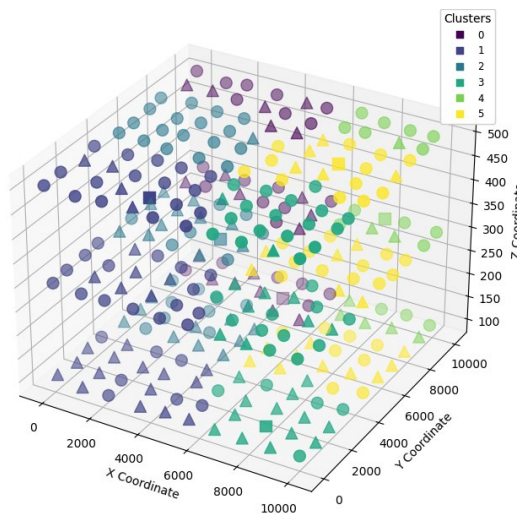
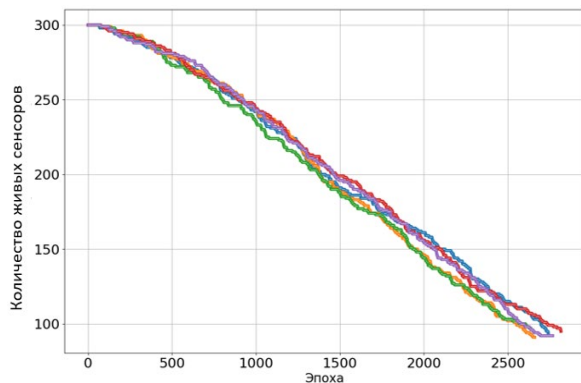


Рис. 5. Результат построения регулярной решетки ПБСС со случайными отклонениями и ее топологической кластеризации (треугольники – сенсоры ретрансляторы; квадраты – сенсоры референсы)

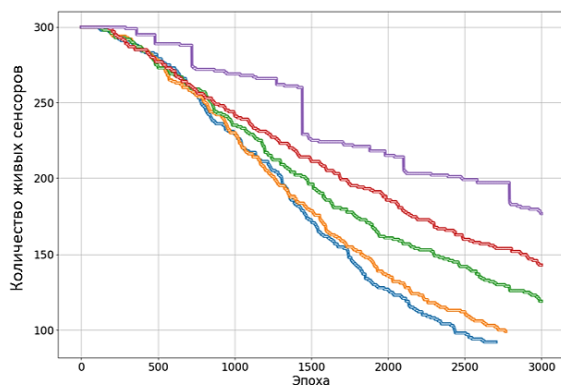
Результаты работы сети на частоте 30 кГц и 70 кГц представлены на рисунке 6.

Из сравнения графиков на рисунке 6 видно, что работа сети на частоте 30 кГц при выбранной топологии оказывается практически не чувствительна к изменению порога кластеризации. Это связано с тем, что выбранные проектные параметры обеспечивают надежную передачу сигнала и увеличение порога не приводит к увеличению

числа кластеров. При 70 кГц мы наблюдаем другую картину – увеличение порога приводит к невозможности сенсоров объединиться в кластеры, сеть распадается на большое число кластеров. Это приводит к увеличению времени жизни сети, но одновременно нарушает ограничение, наложенное на число референсных агентов (кластеров) в сети.



а)



б)

Рис. 6. Зависимость количества живых сенсоров от времени работы сети:

а) – при работе на частоте 30 кГц, б) – при работе на частоте 70 кГц.

Фиолетовая линия – $\delta=0.95$, красная – $\delta=0.8$, зеленая – $\delta=0.7$,

оранжевая – $\delta=0.5$, синяя – без порога

При пороге $\delta=0.95$ время работы ПБСС на частоте 70 кГц ПБСС имеет ступенчатую кривую разряда. Это связано с тем, что при данном

значении порога ПБСС обладает большим начальным количеством кластеров при каждой рекластеризации. По достижении 3000 эпох, количество рабочих сенсоров для $\delta=0.8$ составляет 148, для $\delta=0.7$ количество рабочих сенсоров составляет 125, в то время как для значений $\delta=0.5$ и $\delta=0.0$ (без порога) сеть прекратила работу на 2750 эпохе и 2680 эпохе соответственно.

Результаты сопоставления времени разряда первого сенсора в ПБСС и начального количества кластеров при различных значениях порога δ представлены на рисунке 7.

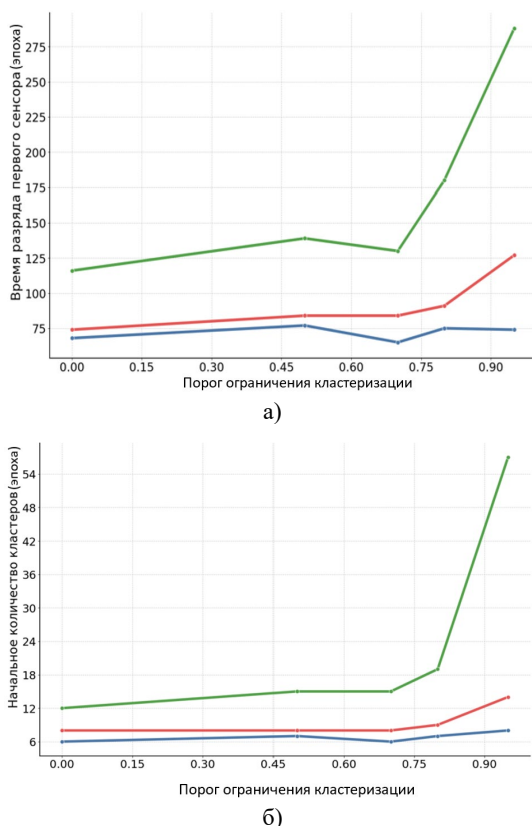


Рис. 7. Результаты работы сети: а) – время разряда первого сенсора; б) – начальное количество кластеров, синяя линия – 30 кГц, красная линия – 50 кГц, зеленая линия – 70 кГц

Результаты, представленные на рисунке 7 наглядно показывают, что при выбранных проектных параметрах сети изменение порога кластеризации от 0 до 0.7 практически не меняет число кластеров и, следовательно не влияет на время жизни сети вне зависимости от частоты передачи сигнала. Для частоты 30 кГц сеть будет разбита на 6 кластеров, для частоты 50 кГц – на 7-8 кластеров, а при частоте 70 кГц – от 12 до 15 кластеров. При пороге выше 0.7 картина меняется – число кластеров начинает расти и к порогу 0.9 достигает для частоты 30 кГц 7-8 кластеров, для частоты 50 кГц – 10-12 кластеров. Однако для 70 кГц мы наблюдаем резкий рост числа кластеров и общее количество достигает 54 кластеров, что не удовлетворяет исходной постановке задачи. Следовательно, несмотря на резкое увеличение времени жизни сети, такой порог не подходит для гибридных ПБСС с мобильным шлюзом.

Из анализа, проведенного выше, можно сделать вывод, что заданный набор проектных параметров, в частности выбранная несущая частота, требует индивидуальной настройки работы алгоритма.

При частоте 70 кГц и пороге $\delta=0.95$ разряд первого сенсора ПБСС происходит с 275 эпохи, что существенно превышает начальный момент разряда ПБСС при частотах 50 и 30 кГц и том же значении порога δ . Аналогичная картина наблюдается и на графике начального количества сенсоров ПБСС (начальное количество сенсоров при пороге $\delta=0.95$ равно 55 (что составляет 18% от всей ПБСС)).

Дополнительно был проведен анализ работы сети при разных расстояниях между сенсорами. Рассматривались расстояния между сенсорами от 800 до 2000 м. Результаты оценки времени жизни сети от расстояния между сенсорами, при разных частотах передачи данных и разных порогах ограничения кластеризации представлены на рисунке 8.

Модель среды, определяемая формулами (2)-(4) характерна водопадоподобным видом снижения вероятности успешной передачи сигнала на определенных расстояниях, зависящих от несущей частоты. На этих же расстояниях на рисунке 9 мы можем наблюдать резкое увеличение времени жизни сети, связанное с увеличением числа кластеров выше допустимого.

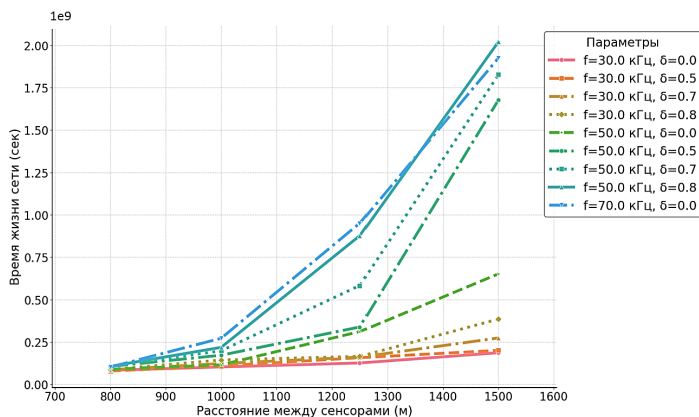


Рис. 8. Зависимость времени жизни сети от расстояния между сенсорами

На рисунке 9 показана зависимость времени жизни сети от порога кластеризации при разных расстояниях между сенсорами в акватории.

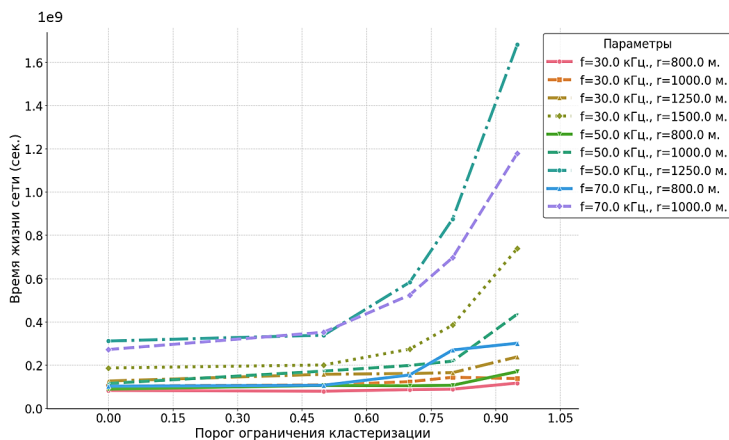


Рис. 9. Зависимость времени жизни сети от порога ограничения кластеризации

На рисунке 9, также как выше на рисунке 7(а) мы наблюдаем увеличение времени жизни сети после превышения порога 0.7 вследствие резкого увеличения общего числа кластеров.

При данных характеристиках ПБСС обладает большим количеством кластеров, в следствии чего время работы сети возрастает.

Исследуем зависимость количества кластеров в сети от порога кластеризации. На рисунке 10 показана динамика числа кластеров при различных частотах в зависимости от выбранного порога кластеризации.

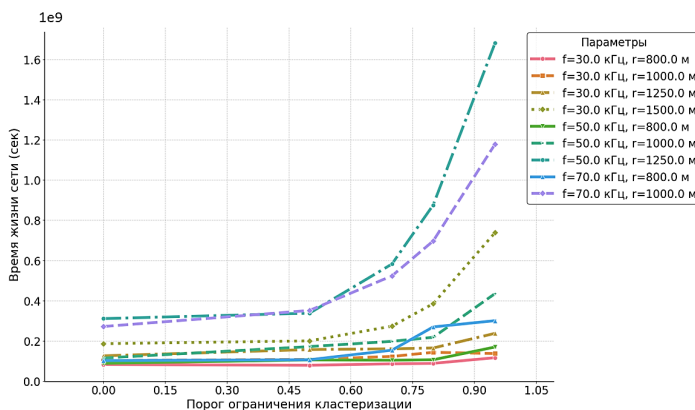


Рис. 10. Зависимость начального количества кластеров от порога ограничения кластеризации

Как видно из рисунка 10 количество кластеров остается в допустимом диапазоне (около 5% от общего количества сенсоров сети) при значениях частоты 30 кГц и 50 кГц, что согласуется с предыдущими графиками.

Таким образом, полученные результаты имитационного моделирования указывают на то, что оптимальным значением порога δ , обеспечивающего надежную работы гибридной ПБСС являются значения от 0.5 до 0.8 при работе сети на частоте 30 кГц и 50 кГц.

12. Заключение. Предлагаемая в данной работе имитационная модель, основанная на модификации таких известных алгоритмов, как Лувенский алгоритм и алгоритм Дейкстры эффективно решает ряд ключевых задач, связанных с управлением ресурсами в подводных беспроводных сенсорных сетях (ПБСС). Во-первых, предложенный алгоритм маршрутизации повышает надёжность передачи данных за счёт оптимизации процессов пересылки, что значительно снижает количество потерянных сообщений. Во-вторых, достигается существенная экономия энергии путём равномерного распределения нагрузки между узлами и рационального использования ретрансляторов. Реализованный протокол связи на основе TDMA позволяет осуществлять параллельную передачу данных в различных

временных слотах, что уменьшает вероятность помех, сокращает общее время работы сети и, соответственно, снижает энергозатраты. Механизм динамической рекластеризации дополнительно повышает устойчивость сети и продлевает срок её автономного функционирования за счёт адаптации к изменениям уровня энергии и структуры сети.

Использование модифицированного Лувенского алгоритма с введением порога кластеризации в подводных сетях оказалось эффективным по следующим причинам. Результаты серии экспериментов показали, что время жизни гибридной подводной сенсорной сети (ПБСС) с мобильным шлюзом существенно зависят от параметров кластеризации, в частности – от порогового значения δ , используемого в Лувенском алгоритме. Если время обхода референсных узлов мобильным шлюзом является критическим параметром (что верно для сетей с высокой частотой съема информации), то предпочтительно выбирать частоты в интервале 30-50 кГц и высокий, порядка $\delta=0.8$ порог кластеризации. Это обеспечит необходимое количество кластеров с надежными связями внутри каждого кластера.

Если предполагается использование несущей частоты более 70 кГц, то в сетях с высокой частотой съема информации предпочтительнее использовать классический лувенский алгоритм, чтобы избежать образования большого числа кластеров. При низкой частоте съема информации возможно использование порога вплоть до $\delta=0.95$, это увеличит время жизни сети, но приведет к возникновению очень большого числа кластеров.

Сочетание модифицированных Лувенского алгоритма и метода Дейкстры в рамках предложенной модели позволило создать устойчивую, адаптивную и энергоэффективную кластерную архитектуру для ПБСС. Сравнительный анализ результатов моделирования показывает преимущество предложенного подхода по таким показателям, как увеличение времени автономной работы сети и повышение надёжности доставки данных по сравнению с решениями, предложенными в других исследованиях.

Таким образом, полученные в работе результаты подтверждают высокую эффективность предложенной модели как с теоретической, так и с практической точек зрения. Модифицированный Лувенский алгоритм с адаптируемым порогом кластеризации, интегрированный с маршрутизацией на основе алгоритма Дейкстры, позволяет гибко адаптировать сетевую топологию к различным условиям среды и режимам эксплуатации. Применение данной модели обеспечивает не

только надёжную передачу данных и оптимальное распределение нагрузки, но и позволяет учитывать физические ограничения подводной среды и особенности проектных требований. Это делает модель перспективной основой для дальнейшей разработки интеллектуальных систем управления ресурсами в подводных сенсорных сетях, ориентированных на длительное автономное функционирование и высокую устойчивость в сложных гидроакустических условиях.

Литература

1. Zorzi M., Casari P., Baldo N., Harris A.F. Energy-efficient routing schemes for underwater acoustic networks // IEEE J. Sel. Areas Commun. 2008. vol. 26. no. 9. pp. 1754–1766.
2. Davis A., Chang H. Underwater wireless sensor networks // Proceedings of the IEEE Oceans. 2012. pp. 1–5.
3. Freitag L., Grund M., Singh S., Partan J., Koski P., Ball K. The WHOI micro-modem: An acoustic communications and navigation system for multiple platforms // Proceedings Oceans MTS/IEEE. 2005. vol. 2. pp. 1086–1092.
4. Rahman M.A., Lee Y., Koo I. EECOR: an energy efficient cooperative opportunistic routing protocol for underwater acoustic sensor networks // IEEE Access. 2017. vol. 5. pp. 14119–14132. DOI: 10.1109/ACCESS.2017.2730233.
5. Akyildiz I.F., Pompili D., Melodia T. Underwater acoustic sensor networks: Research challenges // Ad Hoc Netw. 2005. vol. 3(3). pp. 257–279. DOI: 10.1016/j.adhoc.2005.01.004.
6. Pompili D., Melodia T., Akyildiz I.F. Distributed routing algorithms for underwater acoustic sensor networks // IEEE Trans. Wireless Commun. 2010. vol. 9. no. 9. pp. 2934–2944. DOI: 10.1109/TWC.2010.070910.100145.
7. Lu Q., Liu F., Zhang Y., Jiang S. Routing Protocols for Underwater Acoustic Sensor Networks: A Survey from an Application Perspective // Advances in Underwater Acoustics. 2017. DOI: 10.5772/intechopen.68900.
8. Azam I., Majid A., Khan T., et al. Avoiding Energy Holes in Underwater Wireless Sensor Networks with Balanced Load Distribution // 10th International Conference on Complex, Intelligent and Software Intensive System (CISIS). 2016. pp. 341–350. DOI: 10.1109/CISIS.2016.109.
9. Chaaf A., Muthanna M.S.A., Muthanna A., et al. Energy-Efficient Relay-Based Void Hole Prevention and Repair in Clustered Multi-AUV Underwater Wireless Sensor Network // Security and Communication Networks. 2021. vol. 2021. no. 1. DOI: 10.1155/2021/9969605.
10. Khan Z.A., Karim O.A., Abbas S., Javaid N., Zikria Y.B., Tariq U. Q-learning based energy-efficient and void avoidance routing protocol for underwater acoustic sensor networks // Computer Networks. 2021. vol. 197. DOI: 10.1016/j.comnet.2021.10839.
11. Mhemed R., Phillips W., Comeau F., Aslam N. Void avoiding opportunistic routing protocols for underwater wireless sensor networks: A survey // Sensors. 2022. vol. 22(23). DOI: 10.3390/s22239525.
12. Tokmachev D.A., Chensky A.G., Zolotarev N.S., Poletaev A.S. Modelling of a hydroacoustic modem for underwater communications // J. Phys.: Conf. Ser. 2021. vol. 1728. DOI: 10.1088/1742-6596/1728/1/012021.

13. Chen H., Zhu Y., Zhang W., Wu K., Yuan F. Underwater Acoustic Micromodem for Underwater Internet of Things // *Wireless Communications and Mobile Computing*. 2022. vol. 2022. no. 1. DOI: 10.1155/2022/9148756.
14. Sherlock B., Morozs N., Neasham J., Mitchell P. Ultra-low cost and ultra-low-power, miniature acoustic modems using multipath tolerant spread spectrum techniques // *Electronics*. 2022. vol. 11. no. 9. DOI: 10.3390/electronics11091446.
15. Li J., Tong F., Jiang W., Roy S. Underwater acoustic modem with high precision ranging capability based on second-level cesium atomic clock // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2022. vol. XLVI-3/W1-2022. pp. 75–80, DOI: 10.5194/isprs-archives-XLVI-3-W1-2022-75-2022.
16. Zia I., Poncela J., Otero P. State-of-the-Art Underwater Acoustic Communication Modems: Classifications, Analyses and Design Challenges // *Wireless Personal Communications*. 2021. vol. 116. pp. 1325–1360. DOI: 10.1007/s11277-020-07431-x.
17. Datta A., Dasgupta M. Energy efficient topology control in underwater wireless sensor Networks. *Computers and Electrical Engineering*. 2023. vol. 105. DOI: 10.1016/j.compeleceng.2022.108485.
18. Sutagundar A.V., Halakarnimath B.S., Bhadjantri L.B. et al. A Novel Approach of topology control in underwater sensor networks // *IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*. 2022. pp. 1–5. DOI: 10.1109/MysuruCon55714.2022.9972400.
19. Choudhary M., Goyal N. Dynamic topology control algorithm for node deployment in mobile underwater wireless sensor networks // *Concurrency and Computation Practice and Experience*. 2022. vol. 34(15). DOI: 10.1002/cpe.6942.
20. Hu Y., Hu K., Liu H. et al. An energy-balanced head nodes selection scheme for underwater mobile sensor networks // *EURASIP Journal on Wireless Communications and Networking*. 2022. no. 63. DOI: 10.1186/s13638-022-02141-3.
21. Fedorova T.A., Ryzhov V.A., Semenov N.N. and Sulaiman Sh.A. Optimization of an Underwater Wireless Sensor Network Architecture with Wave Glider as a Mobile Gateway // *Journal of Marine Science and Application*. 2022. vol. 21. pp. 179–196. DOI: 10.1007/s11804-022-00268-9.
22. Ahmad I., Rahman T., Zeb A., et al. Cooperative Energy-Efficient Routing Protocol for Underwater Wireless Sensor Networks // *Sensors*. 2022. vol. 22. DOI: 10.3390/s22186945.
23. Bharny S., Sharma S., Alsharabi N., Eldin E.T., Ghamry N.A. Energy-efficient clustering protocol for underwater wireless sensor networks using optimized glowworm swarm optimization // *Front. Mar. Sci.* 2023. vol. 10. DOI: 10.3389/fmars.2023.1117787.
24. Sheeja R., Iqbal M.M., Sivasankar C. Multi-objective-derived energy efficient routing in wireless sensor network using adaptive black hole-tuna swarm optimization strategy // *Ad Hoc Networks*. 2023. vol. 144. DOI: 10.1016/j.adhoc.2023.103140.
25. Alfajeer A.A., Harous S. A review of routing protocols for underwater wireless sensor networks // *The 3rd International Conference on Distributed Sensing and Intelligent Systems (ICDSIS 2022)*. 2022. pp. 138–145. DOI: 10.1049/icp.2022.2428.
26. Shovon I.I., Shin S. Survey on Multi-Path Routing Protocols of Underwater Wireless Sensor Networks: Advancement and Applications. *Electronics*. 2022. vol. 11(21). DOI: 10.3390/electronics11213467.
27. Kebkal A., Kebkal K., Komar M. Data-link protocol for underwater acoustic networks // *Europe Oceans*. 2005. vol. 2. pp. 1174–1180. DOI: 10.1109/OCEANSE.2005.1513225.
28. Stojanovic M. Optimization of a data link protocol for an underwater acoustic channel // *Europe Oceans*. 2005. vol. 1. pp. 68–73 DOI: 10.1109/OCEANSE.2005.1511686.

29. Guqhaiman A.A., Akanbi O., Aljeadi A., Chow C.E. A Survey on MAC Protocol Approaches for Underwater Wireless Sensor Networks // *IEEE Sensors Journal*. 2021. vol. 21(3). pp. 3916–3932. DOI: 10.1109/JSEN.2020.3024995.
30. Kulla E., Matsuo K., Barolli L. MAC Layer Protocols for Underwater Acoustic Sensor Networks: A Survey // *Lecture Notes in Network and Systems (LNNS)*. 2022. vol. 496. DOI: 10.1007/978-3-031-08819-3_21.
31. Gazi F., Ahmed N., Misra S., Wei W. Reinforcement Learning-Based MAC Protocol for Underwater Multimedia Sensor Networks // *CM Transactions on Sensor Networks*. 2022. vol. 18(3). pp. 1–25. DOI: 10.1145/3484201.
32. Liu X., Du X., Li M., Wang L., Li C. A MAC Protocol of Concurrent Scheduling Based on Spatial Temporal Uncertainty for Underwater Sensor Networks // *Hindawi Journal of Sensors*. 2021. vol. 2021. DOI: 10.1155/2021/5558078.
33. Shahapur S.S., Khanai R. Underwater Sensor Network at physical, data link and network layer – a survey // *International Conference on Communications and Signal Processing (ICCSP)*. 2015. pp. 1449–1453. DOI: 10.1109/ICCSP.2015.7322753.
34. Songzuo L., Iqbal B., Khan I.U., Ahmed N., Qiao G., Zhou F. Full Duplex Physical and MAC Layer-Based Underwater Wireless Communication Systems and Protocols: Opportunities, Challenges, and Future Directions // *J. Mar. Sci. Eng.* 2021. vol. 9(5). DOI: 10.3390/jmse9050468.
35. Casari P., Ardizzon F., Tomasin S. Physical Layer Authentication in Underwater Acoustic Networks with Mobile Devices // *WUWNet '22: Proceedings of the 16th International Conference on Underwater Networks & Systems*. 2022. pp. 1–8. DOI: 10.1145/3567600.3567604.
36. Kaveripakam S., Chinthaginjala R. Energy balanced reliable and effective clustering for underwater wireless sensor networks // *Alexandria Engineering Journal*. 2023. vol. 77. pp. 41–62. DOI: 10.1016/j.aej.2023.06.083.
37. Zhou Z., Peng Z., Cui J., Jiang Z. Handling triple hidden terminal problems for multichannel MAC in long-delay underwater sensor networks // *IEEE Trans. Mob. Comput.* 2012. vol. 11(1). pp. 139–154. DOI: 10.1109/TMC.2011.28.
38. Yahya A., Islam S.U., Zahid M., Ahmed G., Raza M., Pervaiz H. Cooperative routing for energy efficient underwater wireless sensor networks // *IEEE Access*. 2019. vol. 7. pp. 141888–141899. DOI: 10.1109/ACCESS.2019.2941422.
39. Rayen S.J., Arunajsmine J., Neelakandan S., Annamalai R. Social media networks owing to disruptions for effective learning // *Procedia Comput. Sci.* 2020. vol. 172. pp. 145–151. DOI: 10.1016/j.procs.2020.05.022.
40. Li P., Wang S., Zhang H., Zhang E. Improved particle swarm optimization algorithm of clustering in underwater acoustic sensor networks // *Proceedings of the IEEE OCEANS*. 2017. pp. 1–7. DOI: 10.1109/OCEANSE.2017.8084869.
41. Yu W., Chen Y., Wan L., Zhang X., Zhu P., Xu X. An energy optimization clustering scheme for multi-hop underwater acoustic cooperative sensor networks // *IEEE Access*. 2020. vol. 8. pp. 89171–89184. DOI: 10.1109/ACCESS.2020.2993544.
42. Wang M., Chen Y., Sun X., Xiao F., Xu X. Node energy consumption balanced multi-hop transmission for underwater acoustic sensor networks based on clustering algorithm // *IEEE Access*. 2020. vol. 8. pp. 191231–191241. DOI: 10.1109/ACCESS.2020.3032019.
43. Zhang W., Wang J., Han G. et al. A cluster sleep-wake scheduling algorithm based on 3D topology control in underwater sensor networks // *Sensors*. 2019. vol. 19. DOI: 10.3390/s19010156.
44. Vijayalakshmi S.R., Muruganand S. 6 Underwater Wireless Sensor Networks (UWSNs) // *Wireless Sensor Network. Architecture – Applications – Advancements*. 2018. pp. 175–212 DOI: 10.1515/9781683923275-007.

45. Blondel V.D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*. 2008. vol. 2008(10).
46. Dijkstra E.W. A note on two problems in connexion with graphs // *Numerische Mathematik*. 1959. vol. 1. pp. 269–271. DOI: 10.1007/BF01386390.
47. Ramaswami R., Parhi K. Distributed scheduling of broadcasts in a radio network // *Proceedings of the Eighth Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM)*. 1989. vol. 2. pp. 497–504. DOI: 10.1109/INFCOM.1989.101493.
48. Rappaport T. *Wireless Communications: Principles and Practice* // Upper Saddle River, NJ: Prentice Hall, 1996.
49. Thorp W.H. Deep Sound attenuation in the sub and low kilocycle per-second range // *J. Acoust. Soc. Am.* 1965. vol. 38(4). p. 648–654.
50. Cui J.-H., Kong J., Gerla M., Zhou S. The challenges of building scalable mobile underwater wireless sensor networks for aquatic applications // *IEEE Network*. 2006. vol. 20. no. 3. pp. 12–18. DOI: 10.1109/MNET.2006.1637927.
51. Tian K., Zhou C., Zhang J. Improved LEACH Protocol Based on Underwater Energy Propagation Model, Parallel Transmission, and Replication Computing for Underwater Acoustic Sensor Networks. *Sensors* // 2024. vol. 24(2). DOI: 10.3390/s24020556.
52. Fedorova T.A., Ryzhov V.A., Safronov K.S. et al. Energy-Efficient and Reliable Deployment Models for Hybrid Underwater Acoustic Sensor Networks with a Mobile Gateway // *J. Marine. Sci. Appl.* 2024. vol. 23. pp. 960–983. DOI: 10.1007/s11804-024-00444-z.
53. Федорова Т.А., Рыжов В.А., Сафронов К.С. Использование гибридной коммуникационной архитектуры подводной беспроводной сенсорной сети для повышения ее времени жизни и эффективности // *Информатика и автоматизация*. 2024. Т. 23 № 5. С. 1532–1570. DOI: 10.15622/ia.23.5.10.
54. Tian K., Zhou C., Zhang J. Improved LEACH Protocol Based on Underwater Energy Propagation Model, Parallel Transmission, and Replication Computing for Underwater Acoustic Sensor Networks // *Sensors*. 2024. vol. 24(2). DOI: 10.3390/s24020556.

Маевский Андрей Михайлович — канд. техн. наук, заместитель начальника отдела, отдел морской робототехники, Санкт-Петербургский государственный морской технический университет; научный сотрудник, «Научно-производственное предприятие подводных технологий «Океанос». Область научных интересов: системы управления, системы планирования движения, автономные объекты. Число научных публикаций — 30. Maevskiy_andrey@mail.ru; улица Лощманская, 3, 190121, Санкт-Петербург, Россия; р.т.: +7(981)786-9879.

Рыжов Владимир Александрович — д-р техн. наук, профессор, заведующий кафедрой, кафедра прикладной математики и математического моделирования, Санкт-Петербургский государственный морской технический университет. Область научных интересов: математическое моделирование, подводная робототехника, искусственный интеллект. Число научных публикаций — 130. varyzhov@smtu.ru; улица Лощманская, 3, 190121, Санкт-Петербург, Россия; р.т.: +7(921)952-7719.

Федорова Татьяна Александровна — канд. физ.-мат. наук, доцент кафедры, кафедра прикладной математики и математического моделирования, Санкт-Петербургский государственный морской технический университет. Область научных интересов: математическое моделирование, подводные беспроводные сенсорные сети, алгоритмы

планирования, искусственный интеллект. Число научных публикаций — 35. fedorova.tan@gmail.com; улица Лоцманская, 3, 190121, Санкт-Петербург, Россия; р.т.: +7(921)947-4797.

Печайко Иван Александрович — младший научный сотрудник, «Научно-производственное предприятие подводных технологий «Океанос». Область научных интересов: подводные роботы, алгоритмы планирования, манипуляционные системы, искусственный интеллект, симуляция подводных операций. Число научных публикаций — 8. pechaikojohn@gmail.com; Поэтический бульвар, 2, 194295, Санкт-Петербург, Россия; р.т.: +7(906)255-9933.

Буров Никита Михайлович — студент, Санкт-Петербургский государственный университет аэрокосмического приборостроения (ГУАП). Область научных интересов: алгоритмы автономного планирования, алгоритмы кластеризации. Число научных публикаций — 1. burov.nm@yandex.ru; ул. Большая Морская, 67, 190121, Санкт-Петербург, Россия; р.т.: +7(981)786-9879.

A. MAEVSKY, V. RYZHOV, T. FEDOROVA, I. PECHAIKO, N. BUROV
**STOCHASTIC DYNAMIC MODEL OF AN UNDERWATER
WIRELESS SENSOR NETWORK BASED ON A MODIFIED
LOUVAIN CLUSTERING ALGORITHM**

Maevsky A., Ryzhov V., Fedorova T., Pechaiko I., Burov N. Stochastic Dynamic Model of an Underwater Wireless Sensor Network Based on a Modified Louvain Clustering Algorithm.

Abstract. This paper presents an innovative approach to clustering and routing in Underwater Wireless Sensor Networks (UWSNs), based on a modified Louvain algorithm that considers sensor distances, the probability of successful message delivery, and the current energy levels of the nodes. The proposed method incorporates a dynamic reclustering mechanism driven by real-time monitoring of energy resources, allowing the network to adapt to environmental changes and sensor status by redistributing roles and restructuring clusters accordingly. The developed algorithm is designed to enhance energy efficiency, minimize data loss, and reduce the number of retransmissions in the context of limited bandwidth in underwater acoustic communication channels. A TDMA-based MAC protocol is also implemented to prevent collisions by assigning independent time slots to sensors, thereby eliminating interference. The approach addresses key resource management challenges in UWSNs by reducing energy consumption, improving data delivery reliability, shortening overall message transmission time, and extending the network's autonomous operation. The model takes into account the three-dimensional spatial deployment of sensors and optimizes the placement of reference nodes to avoid bottlenecks and excessive energy drain. The primary goal of the study is to construct a network topology that minimizes energy costs and message loss while ensuring efficient routing of data to reference nodes and onward to a mobile sink. The flexibility and adaptability of the proposed solution make it well-suited for real-world underwater applications such as environmental monitoring and ocean exploration.

Keywords: underwater wireless sensor networks, clustering algorithms, Louvain algorithm.

References

1. Zorzi M., Casari P., Baldo N., Harris A.F. Energy-efficient routing schemes for underwater acoustic networks. *IEEE J. Sel. Areas Commun.* 2008. vol. 26. no. 9. pp. 1754–1766.
2. Davis A., Chang H. Underwater wireless sensor networks. *Proceedings of the IEEE Oceans.* 2012. pp. 1–5.
3. Freitag L., Grund M., Singh S., Partan J., Koski P., Ball K. The WHOI micro-modem: An acoustic communications and navigation system for multiple platforms. *Proceedings Oceans MTS/IEEE.* 2005. vol. 2. pp. 1086–1092.
4. Rahman M.A., Lee Y., Koo I. EECOR: an energy efficient cooperative opportunistic routing protocol for underwater acoustic sensor networks. *IEEE Access.* 2017. vol. 5. pp. 14119–14132. DOI: 10.1109/ACCESS.2017.2730233.
5. Akyildiz I.F., Pompili D., Melodia T. Underwater acoustic sensor networks: Research challenges. *Ad Hoc Netw.* 2005. vol. 3(3). pp. 257–279. DOI: 10.1016/j.adhoc.2005.01.004.
6. Pompili D., Melodia T., Akyildiz I.F. Distributed routing algorithms for underwater acoustic sensor networks. *IEEE Trans. Wireless Commun.* 2010. vol. 9. no. 9. pp. 2934–2944. DOI: 10.1109/TWC.2010.070910.100145.

7. Lu Q., Liu F., Zhang Y., Jiang S. Routing Protocols for Underwater Acoustic Sensor Networks: A Survey from an Application Perspective. *Advances in Underwater Acoustics*. 2017. DOI: 10.5772/intechopen.68900.
8. Azam I., Majid A., Khan T., et al. Avoiding Energy Holes in Underwater Wireless Sensor Networks with Balanced Load Distribution. *10th International Conference on Complex, Intelligent and Software Intensive System (CISIS)*. 2016. pp. 341–350. DOI: 10.1109/CISIS.2016.109.
9. Chaaf A., Muthanna M.S.A., Muthanna A., et al. Energy-Efficient Relay-Based Void Hole Prevention and Repair in Clustered Multi-AUV Underwater Wireless Sensor Network. *Security and Communication Networks*. 2021. vol. 2021. no. 1. DOI: 10.1155/2021/9969605.
10. Khan Z.A., Karim O.A., Abbas S., Javaid N., Zikria Y.B., Tariq U. Q-learning based energy-efficient and void avoidance routing protocol for underwater acoustic sensor networks. *Computer Networks*. 2021. vol. 197. DOI: 10.1016/j.comnet.2021.10839.
11. Mhemed R., Phillips W., Comeau F., Aslam N. Void avoiding opportunistic routing protocols for underwater wireless sensor networks: A survey. *Sensors*. 2022. vol. 22(23). DOI: 10.3390/s22239525.
12. Tokmachev D.A., Chensky A.G., Zolotarev N.S., Poletaev A.S. Modelling of a hydroacoustic modem for underwater communications. *J. Phys.: Conf. Ser.* 2021. vol. 1728. DOI: 10.1088/1742-6596/1728/1/012021.
13. Chen H., Zhu Y., Zhang W., Wu K., Yuan F. Underwater Acoustic Micromodem for Underwater Internet of Things. *Wireless Communications and Mobile Computing*. 2022. vol. 2022. no. 1. DOI: 10.1155/2022/9148756.
14. Sherlock B., Morozs N., Neasham J., Mitchell P. Ultra-low cost and ultra-low-power, miniature acoustic modems using multipath tolerant spread spectrum techniques. *Electronics*. 2022. vol. 11. no. 9. DOI: 10.3390/electronics11091446.
15. Li J., Tong F., Jiang W., Roy S. Underwater acoustic modem with high precision ranging capability based on second-level cesium atomic clock. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2022. vol. XLVI-3/W1-2022. pp. 75–80, DOI: 10.5194/isprs-archives-XLVI-3-W1-2022-75-2022.
16. Zia I., Poncela J., Otero P. State-of-the-Art Underwater Acoustic Communication Modems: Classifications, Analyses and Design Challenges. *Wireless Personal Communications*. 2021. vol. 116. pp. 1325–1360. DOI: 10.1007/s11277-020-07431-x.
17. Datta A., Dasgupta M. Energy efficient topology control in underwater wireless sensor Networks. *Computers and Electrical Engineering*. 2023. vol. 105. DOI: 10.1016/j.compeleceng.2022.108485.
18. Sutagundar A.V., Halakarnimath B.S., Bhadjantri L.B. et al. A Novel Approach of topology control in underwater sensor networks. *IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*. 2022. pp. 1–5. DOI: 10.1109/MysuruCon55714.2022.9972400.
19. Choudhary M., Goyal N. Dynamic topology control algorithm for node deployment in mobile underwater wireless sensor networks. *Concurrency and Computation Practice and Experience*. 2022. vol. 34(15). DOI: 10.1002/cpe.6942.
20. Hu Y., Hu K., Liu H. et al. An energy-balanced head nodes selection scheme for underwater mobile sensor networks. *EURASIP Journal on Wireless Communications and Networking*. 2022. no. 63. DOI: 10.1186/s13638-022-02141-3.
21. Fedorova T.A., Ryzhov V.A., Semenov N.N. and Sulaiman Sh.A. Optimization of an Underwater Wireless Sensor Network Architecture with Wave Glider as a Mobile Gateway. *Journal of Marine Science and Application*. 2022. vol. 21. pp. 179–196. DOI: 10.1007/s11804-022-00268-9.

22. Ahmad I., Rahman T., Zeb A., et al. Cooperative Energy-Efficient Routing Protocol for Underwater Wireless Sensor Networks. *Sensors*. 2022. vol. 22. DOI: 10.3390/s22186945.
23. Bharny S., Sharma S., Alsharabi N., Eldin E.T., Ghamry N.A. Energy-efficient clustering protocol for underwater wireless sensor networks using optimized glowworm swarm optimization. *Front. Mar. Sci.* 2023. vol. 10. DOI: 10.3389/fmars.2023.1117787.
24. Sheeja R., Iqbal M.M., Sivasankar C. Multi-objective-derived energy efficient routing in wireless sensor network using adaptive black hole-tuna swarm optimization strategy. *Ad Hoc Networks*. 2023. vol. 144. DOI: 10.1016/j.adhoc.2023.103140.
25. Alfajeer A.A., Harous S. A review of routing protocols for underwater wireless sensor networks. The 3rd International Conference on Distributed Sensing and Intelligent Systems (ICDSIS 2022). 2022. pp. 138–145. DOI: 10.1049/icp.2022.2428.
26. Shovon I.I., Shin S. Survey on Multi-Path Routing Protocols of Underwater Wireless Sensor Networks: Advancement and Applications. *Electronics*. 2022. vol. 11(21). DOI: 10.3390/electronics11213467.
27. Kebkal A., Kebkal K., Komar M. Data-link protocol for underwater acoustic networks. *Europe Oceans*. 2005. vol. 2. pp. 1174–1180. DOI: 10.1109/OCEANSE.2005.1513225.
28. Stojanovic M. Optimization of a data link protocol for an underwater acoustic channel. *Europe Oceans*. 2005. vol. 1. pp. 68–73. DOI: 10.1109/OCEANSE.2005.1511686.
29. Guqhaiman A.A., Akanbi O., Aljeadi A., Chow C.E. A Survey on MAC Protocol Approaches for Underwater Wireless Sensor Networks. *IEEE Sensors Journal*. 2021. vol. 21(3). pp. 3916–3932. DOI: 10.1109/JSEN.2020.3024995.
30. Kulla E., Matsuo K., Barolli L. MAC Layer Protocols for Underwater Acoustic Sensor Networks: A Survey. *Lecture Notes in Network and Systems (LNNS)*. 2022. vol. 496. DOI: 10.1007/978-3-031-08819-3_21.
31. Gazi F., Ahmed N., Misra S., Wei W. Reinforcement Learning-Based MAC Protocol for Underwater Multimedia Sensor Networks. *CM Transactions on Sensor Networks*. 2022. vol. 18(3). pp. 1–25. DOI: 10.1145/3484201.
32. Liu X., Du X., Li M., Wang L., Li C. A MAC Protocol of Concurrent Scheduling Based on Spatial Temporal Uncertainty for Underwater Sensor Networks. *Hindawi Journal of Sensors*. 2021. vol. 2021. DOI: 10.1155/2021/5558078.
33. Shahapur S.S., Khanai R. Underwater Sensor Network at physical, data link and network layer – a survey. *International Conference on Communications and Signal Processing (ICCSP)*. 2015. pp. 1449–1453. DOI: 10.1109/ICCSP.2015.7322753.
34. Songzuo L., Iqbal B., Khan I.U., Ahmed N., Qiao G., Zhou F. Full Duplex Physical and MAC Layer-Based Underwater Wireless Communication Systems and Protocols: Opportunities, Challenges, and Future Directions. *J. Mar. Sci. Eng.* 2021. vol. 9(5). DOI: 10.3390/jmse9050468.
35. Casari P., Ardizzon F., Tomasin S. Physical Layer Authentication in Underwater Acoustic Networks with Mobile Devices. *WUWNet '22: Proceedings of the 16th International Conference on Underwater Networks & Systems*. 2022. pp. 1–8. DOI: 10.1145/3567600.3567604.
36. Kaveripakam S., Chinthajinjala R. Energy balanced reliable and effective clustering for underwater wireless sensor networks. *Alexandria Engineering Journal*. 2023. vol. 77. pp. 41–62. DOI: 10.1016/j.aej.2023.06.083.
37. Zhou Z., Peng Z., Cui J., Jiang Z. Handling triple hidden terminal problems for multichannel MAC in long-delay underwater sensor networks. *IEEE Trans. Mob. Comput.* 2012. vol. 11(1). pp. 139–154. DOI: 10.1109/TMC.2011.28.

38. Yahya A., Islam S.U., Zahid M., Ahmed G., Raza M., Pervaiz H. Cooperative routing for energy efficient underwater wireless sensor networks. IEEE Access. 2019. vol. 7. pp. 141888–141899. DOI: 10.1109/ACCESS.2019.2941422.
39. Rayen S.J., Arunajsmine J., Neelakandan S., Annamalai R. Social media networks owing to disruptions for effective learning. Procedia Comput. Sci. 2020. vol. 172. pp. 145–151. DOI: 10.1016/j.procs.2020.05.022.
40. Li P., Wang S., Zhang H., Zhang E. Improved particle swarm optimization algorithm of clustering in underwater acoustic sensor networks. Proceedings of the IEEE OCEANS. 2017. pp. 1–7. DOI: 10.1109/OCEANSE.2017.8084869.
41. Yu W., Chen Y., Wan L., Zhang X., Zhu P., Xu X. An energy optimization clustering scheme for multi-hop underwater acoustic cooperative sensor networks. IEEE Access. 2020. vol. 8. pp. 89171–89184. DOI: 10.1109/ACCESS.2020.2993544.
42. Wang M., Chen Y., Sun X., Xiao F., Xu X. Node energy consumption balanced multi-hop transmission for underwater acoustic sensor networks based on clustering algorithm. IEEE Access. 2020. vol. 8. pp. 191231–191241. DOI: 10.1109/ACCESS.2020.3032019.
43. Zhang W., Wang J., Han G. et al. A cluster sleep-wake scheduling algorithm based on 3D topology control in underwater sensor networks. Sensors. 2019. vol. 19. DOI: 10.3390/s19010156.
44. Vijayalakshmi S.R., Muruganand S. 6 Underwater Wireless Sensor Networks (UWSNs). Wireless Sensor Network. Architecture – Applications – Advancements. 2018. pp. 175–212 DOI: 10.1515/9781683923275-007.
45. Blondel V.D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment. 2008. vol. 2008(10).
46. Dijkstra E.W. A note on two problems in connexion with graphs. Numerische Mathematik. 1959. vol. 1. pp. 269–271. DOI: 10.1007/BF01386390.
47. Ramaswami R., Parhi K. Distributed scheduling of broadcasts in a radio network. Proceedings of the Eighth Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM). 1989. vol. 2. pp. 497–504. DOI: 10.1109/INFCOM.1989.101493.
48. Rappaport T. Wireless Communications: Principles and Practice. Upper Saddle River, NJ: Prentice Hall, 1996.
49. Thorp W.H. Deep Sound attenuation in the sub and low kilocycle per-second range. J. Acoust. Soc. Am. 1965. vol. 38(4). p. 648–654.
50. Cui J.-H., Kong J., Gerla M., Zhou S. The challenges of building scalable mobile underwater wireless sensor networks for aquatic applications. IEEE Network. 2006. vol. 20. no. 3. pp. 12–18. DOI: 10.1109/MNET.2006.1637927.
51. Tian K., Zhou C., Zhang J. Improved LEACH Protocol Based on Underwater Energy Propagation Model, Parallel Transmission, and Replication Computing for Underwater Acoustic Sensor Networks. Sensors. 2024. vol. 24(2). DOI: 10.3390/s24020556.
52. Fedorova T.A., Ryzhov V.A., Safronov K.S. et al. Energy-Efficient and Reliable Deployment Models for Hybrid Underwater Acoustic Sensor Networks with a Mobile Gateway. J. Marine. Sci. Appl. 2024. vol. 23. pp. 960–983. DOI: 10.1007/s11804-024-00444-z.
53. Fedorova T.A., Ryzhov V.A., Safronov K.S. [Using Hybrid Communication Architecture of Underwater Wireless Sensor Network to Improve Its Lifetime and Efficiency]. Informatics and Automation. 2024. no. 23(5). pp. 1532–1570. DOI: 10.15622/ia.23.5.10. (In Russ.).
54. Tian K., Zhou C., Zhang J. Improved LEACH Protocol Based on Underwater Energy Propagation Model, Parallel Transmission, and Replication Computing for

Maevsky Andrey — Ph.D., Deputy head of the department, Marine robotics department, Saint Petersburg State Marine Technical University; Researcher, Research and Production Enterprise for Underwater Technologies «Okeanos». Research interests: control systems, motion planning systems, autonomous systems. The number of publications — 30. Maevskiy_andrey@mail.ru; 3, Lotsmanskaya St., 190121, St. Petersburg, Russia; office phone: +7(981)786-9879.

Ryzhov Vladimir — Ph.D., Dr.Sci., Professor, Head of the department, Department of applied mathematics and mathematical modeling, Saint Petersburg State Marine Technical University. Research interests: mathematical modeling, underwater robotics, artificial intelligence. The number of publications — 130. varyzhov@smtu.ru; 3, Lotsmanskaya St., 190121, St. Petersburg, Russia; office phone: +7(921)952-7719.

Fedorova Tatyana — Ph.D., Associate professor, Department of applied mathematics and mathematical modeling, Saint Petersburg State Marine Technical University. Research interests: mathematical modeling, underwater wireless sensor networks, path planning algorithms, artificial intelligence. The number of publications — 35. fedorova.tan@gmail.com; 3, Lotsmanskaya St., 190121, St. Petersburg, Russia; office phone: +7(921)947-4797.

Pechaiko Ivan — Junior researcher, Research and Production Enterprise for Underwater Technologies «Okeanos». Research interests: underwater robots, path planning algorithms, manipulation systems, artificial intelligence, simulation of underwater operations. The number of publications — 8. pechaikojohn@gmail.com; 2, Poeticheskoy Av., 194295, St. Petersburg, Russia; office phone: +7(906)255-9933.

Burov Nikita — Student, Saint Petersburg State University of Aerospace Instrumentation (SUAI). Research interests: autonomous planning algorithms, clustering algorithms. The number of publications — 1. burov.nm@yandex.ru; 67, Bolshaya Morskaya St., 190121, St. Petersburg, Russia; office phone: +7(981)786-9879.

Руководство для авторов

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Информатика и автоматизация» <http://ia.spcras.ru/>. При регистрации авторам рекомендуется заполнить все предложенные поля данных. Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше или LaTeX. Объем основного текста (до раздела Литература) - от 20 до 30 страниц включительно. Переносы запрещены. Номера страниц не проставляются. Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материал а выделение подразделов. В основную часть опускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

Author guidelines

Interaction between each potential author and the Editorial board is realized through the personal account on the website of the journal "Informatics and Automation" <http://ia.spcras.ru/>. At the registration the authors are requested to fill out all data fields in the proposed form. The submissions should be prepared using MS Word 2007, LaTeX. The text of the paper in the main part should not exceed 30 pages. Pages are not numbered; hyphenations are not allowed. Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered in more detail at the journal web.

Signed to print 18.08.2025. Passed for print 19.08.2025.

Printed in Publishing center GUAP.

Address: 67 litera A, B. Morskaya, St. Petersburg, 190000, Russia

Founder and Publisher: SPC RAS.

Address: 39 litera A, 14th Line V.O., St. Peterburg, 199178, Russia.

The journal is registered in the Federal Service for Supervision of Communications,

Information Technology, and Mass Media,

Registration Certificate (registration number) ПИ № ФС77-79228 dated September 25, 2020

Subscription Index П5513, Russian Post Catalog

Подписано к печати 18.08.2025. Дата выхода в свет 19.08.2025.

Формат 60х90 1/16. Усл. печ. л. 14,5. Заказ № 257. Тираж 300 экз., цена свободная.

Отпечатано в Редакционно-издательском центре ГУАП.

Адрес типографии: Б. Морская, д. 67, лит. А, г. Санкт-Петербург, 190000, Россия

Учредитель и издатель: СПб ФИЦ РАН.

Адрес учредителя и издателя: 14-я линия В.О., д. 39, лит. А, г. Санкт-Петербург, 199178, Россия

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций, свидетельство о регистрации (регистрационный номер) ПИ № ФС77-79228 от 25 сентября 2020 г.

Подписной индекс П5513 по каталогу «Почта России»