СИСТЕМЫ АНАЛИЗА И ОБРАБОТКИ ДАННЫХ

ISSN 2782-2001

№ 3 (103) **2026**

УДК 681.51(06)
С 409

Учредитель

ФГБОУ ВО «Новосибирский государственный технический университет»

Редакционная коллегия

Главный редактор

Лемешко Борис Юрьевич, д-р техн. наук, проф., НГТУ, г. Новосибирск

Заместитель главного редактора

Вострецов Алексей Геннадьевич, д-р техн. наук, проф., НГТУ, г. Новосибирск

Ответственный секретарь

Воевода Александр Александрович, д-р техн. наук, проф., НГТУ, г. Новосибирск

Члены редколлегии

Володин Андрей Игоревич, канд. физ.-мат. наук, д-р философии по статистике, проф., Университет Регины, Канада
Воронин Владимир Викторович, д-р техн. наук, доцент, ТОГУ, г. Хабаровск
Гайдук Анатолий Романович, д-р техн. наук, проф., ЮФУ, г. Таганрог
Генералов Владимир Михайлович, д-р техн. наук, ФБУН ГНЦ ВБ «Вектор» Роспотребнадзора, г. Новосибирск
Григорьев Юрий Дмитриевич, д-р техн. наук, проф., СПбГЭТУ «ЛЭТИ», г. Санкт-Петербург
Гриф Михаил Геннадьевич, д-р техн. наук, проф., НГТУ, г. Новосибирск
Дмитриев Александр Капитонович, д-р физ.-мат. наук, НГТУ, г. Новосибирск
Комаров Андрей Константинович, д-р физ.-мат. наук, проф., Институт автоматики и электрометрии СО РАН, г. Новосибирск
Крюков Андрей Васильевич, д-р техн. наук, проф., ИрГУПС, г. Иркутск
Лисицына Лилия Ивановна, д-р техн. наук, проф., НГТУ, г. Новосибирск
Наумова Елена Николаевна, д-р техн. наук, проф., Университет Тафтса, США
Носек Ярослав, д-р техн. наук, проф., Технический университет, г. Либерец, Чехия
Персова Марина Геннадьевна, д-р техн. наук, проф., НГТУ, г. Новосибирск
Ронжин Андрей Леонидович, д-р техн. наук, проф., СПб ФИЦ РАН, г. Санкт-Петербург
Рубан Анатолий Иванович, д-р техн. наук, проф., СФУ, г. Красноярск
Рудяк Валерий Яковлевич, д-р физ.-мат. наук, проф., НГАСУ, г. Новосибирск
Сидоров Денис Николаевич, д-р физ.-мат. наук, проф., Институт систем энергетики им. Мелентьева СО РАН, г. Иркутск
Соловейчик Юрий Григорьевич, д-р техн. наук, проф., НГТУ, г. Новосибирск
Тудевдагва Уранчимэг, д-р техн. наук, проф., Хемницкий технический университет, Германия
Фомин Василий Михайлович, д-р физ.-мат. наук, проф., Институт теоретической и прикладной механики СО РАН, г. Новосибирск
Хабаров Валерий Иванович, д-р техн. наук, проф., СГУПС, г. Новосибирск
Ханова Анна Алексеевна, д-р техн. наук, проф., АГТУ, г. Астрахань
Хлебников Михаил Владимирович, д-р физ.-мат. наук, ИПУ им. В.А. Трапезникова РАН, г. Москва
Шандаров Станислав Михайлович, д-р физ.-мат. наук, проф., ТУСУР, г. Томск

Адрес редакции и издателя: 630073, г. Новосибирск, пр. К. Маркса, 20, корп. 7, комн. 409, 417
Тел. (383) 346-56-88, тел., факс (383) 344-49-98
E-mail: nauchvest@corp.nstu.ru
Web Site: http://journals.nstu.ru/vestnik/

16+

Индекс: 82962 («РОСПЕЧАТЬ») **УДК 681.51(06)**

© Коллектив авторов, 2026
© Новосибирский государственный технический университет, 2026

СИСТЕМЫ АНАЛИЗА И ОБРАБОТКИ ДАННЫХ

ISSN 2782-2001

№ 3 (103) **2026**

СОДЕРЖАНИЕ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

Гомбоин А.Б., Кориков А.М. Оптимизация рендеринга картографических данных в мобильных геоинформационных системах на фреймворке Flutter 7

Зеркаль С.М., Путята Е.А. Численное решение обратной кинематической задачи в линеаризованной постановке для неоднородных сред с использованием рефрагированных и рассеянных волн 23

Куликов В.В., Куцый Н.Н., Осипова Е.А. Схема придания автоматическим системам с широтно-импульсной модуляцией 1–5-го рода свойства суперпозиции 39

Кузьминых И.А., Гаврилов А.Б., Толстиков А.С., Тиссен В.М. Последовательное применение метода Прони для повышения точности прогнозирования временных рядов 57

Мамикьян М.С., Бекенева Я.А. Обработка текстовых данных в медицинских диалоговых системах на примере виртуального ассистента эндокринолога 75

Мезенцев Ю.А., Борисова А.А. Применение алгоритма муравьиной колонии в оптимизации управления мультипроектными разработками 97

Степанова А.И., Хальясмаа А.И., Матренин П.В. Анализ влияния гиперпараметров ансамблевых моделей на процесс их обучения в задаче прогнозирования электропотребления предприятия газовой промышленности 115

Хорев В.С., Цапко С.Г., Кузьминская Е.В., Командинов П.А. Достоверность измерений размеров канатной проволоки с использованием цифровых технологий и нейронных сетей 131

КОМПЬЮТЕРНЫЕ НАУКИ И ИНФОРМАТИКА

Лапатин И.Л., Назаров А.А., Пауль С.В., Салимзянова Д.Д. Характеристики потока обработанных задач модели вычислительного узла при конкуренции за ресурс 151

Правила для авторов 169

Корректор *Л.Н. Киншт*
Компьютерная верстка *С.И. Ткачева*

Новосибирский государственный технический университет
630073, г. Новосибирск, пр. К. Маркса, 20

Налоговая льгота – Общероссийский классификатор продукции
Издание соответствует коду 95 2000 ОК 005-93 (ОКП)

Журнал зарегистрирован Федеральной службой по надзору в сфере связи,
информационных технологий и массовых коммуникаций. Свидетельство о регистрации
средства массовой информации ПИ № ФС77-76010 от 03 июля 2019 г.
Периодичность издания – 4 номера в год.
Подписано в печать 21.09.2026. Выход в свет 25.09.2026.
Формат 70 × 108 1/16. Бумага офсетная. Тираж 300 экземпляров.
Уч.-изд. л. 15,05. Печатных листов 10,75. Издательский № 130. Заказ № 173.
Цена свободная

Отпечатано в типографии
Новосибирского государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20

ANALYSIS AND DATA PROCESSING SYSTEMS

ISSN 2782-2001

№ 3 (103) **2026**

УДК 681.51(06)
С 409

Journal was established by

Novosibirsk State Technical University

Editorial board

Chief editor

Lemeshko B. Yu., D. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Deputy chief editor

Vostretsov A. G., D. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Executive Secretary

Voevoda A. A., D. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

The members of the editorial Board

Volodin A. I., D. Sc. (Phys. & Math.), Doctor of Philosophy in Statistics from the University of Regina, Professor of Statistics University of Regina, Canada
Voronin V. V., D. Sc. (Eng.), Pacific national University, Khabarovsk, RF
Gaiduk A. R., D. Sc. (Eng.), Southern Federal University, Taganrog, RF
Generalov V. M., D. Sc. (Eng.), FBUN GNTS VB "Vektor" Rospotrebnadzora, Novosibirsk, RF
Grigor'ev Yu. D., D. Sc. (Eng.), Saint-Petersburg Electrotechnical University "LETI", Saint-Petersburg, RF
Grif M. G., D. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF
Dmitriev A. K., D. Sc. (Phys. & Math.), Novosibirsk State Technical University, Novosibirsk, RF
Komarov A. K., D. Sc. (Phys. & Math.), professor, Institute of Automation and Electrometry SB RAS, Novosibirsk, RF
Kryukov A. V., D. Sc. (Eng.), Irkutsk State Transport University, Irkutsk, RF
Lisitcina L. I., D. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF
Naumova E. N., PhD. (Math.), Tufts University, USA
Nosek J., D. Sc. (Eng.), Technical University of Liberec, Czech Republic
Persova M. G., D. Sc. (Eng.), professor, Novosibirsk State Technical University, Novosibirsk, RF
Ronzhin A. L., D. Sc. (Eng.), Federal Research Center Russian Academy of sciences, Saint-Petersburg, RF
Ruban A. I., D. Sc. (Eng.), Siberian Federal University, Krasnoyarsk, RF
Rudyak V. Ya., D. Sc. (Phys. & Math.), professor, Novosibirsk State University of Architecture and Civil Engineering, Novosibirsk, RF
Sidorov D. N., D. Sc. (Phys. & Math.), Melentiev Energy Systems Institute, Irkutsk, RF
Soloveichik Yu. G., D. Sc. (Eng.), professor, Novosibirsk State Technical University, Novosibirsk, RF
Tudevdagva U., D. Sc. (Eng.), Technische Universität Chemnitz, Germany
Fomin V. M., D. Sc. (Phys. & Math.), Khristianovich Institut of theoretical and applied mechanics SB RAS, Novosibirsk, RF
Khabarov V. I., D. Sc. (Eng.), Siberian Transport University, Novosibirsk, RF
Khanova A. A., D. Sc. (Eng.), Astrakhan State Technical University, Astrakhan, RF
Khlebnikov M. V., D. Sc. (Phys. & Math.), Institute of Control Sciences, Moscow, RF
Shandarov S. M., D. Sc. (Phys. & Math.), Tomsk State University of Control Systems and Radioelectronics, Tomsk, RF

Editorial Adress: 20 K. Marx Prospekt, building 7, rooms 409 and 417, Novosibirsk, 630073
Tel. (383) 346-56-88, Fax (383) 344-49-98
E-mail: nauchvest@corp.nstu.ru
Web Site: http://journals.nstu.ru/vestnik/

УДК 681.51(06)

© Authors, 2026
© Novosibirsk State Technical University, 2026

ANALYSIS AND DATA PROCESSING SYSTEMS

ISSN 2782-2001

№ 3 (103) **2026**

CONTENTS

INFORMATION TECHNOLOGIES AND TELECOMMUNICATIONS

Gomboin A.B., Korikov A.M. Optimization of cartographic data rendering in mobile geoinformation systems on the Flutter framework 7

Zerkal S.M., Putyata E.A. Numerical solution of the inverse kinematic problem in a linearized formulation for inhomogeneous media using refractive and scattered waves 23

Kulikov V.V., Kutsyy N.N., Osipova E.A. Scheme for imparting superposition property to automatic systems with pulse-width modulation of the 1st – 5th kind 39

Kuzminykh I.A., Gavrilov A.B., Tolstikov A.S., Tissen V.M. Sequential application of the Prony's method for improved time series forecasting accuracy 57

Mamikyan M.S., Bekeneva Ya.A. Processing text data in medical dialogue systems using the example of a virtual assistant for an endocrinologist 75

Mezentsev Y.A., Borisova A.A. Application of the ant colony algorithm in optimizing the management of multi-project development 97

A.I. Stepanova, Khalyasmaa A.I., Matrenin P.V. Analysis of the impact of hyperparameters of ensemble models on their training process in the task of predicting electricity consumption for a gas industry enterprise 115

Khorev V.S., Tsapko S.G., Kuzminskaya E.V., Komandinov P.A. Reliability of measuring rope wire dimensions using digital technologies and neural networks 131

COMPUTER SCIENCE AND INFORMATION SCIENCE

Lapatin I.L., Nazarov A.A., Paul S.V., Salimzyanova D.D. Characteristics of the output flow of processed tasks in a computational node model under resource competition 151

Rules for authors 169

Editor *L.N. Kinsht*
Computer imposition *S.I. Tkacheva*

Novosibirsk State Technical University
20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation

The journal is registered in the Federal Service for Supervision of Communications, Information Technology, and Mass Media. The certificate of registration of mass media ПИ № ФС77-76010 dated July 03, 2019.
Publication frequency – 4 numbers a year.
Signed in print September 21, 2026.
Date of publication September 25, 2026. Format 70 × 108 1/16
Offset Paper. Circulation is 300 copies. Educational-ed. liter. 15,05. Printed pages 10,75.
Publishing number 130. Order number 173.

It is printed in printing house of Novosibirsk State Technical University
20 K. Marx Prospekt, Novosibirsk, 630073

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 7–22

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 7–22

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 004.94

DOI: 10.17212/2782-2001-2026-3-7-22

Оптимизация рендеринга картографических данных в мобильных геоинформационных системах на фреймворке Flutter*

А.Б. ГОМБОИН[a], А.М. КОРИКОВ[b]

РФ, 634050, г. Томск, пр. Ленина, 40, Томский государственный университет систем управления и радиоэлектроники

[a] *a.gomboin@mail.ru* [b] *anatolii.m.korikov@tusur.ru*

В настоящей статье рассматривается задача оптимизации рендеринга сложных полигональных объектов в мобильных геоинформационных системах, реализованных на кросс-платформенном фреймворке Flutter. В условиях ограниченных вычислительных ресурсов мобильных устройств ГИС и высоких требований к отзывчивости интерфейса стандартные подходы к визуализации картографических данных, применяемые в библиотеке flutter_map, приводят к критической деградации производительности мобильных ГИС: снижению частоты кадров (FPS – Frames Per Second), росту потребления памяти и нестабильной реакции на пользовательские жесты.

Для обеспечения эффективной работы мобильных ГИС в статье предлагается два метода: отложенная загрузка полигонов для отсрочки пересчета до завершения взаимодействия пользователя с картой (картографическими данными) и кластеризация полигонов по географической близости с последующим геометрическим объединением в составные объекты. Экспериментальное исследование, проведенное на наборе из 5000 полигонов на реальном мобильном устройстве, показало, что комбинированное применение методов оптимизации позволило увеличить средний FPS, снизить потребление оперативной памяти и улучшить оценку пользовательского опыта. Результаты экспериментального исследования подтверждают, что даже в условиях архитектурных ограничений Flutter возможно достижение высокой производительности мобильных ГИС за счет оптимизации логики отображения данных.

Предложенные методы оптимизации рендеринга картографических данных не требуют глубокой модификации существующего кода – они легко интегрируются в архитектуру библиотеки flutter_map и могут быть адаптированы под любые типы полигонов и масштабы, что делает фреймворк Flutter перспективной платформой для разработки профессиональных мобильных ГИС-приложений, в том числе для работы в полевых условиях с минимальными ресурсами.

* *Статья получена 05 октября 2025 г.*

Ключевые слова: рендеринг, мобильное приложение, геоинформационные системы (ГИС), мобильная ГИС, Flutter, flutter_map, полигоны, оптимизация, отложенная загрузка полигонов, кластеризация полигонов, комбинирование методов оптимизации, viewport, геопространственные данные, OpenStreetMap

ВВЕДЕНИЕ

Мобильные приложения, установленные в смартфоне или планшетном компьютере, навигаторе и других подобных переносных устройствах, – это постоянные спутники современного человека, обеспечивающие ему неограниченные возможности и комфорт в повседневной жизни. Особенно полезными являются приложения геоинформационных систем. Они помогают человеку ориентироваться на местности, успешно решают задачи местоопределения и планирования маршрута от старта до финиша, указывают рациональное направление движения, обеспечивают информацией об окружающей среде. Такие приложения полезны во многих областях деятельности – в геологии, добыче нефти, газа и других полезных ископаемых. Трудно переоценить роль мобильных геоинформационных систем (ГИС) для людей, постоянно работающих в различной местности: им нужно постоянно пополнять знания о своем местонахождении и, возможно, в большем объеме, чтобы качественно выполнять свою работу. Как правило, на местности, где ведется поиск и разработка месторождений и/или другая подобная деятельность, покрытия Сети или недостаточно, или она полностью отсутствует, что накладывает ограничения на использование мобильных ГИС, так как они требуют постоянного подключения к интернету. Можно скачать различные карты местности и создать картографическую базу данных (КБД), но пользователям КБД нужна информация, связанная с их непосредственной деятельностью (например, информация о газопроводах, административных зданиях и других объектах, которые отсутствуют на обычных картах местности). Такую картографическую информацию можно загрузить на специальное мобильное приложение, которое позволяет отображать заранее заготовленные данные в специальном формате и представлять их в векторной форме. При использовании таких приложений возникает проблема, связанная с временными задержками работы пользовательского интерфейса, так как пользователям для работы необходимо большое количество картографических данных.

Вышеизложенное доказывает актуальность проблемы оптимизации рендеринга сложных картографических данных, в частности полигональных объектов в мобильном приложении на Flutter [1]. Для ее решения предлагается несколько подходов к оптимизации рендеринга: отложенная загрузка объектов на карту местности и применение кластеризации объектов в один объект.

Экспериментальная оценка показала улучшение скорости рендеринга и снижение нагрузки на память, что значительно повышает пользовательский опыт. Эти результаты предоставляют практические рекомендации для разработчиков, стремящихся сбалансировать визуальное качество и производительность в системах картографии на Flutter.

В первом разделе настоящей статьи рассматривается постановка задачи, анализируются особенности фреймворка, языка программирования и библио-

теки, применяемые для отображения карты местности и данных на ней, а также анализируется существующий метод кластеризации. Во втором разделе представлены решения по загрузке данных на карту местности в момент их непосредственного использования и объединению полигональных объектов в один объект для уменьшения нагрузки на рендеринг. В третьем разделе представлены результаты экспериментальных исследований.

1. ПОСТАНОВКА ЗАДАЧИ

Разработать метод оптимизации рендеринга сложных геопространственных полигонов с помощью инструментов библиотеки flutter_map, обеспечивающий сохранение функциональной точности, интерактивности и визуальной достоверности при одновременном снижении вычислительной нагрузки на рендер-конвейер Flutter, с целью повышения производительности и стабильности работы мобильных приложений ГИС при работе с массивами полигонов.

Современные ГИС, интегрируемые в мобильные и веб-приложения, всё чаще требуют визуализации крупномасштабных, многослойных, высокодетализированных пространственных объектов (в частности, полигонов, представляющих административные границы, зоны экологического мониторинга, территории инфраструктурных объектов, зоны риска и прочие географические сущности). При этом возросшие требования к интерактивности, отзывчивости и кросс-платформенной совместимости вынуждают разработчиков отказываться от традиционных нативных SDK [2] (Software Development Kit – набор инструментов для разработки программного обеспечения (ПО) под конкретную платформу, операционную систему, язык программирования, аппаратное обеспечение, облачные сервисы и фреймворки) в пользу универсальных фреймворков, таких как Flutter.

Однако применение Flutter для реализации картографических компонентов с высокой плотностью полигонов сталкивается с фундаментальными архитектурными ограничениями, обусловленными его декларативной моделью рендеринга. Библиотека flutter_map использует подход, при котором каждый полигон устанавливается как самостоятельный виджет PolygonLayer [2], содержащий собственный CustomPaint [2] и дерево геометрических примитивов. Такая реализация, хотя и обеспечивает гибкость в управлении стилями и событиями, приводит к экспоненциальному росту нагрузки на рендер-конвейер Flutter при увеличении числа полигонов. Это порождает критические деградации производительности: увеличение времени рендеринга, частоту кадров ниже 20 FPS (Frames Per Second – количество кадров в секунду), значительное потребление оперативной памяти (приблизительно 500 МБ и выше), а также нестабильную реакцию на пользовательские взаимодействия: скроллинг, масштабирование, клики.

Таким образом, возникает научно-практическая проблема – отсутствие эффективных методов оптимизации рендеринга сложных, многовершинных полигонов в рамках реактивной архитектуры Flutter, что препятствует внедрению высоконагруженных мобильных приложений ГИС в условиях ограниченных вычислительных ресурсов мобильных устройств.

1.1. ФРЕЙМВОРК FLUTTER И ЯЗЫК ПРОГРАММИРОВАНИЯ DART

Flutter – это открытый фреймворк для создания нативных пользовательских интерфейсов, разработанный компанией Google и выпущенный в 2017 году. Его ключевая особенность заключается в использовании единой кодовой базы для генерации приложений на разные операционные системы: iOS, Android, Web, Windows, macOS и Linux, что делает его особенно привлекательным для проектов с высокими требованиями к кросс-платформенности и скорости развертывания. В отличие от традиционных подходов к разработке кросс-платформенных мобильных приложений, Flutter использует собственный высокопроизводительный рендер-движок Skia, который обеспечивает прямую отрисовку UI-элементов на GPU без привлечения нативных компонентов операционной системы. Это позволяет достичь плавности анимаций до 120 FPS даже на устройствах среднего класса, что является критически важным для интерактивных геоинформационных систем [2–4]. В традиционных подходах, таких как WebView-интеграция, пользовательский интерфейс формируется с использованием веб-технологий и отображается внутри встроенного браузерного компонента операционной системы, что сопряжено с накладными расходами на взаимодействие между веб-контекстом и нативной средой. Альтернативный архитектурный паттерн native-bridge, применяемый в React Native и Xamarin, бизнес-логика которых пишется на кросс-платформенных языках JavaScript, C#, для взаимодействия с нативными компонентами операционной системы используют специальный «мост» для передачи сообщений между кросс-платформенным кодом и нативными API, однако асинхронная природа такого взаимодействия может приводить к задержкам при частом обмене данными. Это позволяет достичь плавности анимаций до 120 FPS даже на устройствах среднего класса, что является критически важным для интерактивных геоинформационных систем [2–4].

Основой Flutter является язык программирования Dart, разработанный Google как оптимизированный клиентский язык с поддержкой как компиляции в нативный код AOT – Ahead-of-Time, так и виртуальной машины JIT – Just-in-Time для целей разработки. Язык Dart реализует концепцию реактивного программирования через механизм widgets – декларативных, иммутабельных и рекурсивно-составных единиц интерфейса. Каждый виджет в Flutter представляет собой чистую функцию от состояния State в StatefulWidget или константное описание const Widget в StatelessWidget. Такой подход обеспечивает предсказуемость изменений интерфейса и упрощает трассировку источников перестроения дерева виджетов, но одновременно создает сложности при работе с динамическими, вычислительно тяжелыми элементами, такими как картографические полигоны с тысячами вершин.

Рендеринг в Flutter происходит в три этапа: 1) построение дерева виджетов, 2) создание дерева элементов, 3) построение дерева слоев. На последнем этапе все графические примитивы преобразуются в команды Skia, которые передаются на GPU. Однако при обработке сложных геометрических объектов (например, полигонов с десятками тысяч координат) дерево виджетов может содержать сотни тысяч экземпляров CustomPaint, Path, Polygon и других компонентов, что приводит к экспоненциальному росту времени на layout

(этап конвейера рендеринга, на котором фреймворк вычисляет размеры, координаты и иерархические ограничения для каждого визуального элемента) и paint-фазы (этап, следующий за layout, на котором формируются низкоуровневые графические инструкции для движка рендеринга), вызывая задержки и пропуски кадров [5].

1.2. БИБЛИОТЕКА FLUTTER_MAP

Для реализации карты и ее функций для Flutter существуют различные библиотеки, одной из которых является flutter_map [6] – одна из наиболее популярных open-source-альтернатив коммерческим решениям (например, Mapbox GL Native или Google Maps SDK). Flutter_map представляет собой порт JavaScript-библиотеки Leaflet на платформу Dart, адаптированный для работы в рамках реактивной архитектуры Flutter. Она предоставляет модульную структуру, основанную на следующих компонентах:

1) MapController – управление центром, масштабом, поворотом и другими функциями;
2) TileLayer – загрузка и отображение растровых или векторных тайлов XYZ, TMS, WMTS;
3) MarkerLayer – отображение точечных объектов;
4) PolylineLayer и PolygonLayer – отображение линий и замкнутых полигонов.

Полигоны в flutter_map реализованы через класс Polygon, который наследуется от Layer и содержит список координат List<LatLng>, представляющих вершины многоугольника. При каждом обновлении состояния карты, например при изменении масштаба или перемещении, flutter_map пересчитывает положение всех полигонов относительно текущего viewport (прямоугольная область экрана или координатного пространства, в пределах которой осуществляется отрисовка и отображение контента). Flutter_map применяет проекцию Mercator и передает ее в CustomPaint-виджет для отрисовки через Canvas.drawPath().

Однако здесь возникает фундаментальное архитектурное противоречие: flutter_map, будучи реализован как набор виджетов, не учитывает специфику геопространственных данных. Каждый полигон устанавливается как отдельный Widget, что приводит к проблемам.

При отображении большого количества полигонов создается большое количество экземпляров PolygonLayer, каждый из которых содержит свои CustomPaint, Path, Paint, и связанные с ними RenderBox. Это превышает допустимые пределы производительности Flutter, рекомендованные в официальной документации фреймворка: более 1000 виджетов на экране уже считается критическим порогом.

Полигоны рендерятся с одинаковой детализацией на всех масштабах, даже если они находятся за пределами области видимости или имеют слишком мелкие детали, невидимые при текущем zoom-уровне (дискретная или непрерывная величина, определяющая степень приближения / удаления карты относительно исходного масштаба).

Любое изменение состояния (например, изменение цвета одного полигона) вызывает перестроение всего PolygonLayer, поскольку Flutter не имеет встроенной системы дифференциального обновления для сложных геометрических примитивов.

Хранение координат полигонов в виде списков List<LatLng> в Dart-контейнерах без оптимизации памяти (например, через Float32List или ByteData) приводит к фрагментации кучи и увеличению GC-пауз (временные остановки выполнения основного потока приложения, вызванные работой автоматического сборщика мусора среды выполнения Dart). При работе с крупными наборами данных (например, границы регионов) объем потребляемой памяти может достигать 200…500 Мб на устройстве.

Такие проблемы вполне решаемые, а некоторые из них далее детализируются, чтобы показать эффективность разработанных подходов и продемонстрировать оптимизированную обработку больших объемов геоданных в реальном времени.

1.3. ГЕОМЕТРИЧЕСКАЯ МОДЕЛЬ ПОЛИГОНОВ: ПРЕДСТАВЛЕНИЕ, ПРОЕКЦИИ И РЕНДЕРИНГ, ПРОБЛЕМЫ РЕАЛИЗАЦИИ ГИС

Полигон в ГИС – это двумерная геометрическая фигура, определяемая замкнутой последовательностью точек или вершин на поверхности земного эллипсоида. В контексте цифровых карт каждая вершина выражается в географических координатах (широте и долготе) в системе WGS84 (World Geodetic System 1984 – всемирная система геодезических параметров Земли 1984 года, в число которых входит система геоцентрических координат) или EPSG:4326 (European Petroleum Survey Group – публичный реестр геодезических данных, систем пространственной привязки, земных эллипсоидов, преобразований координат и связанных с ними единиц измерения) [7]. Для отображения на плоском экране эти координаты должны быть спроецированы в плоскую систему, чаще всего в проекцию Меркатора – EPSG:3857, которая используется большинством картографических онлайн-сервисов: OpenStreetMap, Google Maps, Yandex.Map [7].

Проекция Меркатора сохраняет углы, но искажает площади и расстояния, особенно на высоких широтах. Это требует корректной трансформации координат перед рендерингом. В FlutterMap эта трансформация выполняется методом LatLngToPixel(), который преобразует географические координаты в пиксельные позиции на основе текущего масштаба, центра карты и размера экрана.

Сложность рендеринга возрастает, когда полигоны становятся многосвязными с внутренними «дырами» или содержат тысячи вершин (например, границы городов с учетом микрорайонов, береговых линий, лесных массивов). В таких случаях стандартный алгоритм отрисовки через Canvas.drawPath() становится неэффективным по следующим причинам:

1) каждый вызов drawPath() требует копирования всей последовательности вершин в буфер GPU;
2) отсутствует автоматическое сглаживание для малых масштабов;
3) не применяются алгоритмы упрощения геометрии (например, Douglas-Peucker);

4) не реализуется culling – отсечение невидимых полигонов за пределами view frustum (объем видимости виртуальной камеры).

Кроме того, Flutter не предоставляет встроенного механизма для батчинга – пакетной обработки геометрических примитивов. В отличие от WebGL (Web-based Graphics Library – кросс-платформенный API для отображения 3D-графики в браузере) или OpenGL ES (Open Graphics Library for Embedded Systems – программный интерфейс к графическому оборудованию, который позволяет указывать объекты и операции для создания высококачественных графических изображений, в частности цветных изображений трехмерных объектов), где можно отправить сотни полигонов одним вызовом glDrawElements(), Flutter вынуждает разработчика использовать либо CustomPaint, либо сторонние пакеты типа flutter_svg или flutter_map_3d, которые не решают проблемы производительности, а лишь добавляют дополнительный уровень абстракции.

1.4. АНАЛИЗ СУЩЕСТВУЮЩЕГО МЕТОДА КЛАСТЕРИЗАЦИИ

Если использовать популярные онлайн-карты (Google Maps, «Яндекс.Карты»), то можно заметить, что при уменьшении масштаба отдельные метки (например, кафе или АЗС) начинают сливаться в цифры: «12», «47», «203». Это кластеризация маркеров – один из самых известных и отлаженных приемов в геовизуализации. Она работает просто и эффективно. Система берет все точки в текущем viewport, группирует их поблизости (например, в ячейки сетки), заменяет каждую группу на одну метку с числом объектов внутри. Для точек это идеально. Геометрия тривиальна – одна пара координат, объединение – просто подсчет, визуализация – одна иконка и текст. Но важно отметить, что всё это работает только с точками. Эта технология реализована как для библиотеки flutter_map с помощью отдельного пакета flutter_map_marker_cluster [8], так и для других библиотек. Заметим, что эти пакеты не работают с полигонами, и готовые решения для работы с полигонами отсутствуют, поэтому кластеризация полигонов в мобильных ГИС сохраняет свою актуальность и в настоящее время.

2. МЕТОДЫ ОПТИМИЗАЦИИ РЕНДЕРИНГА ПОЛИГОНАЛЬНЫХ ОБЪЕКТОВ В FLUTTER MAP

2.1. ОТЛОЖЕННАЯ ЗАГРУЗКА И ОТРИСОВКА ПОЛИГОНОВ

Отложенная загрузка – стратегия, при которой объекты обрабатываются и отображаются только в момент их фактической необходимости вместо их предварительной обработки в полном объеме [9]. В контексте визуализации полигонов на карте это означает, что рендеринг происходит только для тех объектов, которые находятся в текущей области видимости пользователя – так называемом viewport, что также называют экстентом.

Отложенная загрузка основана на простом, но эффективном принципе: не отрисовывать то, что пользователь не видит. При перемещении или масштабировании карты viewport динамически изменяется, и система должна пересчи-

тывать, какие полигоны попадают в его границы. Все полигоны, находящиеся за пределами viewport, исключаются из процесса построения дерева виджетов и рендеринга, что позволяет существенно снизить нагрузку на поток, отвечающий за рендеринг пользовательского интерфейса, и GPU (Graphics Processing Unit – специализированный графический процессор, который обрабатывает данные, связанные с отображением изображений, видео и 3D-графики).

В рамках архитектуры Flutter, где каждый виджет – это объект в памяти, а каждый Polygon в flutter_map порождает собственный CustomPaint и RenderObject, отложенная загрузка становится критически важной при работе с тысячами полигонов. Без нее даже на мощных устройствах возникают задержки, пропуски кадров и повышенное потребление памяти.

В качестве примера нами реализовано приложение с подключенной библиотекой flutter_map. Это приложение принимает данные в виде файлов с расширением sqlite, сохраняющим данные по объектам, которые должны быть приведены на карте. Эти данные получаем SQL-запросом (Structured Query Language – язык структурированных запросов для работы с реляционными базами данных) к загруженному файлу. Далее эти данные кешируются в оперативную память и хранятся в переменной для удобного доступа к ним [10].

Данные, которые поместили в оперативную память, можем передавать в библиотеку flutter_map постепенно, по ходу движения пользователя по карте. Для этого необходимо знать границы экрана и соотнести их с границами на карте. В библиотеке имеется инструмент, который дает нужную информацию по границам. Библиотека передает их в виде объекта LatLngBounds. Благодаря этим данным можно получать только необходимые полигоны, а не целиком все сразу, и уменьшить нагрузку на рендеринг.

На рис. 1 представлен фрагмент кода, показывающий процесс выбора полигонов, которые находятся на экране. Также стоит отметить, что отображаются полигоны, имеющие несколько точек на экране. Точное количество пересчитывается в пропорции от общего числа внешних точек полигона, чтобы не засорять рендеринг.

В рассматриваемом коде имеется переменная visiblePolygonsController, которая является типом Stream, для потоковой передачи данных flutter_map. В этой библиотеке также есть встроенный инструмент для работы с потоковыми данными, благодаря которому легко можно передавать сразу на карту необходимые полигоны, и она будет своевременно предупреждена о новых данных.

Для такой реализации использовался также паттерн дебаунсер, так как он помогает сгладить рывки интерфейса, снизить нагрузку на процессоры устройства и избежать лишних вычислений при частых изменениях viewport, например, при быстром скроллинге или масштабировании карты.

Дебаунсер – это программный паттерн, который откладывает выполнение функции до тех пор, пока не пройдет определенный интервал времени без новых вызовов [11–13]. Если на этом интервале поступает новый вызов – таймер сбрасывается и выполнение вновь откладывается.

Этот паттерн помогает в ситуациях, когда при перемещении карты во flutter_map событие изменения позиции может срабатывать десятки раз

в секунду, особенно при перемещении по карте или различных скроллингах. Паттерн начинает свою работу после того, как пользователь остановился на каком-то месте на карте, и если в течение определенного времени не было изменений в карте, то подгружаются данные.

```
/// Метод возвращает список полигонов, которые находятся в зоне экрана
///
/// [bounds] - текущие границы видимой области карты
List<Polygon> getVisiblePolygons(LatLngBounds bounds) {
  // Фильтруем исходный список полигонов, оставляя только те,
  // которые попадают в заданные географические границы
  return listPolygones.where((polygon) {
    return _isPolygonInBounds(polygon, bounds);
  }).toList();
}

/// Проверяет, находится ли полигон в пределах заданных границ карты
///
/// Возвращает `true`, если хотя бы одна вершина полигона
/// попадает в область [bounds]
bool _isPolygonInBounds(Polygon polygon, LatLngBounds bounds) {
  // Проходим по всем точкам (вершинам) полигона
  for (final point in polygon.points) {
    // Если хотя бы одна точка находится внутри границ,
    // считаем весь полигон видимым
    if (bounds.contains(point))
      return true;
  }
  // Если ни одна точка не попала в границы, полигон невидим
  return false;
}

/// Обновляет поток видимых полигонов и уведомляет подписчиков
void updateVisiblePolygons(LatLngBounds bounds) {
  // Получаем отфильтрованный список полигонов для текущей области
  final visiblePolygons = getVisiblePolygons (bounds);
  // Отправляем обновленный список в StreamController
  _visiblePolygonsController.add(visiblePolygons);
}
```

Рис. 1. Фрагмент кода для выбора полигонов

Fig. 1. Code snippet for selecting polygons

2.2. КЛАСТЕРИЗАЦИЯ ПОЛИГОНОВ

При визуализации больших объемов геопространственных данных на мобильных устройствах, особенно в условиях ограниченных вычислительных ресурсов, возникает необходимость не только отбирать объекты для отображения, как при ленивой загрузке, но и сокращать их количество за счет семантического или геометрического объединения. Для этого применяется метод кластеризации и агрегации полигонов – подход, при котором множество отдельных полигонов, близко расположенных или логически связанных, объединяются в один составной объект, отображаемый на карте как единое целое [14–18].

Этот метод особенно эффективен на мелких масштабах, когда детализация отдельных объектов теряется, а их визуальное скопление приводит к перегруженности карты и снижению читаемости. Агрегация позволяет сохранить смысловую нагрузку данных, одновременно снижая нагрузку на рендер-конвейер Flutter.

Кластеризация делится на две части: геометрическое объединение и семантическая агрегация.

Геометрическое объединение – физическое слияние пересекающихся или близко расположенных полигонов в один составной полигон с общей границей.

Семантическая агрегация – объединение полигонов по признаку принадлежности к одной категории, региону или типу (например, все полигоны зон застройки в районе, один полигон района с меткой «количество входящих в него полигонов»).

В обоих случаях результат – уменьшение количества рендер-объектов, что напрямую снижает количество виджетов Polygon в дереве Flutter, число вызовов Canvas.drawPath(), объем памяти, занимаемый геометрией и стилями, а также нагрузку на графический процессор и поток пользовательского интерфейса.

В отличие от точечных объектов, для которых кластеризация реализована во многих библиотеках, для полигонов готовых решений нет, что делает эту задачу особенно актуальной.

В реализованном нами приложении использована следующая логика: группирование полигонов по географической близости; группирование по идентичным семантическим признакам (к примеру, лес, озеро, промышленная зона и т. д.).

Реализация представляет собой объединение по расстоянию между полигонами: берутся координаты каждого полигона и вычисляется площадь, самый большой полигон остается как полигон, который будет отображаться; задается расстояние между полигонами. Если расстояние меньше 10 пикселей, то полигоны объединяются в один. Для семантического объединения используем данные объектов, которые заранее известны.

На рис. 2, *а* показано, как были отрисованы полигоны до применения кластеризации; на рис. 2, *б* представлены полигоны после применения кластеризации; рис. 2, *б* визуально представляет суть кластеризации полигонов в мобильном приложении. Видно, что полигоны скрыты за самым большим полигоном, но информация о их существование не теряется.

а *б*

Рис. 2. Полигоны до применения кластеризации (*а*), после применения кластеризации (*б*)

Fig. 2. Polygons before clusterization (*a*), after clusterization is applied (*b*)

3. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ПОДХОДОВ ОПТИМИЗАЦИИ РЕНДЕРИНГА

Для оценки эффективности предложенных методов оптимизации – ленивой загрузки и кластеризации полигонов – была проведена серия экспериментов на симуляторе мобильного устройства Pixel 8 Pro с версией андроид 13 в условиях, максимально приближенных к эксплуатационным: отсутствие подключения к Wi-Fi, фоновые процессы, запущенные приложения.

В качестве тестового набора данных использовался массив из 5000 полигонов, представляющих условные участки промышленной застройки. Каждый полигон содержал в среднем 8–12 вершин. Полигоны были сгруппированы в плотные кластеры, примерно по 50–200 объектов на квадратный километр, при этом имитировались реальные условия размещения инфраструктурных объектов. Также запускалось приложение без использования группировки в кластеры и без использования отложенной загрузки. Результаты экспериментального исследования представлены в таблице.

Результаты оптимизации

Optimization results

Показатель	Базовая реализация	Оптимизированная реализация	Улучшение
FPS (средний)	18…25	52…58	в 2,5…3 раза
Пиковое потребление RAM	580 МБ	190 МБ	–67 %
Время отклика на жесты	300…500 мс	40…80 мс	в 6…7 раз
Количество рендер-объектов	5000	85…120 (кластеры)	в 40…60 раз

Экспериментальное исследование подтвердило высокую эффективность комбинированного применения ленивой загрузки и кластеризации полигонов по географической близости в мобильных ГИС-приложениях на Flutter с использованием библиотеки flutter_map.

Достигнуто трехкратное увеличение FPS и снижение потребления памяти на 67 %. Интерфейс стал отзывчивым и плавным даже на устройствах среднего класса. Приложение стало энергоэффективным и стабильным при длительной работе.

Результаты показывают, что даже в условиях архитектурных ограничений Flutter – декларативный рендеринг, отсутствие батчинга на уровне фреймворка – можно добиться высокой производительности за счет оптимизации логики отображения данных, а не переписывания ядра приложения.

ЗАКЛЮЧЕНИЕ

В условиях стремительного расширения сферы применения мобильных ГИС и постоянно возрастающих требований к их производительности, отзывчивости и энергоэффективности разработка оптимизированных методов визуализации картографических данных становится не только актуальной научно-технической задачей, но и ключевым фактором успешного внедрения профессиональных решений в полевых условиях применения ГИС. Особенно это актуально для отраслей, где стабильность и скорость работы мобильных приложений ГИС напрямую влияют на эффективность выполнения задач: геологии, нефтегазовой разведки, логистики, экологического мониторинга и т. п.

В данной работе решена научно-практическая задача оптимизации рендеринга сложных полигональных объектов в мобильном приложении на Flutter с учетом архитектурных особенностей фреймворка и ограничений мобильных устройств. Разработаны и экспериментально проверены два ключевых метода.

Отложенная загрузка полигонов с применением дебаунсера позволила исключить из рендеринга невидимые объекты и сгладить частоту пересчетов при активном взаимодействии с картой местности.

Кластеризация и агрегация полигонов по географической близости обеспечила снижение количества рендер-объектов за счет семантически корректного объединения соседних полигонов в составные объекты.

Экспериментальная оценка показала значительное улучшение ключевых метрик производительности: средний FPS увеличился, потребление

оперативной памяти снизилось, время отклика на пользовательские жесты уменьшилось.

Результаты исследования подтверждают, что даже в условиях декларативной архитектуры Flutter, где каждый виджет – это отдельный объект, а рендеринг основан на перестройке дерева, можно добиться высокой производительности за счет логики управления данными, а не переписывания фреймворка или перехода на нативную разработку. Это делает Flutter перспективной платформой для создания высоконагруженных ГИС-приложений, особенно в условиях кросс-платформенности и ограниченных ресурсов.

Предложенные методы не требуют глубокой модификации существующего кода – они легко интегрируются в архитектуру flutter_map и могут быть адаптированы под любые типы полигонов и масштабы.

СПИСОК ЛИТЕРАТУРЫ

1. *Гомбоин А.Б., Кориков А.М.* Параллельная обработка географических данных в мобильном приложении «Мобильная ГИС» // Доклады ТУСУР. – 2024. – Т. 27, № 2. – С. 51–56. – DOI: 10.21293/1818-0442-2024-27-2-51-56.

2. Official Flutter Documentation. – URL: https://docs.flutter.dev (accessed: 01.09.2025).

3. Dart: Language tour и performance best practices. – URL: https://dart.dev/guides/language/language-tour (accessed: 01.09.2025).

4. *Колтыгин Д.С., Седельников И.А., Крумин О.К.* Разработка кросс-платформенного приложения для управления робототехническим комплексом с помощью Flutter // Системы анализа и обработки данных. – 2025. – № 1 (97). – С. 7–26. – DOI: 10.17212/2782-2001-2025-1-7-26.

5. Google. Flutter Performance Best Practices. – URL: https://docs.flutter.dev/perf/rendering (accessed: 01.09.2025).

6. Flutter Map documentation. – URL: https://github.com/fleaflet/flutter_map (accessed: 01.09.2025).

7. *Dhonju H.K., Walsh K.B., Bhattarai T.* Web mapping for farm management information systems: A review and Australian orchard case study // Agronomy. – 2023. – Vol. 10. – P. 2563.

8. Flutter Map Marker Cluster. – URL: https://github.com/lpongetti/flutter_map_marker_cluster (accessed: 01.09.2025).

9. Optimizing Flutter apps for performance: A practical guide. – URL: https://medium.com/flutter/optimizing-flutter-apps-for-performance-7d7b3b45a1a2 (accessed: 01.09.2025).

10. *Silva Y.N., Almeida I., Queiroz M.* SQL: From traditional databases to big data // Proceedings of the 47th ACM Technical Symposium on Computing Science Education. – ACM, 2016. – P. 413–418.

11. *Yang L., Lan Q.* Front-end performance optimisation: An exploration of debounce and throttling strategies in image rendering // 2024 3rd International Conference on Cloud Computing, Big Data Application and Software Engineering. – IEEE, 2024. – P. 504–508.

12. *Ganssle J.G.* A guide to debouncing / The Ganssle Group. – Baltimore, MD, US, 2004. – 22 p.

13. *Нуждин Д.Г.* Оптимизация производительности мобильных приложений: стратегии и лучшие практики // Вестник науки. – 2023. – № 11, т. 3. – С. 850–862. – DOI: 10.24412/2712-8849-2023-1168-850-862.

14. *Douglas D.H., Peucker T.K.* Algorithms for the reduction of the number of points required to represent a digitized line or its caricature // The Canadian Cartographer. – 1973. – Vol. 10. – P. 112–122.

15. *Полищук Ю.М., Кочергин Г.А.* Использование геоинформационных систем для комплексного анализа пространственных данных на основе многомерной кластеризации // Геоинформатика. – 2011. – № 2. – С. 11–15.

16. *Колесенков А.Н.* Современные подходы к обработке данных при построении геоинформационных систем экологического мониторинга // Известия Тульского государственного университета. Технические науки. – 2016. – № 9. – С. 103–112.

17. *Hu Z., Li B., Luo J.* Flutter: Scheduling tasks closer to data across geo-distributed datacenters // IEEE INFOCOM 2016 – The 35th Annual IEEE International Conference on Computer Communications. – IEEE, 2016. – P. 1–9.

18. Development of a platform-independent renderer for the rendering of openstreetmap indoor maps in Flutter / J. Richter, R. Thomas, D. Lange, T. Graichen, U. Heinkel // Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP). – 2022. – Vol. 1. – P. 147–154.

Гомбоин Александр Булатович, аспирант кафедры автоматизированных систем управления Томского государственного университета систем управления и радиоэлектроники. Основное направление научных исследований – мобильные геоинформационные системы и их оптимизация. Имеет 10 научных публикаций. E-mail: a.gomboin@mail.ru

Кориков Анатолий Михайлович, доктор технических наук, профессор, заслуженный деятель науки РФ, академик МАН ВШ и МАИ, профессор кафедры автоматизированных систем управления Томского государственного университета систем управления и радиоэлектроники (ТУСУР), заслуженный профессор ТУСУР. Основное направление научных исследований – автоматизация и оптимизация систем обработки информации и управления. Имеет более 300 научных публикаций, в том числе 12 монографий. E-mail: anatolii.m.korikov@tusur.ru

Gomboin Alexander B., PhD student at the Department of Automated Control Systems, Tomsk State University of Control Systems and Radio Electronics. The main area of his scientific research is mobile geoinformation systems and their optimization. He has 10 scientific publications. E-mail: a.gomboin@mail.ru

Korikov Anatoly M., Doctor of Engineering Sciences, Professor, Professor at the Automated Control Systems Department in Tomsk State University of Control Systems and Radioelectronics (TUSUR), Honored Scientist of the Russian Federation, Academician of the Higher School. The main field of his scientific research is automation and optimization of information processing and control systems. He has more than 300 scientific publications, including 12 monographs. E-mail: anatolii.m.korikov@tusur.ru

DOI: 10.17212/2782-2001-2026-3-7-22

Optimization of cartographic data rendering in mobile geoinformation systems on the Flutter framework[*]

A.B. GOMBOIN[a], *A.M. KORIKOV*[b]

Tomsk State University of Control Systems and Radio Electronics, 40 Lenin Avenue, Tomsk, 634050, Russian Federation

[a] *a.gomboin@mail.ru* [b] *anatolii.m.korikov@tusur.ru*

Abstract

This article discusses the problem of optimizing the rendering of complex polygonal objects in mobile geographic information systems (GIS) implemented on the Flutter cross-platform framework. In conditions of limited computing resources of mobile GIS devices and high

[*] *Received 05 October 2025.*

requirements for interface responsiveness, standard approaches to visualization of cartographic data used in the flutter_map library lead to critical degradation of mobile GIS performance: a reduced frame rate (FPS – Frames Per Second), an increased memory consumption and an unstable response to user gestures.

To ensure an efficient operation of mobile GIS, the article proposes two methods: lazy loading of polygons to delay recalculation until the user's interaction with the map (cartographic data) is completed, and clustering polygons by geographical proximity followed by geometric integration into composite objects. An experimental study conducted on a set of 5,000 polygons using a real mobile device showed that the combined use of optimization methods made it possible to increase average FPS, reduce RAM consumption, and improve user experience assessment. The results of the experimental study confirm that even under the conditions of Flutter's architectural limitations, it is possible to achieve high performance of mobile GIS by optimizing the logic of data display.

The proposed methods for optimizing the rendering of cartographic data do not require deep modification of the existing code – they are easily integrated into the architecture of the flutter_map library and can be adapted to any type of polygon and scale, which makes the Flutter framework a promising platform for developing professional mobile GIS applications, including for work in the field with minimal resources.

Keywords: rendering, map, mobile application, GIS, mobile GIS, Flutter, flutter_map, polygons, optimization, lazy loading of polygons, clusterization of polygons, combination of optimization methods, viewport, geospatial data, OpenStreetMap

REFERENCES

1. Gomboin A.B., Korikov A.M. Parallel'naya obrabotka geograficheskikh dannykh v mobil'nom prilozhenii «Mobil'naya GIS» [Parallel processing of geographical data in a mobile application]. *Doklady TUSUR = Proceedings of TUSUR University*, 2024, vol. 27, no. 2, pp. 51–56. DOI: 10.21293/1818-0442-2024-27-2-51-56.

2. *Official Flutter Documentation*. Available at: https://docs.flutter.dev (accessed 01.09.2025).

3. *Dart: Language tour u performance best practices*. Available at: https://dart.dev/guides/language/language-tour (accessed 01.09.2025).

4. Koltygin D.S., Sedelnikov I.A., Krumin O.K. Razrabotka kross-platformennogo prilozheniya dlya upravleniya robototekhnicheskim kompleksom s pomoshch'yu Flutter [Development of a cross-platform application for controlling a robotic complex using Flutter]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2025, no. 1 (97), pp. 7–26. DOI: 10.17212/2782-2001-2025-1-7-26.

5. Google. *Flutter Performance Best Practices*. Available at: https://docs.flutter.dev/perf/rendering (accessed 01.09.2025).

6. *Flutter Map documentation*. Available at: https://github.com/fleaflet/flutter_map (accessed 01.09.2025).

7. Dhonju H.K., Walsh K.B., Bhattarai T. Web mapping for farm management information systems: A review and Australian orchard case study. *Agronomy*, 2023, vol. 10, p. 2563.

8. *Flutter Map Marker Cluster*. Available at: https://github.com/lpongetti/flutter_map_marker_cluster (accessed 01.09.2025).

9. *Optimizing Flutter apps for performance: A practical guide*. Available at: https://medium.com/flutter/optimizing-flutter-apps-for-performance-7d7b3b45a1a2 (accessed 01.09.2025).

10. Silva Y.N., Almeida I., Queiroz M. SQL: From traditional databases to big data. *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 2016, pp. 413–418.

11. Yang L., Lan Q. Front-end performance optimisation: An exploration of debounce and throttling strategies in image rendering. *2024 3rd International Conference on Cloud Computing, Big Data Application and Software Engineering*. IEEE, 2024, pp. 504–508.

12. Ganssle J.G. *A guide to debouncing*. Ganssle Group. Baltimore, MD, US, 2004. 22 p.

13. Nuzhdin D.G. Optimizatsiya proizvoditel'nosti mobil'nykh prilozhenii: strategii i luchshie praktiki [Optimization of mobile application performance: strategies and best practices]. *Vestnik nauki*, 2023, no. 11, vol. 3, pp. 850–862. DOI: 10.24412/2712-8849-2023-1168-850-862. (In Russian).

14. Douglas D.H., Peucker T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 1973, vol. 10, pp. 112–122.

15. Polishchuk Yu.M., Kochergin G.A. Ispol'zovanie geoinformatsionnykh sistem dlya kompleksnogo analiza prostranstvennykh dannykh na osnove mnogomernoi klasterizatsii [Use of geoinformation systems for complex analysis of spatial data based on multidimensional clusterization]. *Geoinformatika*, 2011, no. 2, pp. 11–15. (In Russian).

16. Kolesenkov A.N. Sovremennye podkhody k obrabotke dannykh pri postroenii geoinformatsionnykh sistem ekologicheskogo monitoringa [Modern approaches to the processing of data in the construction of information systems of environmental monitoring]. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki = News of the Tula state university. Technical sciences*, 2016, no. 9, pp. 103–112.

17. Hu Z., Li B., Luo J. Flutter: Scheduling tasks closer to data across geo-distributed datacenters. *IEEE INFOCOM 2016 – The 35th Annual IEEE International Conference on Computer Communications*. IEEE, 2016, pp. 1–9.

18. Richter J., Thomas R., Lange D., Graichen T., Heinkel U. Development of a platform-independent renderer for the rendering of openstreetmap indoor maps in Flutter. *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, 2022, vol. 1, pp. 147–154.

Для цитирования:

Гомбоин А.Б., Кориков А.М. Оптимизация рендеринга картографических данных в мобильных геоинформационных системах на фреймворке Flutter // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 7–22. – DOI: 10.17212/2782-2001-2026-3-7-22.

For citation:

Gomboin A.B., Korikov A.M. Optimizatsiya renderinga kartograficheskikh dannykh v mobil'nykh geoinformatsionnykh sistemakh na freimvorke Flutter [Optimization of cartographic data rendering in mobile geoinformation systems on the Flutter framework]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 7–22. DOI: 10.17212/2782-2001-2026-3-7-22.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 23–38

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 23–38

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.6

DOI: 10.17212/2782-2001-2026-3-23-38

Численное решение обратной кинематической задачи в линеаризованной постановке для неоднородных сред с использованием рефрагированных и рассеянных волн*

С.М. ЗЕРКАЛЬ[a], Е.А. ПУТЯТА[b]

РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет

[a] *zerkal@corp.nstu.ru* [b] *putyata.2018@stud.nstu.ru*

Исследована обратная кинематическая задача для вертикально-неоднородных сред, в которых скорость распространения волн зависит только от глубины, с использованием рефрагированных и рассеянных волновых сигналов. В отличие от классических постановок обратных кинематических задач, где основное внимание уделяется прямым или отраженным волнам, в настоящей работе ключевым объектом анализа выступают рассеянные волны, в которых траектория была изменена в результате взаимодействия с локальными неоднородностями среды. Такой подход расширяет область применимости метода на среды, демонстрирующие волноводные эффекты, где волны «запираются» в слое и распространяются на большие расстояния с минимальными потерями. Разработан вычислительный алгоритм на основе линеаризации функционала Ферма, позволяющий определять скоростное распределение $V(z)$ в таких средах. Для прямой задачи применены как аналитические решения (в линейном случае $V(z) = 1 + Bz$), так и численные методы (для нелинейных случаев), включая метод Рунге – Кутты 4-го порядка. Реализованное решение показало возможность восстановления скоростного распределения среды по измеренному времени регистрации волновых сигналов; выявлены погрешности, связанные с неточностью линеаризации. Анализ показал, что ошибка линеаризации растет с увеличением глубины, что, в свою очередь, может быть компенсировано введением дополнительных регуляризирующих процедур, таких как, например, сглаживание шумов с помощью сплайнов. Полученные результаты могут быть интересны для решения задач сейсмики, акустического зондирования и неразрушающего контроля. Перспективы дальнейших исследований связаны прежде всего с переходом к более сложным моделям сред с применением более совершенных методов регуляризации, а также с гибридизацией данного подхода с моделями искусственного интеллекта. Всё это открывает перспективы к расширению области применимости и разрешающей способности метода.

* *Статья получена 10 марта 2026 г.*

Ключевые слова: обратная кинематическая задача, рефрагированные волны, рассеяние, компьютерное моделирование, линеаризация функционала Ферма, волноводные среды, численные методы, вертикально-неоднородные среды

ВВЕДЕНИЕ

Дистанционная диагностика удаленных объектов привлекала внимание ученых с давних пор, и ее актуальность со временем только возрастает. Причем с развитием «человеческого знания» естественно становились доступными всё более удаленные объекты, возрастала детальность, разрешающая способность и границы применимости таких исследований. Часто разрабатываемые технологии имели «двойное назначение» – как в гражданском, так и в военном отношении (например, решение локационных задач, в том числе в гидроакустике) [1]. Достаточно сказать, что серьезный толчок к развитию сейсмических исследований возник в результате попыток с использованием поверхностных волн локализовать тяжелую артиллерию противника во время Первой мировой войны. Незадолго до этих событий, в 1905–1907 годах, геофизик Е. Вихерт и математик Г. Герготц получили первый результат, относящийся к «обратной кинематической задаче сейсмики», являющейся одной из основных задач теории распространения волн. Обратная кинематическая задача (ОКЗ) в общем случае заключается в определении скоростного распределения распространения сигнала в заданной среде для рефрагированных волн в некоторой области исследуемой среды по зарегистрированному времени пробега этих волн от источника до приемника колебаний, расположенных на границе отмеченной области, например поверхности Земли [2]. Исследования выполняются в «приближении геометрической оптики», что в сейсмике имеет название «геометрическая сейсмика».

Обратная кинематическая задача – типичная обратная задача математической физики. С философских позиций ОКЗ характеризуется обращением причинно-следственных связей: по следствию – времени прихода рефрагированного волнового сигнала (результату решения прямой задачи) – определяется его скоростное распространение в заданной среде, т. е. причина вместе с генерируемым зондирующим сигналом, породившими это следствие. Поэтому уместно отметить корреляцию с принятым в области искусственного интеллекта понятием «целенаправленный поиск в пространстве состояний»: прямая задача соответствует поиску от исходных данных к цели, а обратная – от цели к исходным данным. В этом случае исходные данные – скоростное распределение распространения сигнала в исследуемой среде, а цель – вектор измерений (время прихода рефрагированных волн), называемый в теории распространения волн (сейсмоакустике) годографом. Как обратная задача математической физики, ОКЗ является некорректной и требует для своего решения регуляризирующего воздействия. При этом она может иметь переопределенные постановки, а степень некорректности – как сильную, так и слабую – в соответствии с классификацией, принятой в отечественной научной школе по «условно-корректным задачам» [3].

Основным интегральным соотношением ОКЗ является функционал Ферма (1) – известная формула «Задачи геометрической оптики в пространстве» [4, 8]. Функционал Ферма связывает $V = V(x, y, z)$ – скоростное распре-

деление в исследуемой среде, $t = t(x^0, x^1)$ – время прихода рефрагированной волны от источника $x^0 = (x_0, y_0, z_0)$ до приемника $x^1 = (x_1, y_1, z_1)$, вдоль лучевой траектории $\Gamma = \Gamma(x^0, x^1)$:

$$t = t(x^0, x^1) = \int\limits_{\Gamma(x^0, x^1)} \frac{dS}{V}. \qquad (1)$$

Здесь *ds* – элемент дуги кривой $\Gamma(x^0, x^1)$. Функционал Ферма в соответствии с понятиями вариационного исчисления позволяет, зная подынтегральную функцию, найти $\Gamma(x^0, x^1)$ как решение соответствующей системы уравнений Эйлера. Таким образом, $\Gamma(x^0, x^1)$ являются геодезическими.

Если поставить ОКЗ как определение функции V по известным значениям t (естественно, при этом необходимо знать геодезические Г), то в такой постановке ОКЗ представляет собой классическую задачу «интегральной геометрии» [4], для решения которой предпочтительным является получение формулы обращения, т. е. для оператора A ($t = AV$) построение обратного оператора A^{-1}:

$$t = AV, \;\; A^{-1}t = A^{-1}AV, \;\; V = A^{-1}t. \qquad (2)$$

При решении практических задач точная информация, как правило, отсутствует, и прежде всего это относится к лучевым траекториям. Таким образом, построение обратного оператора A^{-1} становится нереализуемым. Решать такие задачи возможно только приближенно, используя априорную информацию об исследуемой среде, а точнее, об искомом скоростном распределении.

Важным обстоятельством является то, что для построения A^{-1} значения функции t недостаточно, необходимо знать еще и функции Г. При практическом применении значение t – это измеренное время прихода волн. Однако точные значения параметров среды (например, скорость и зависящую от нее лучевую траекторию) получить невозможно, поэтому в общей постановке ОКЗ не имеет решения.

Однако, как отмечалось выше, знание так называемой априорной информации об исследуемой среде, о распространении зондирующего волнового сигнала позволяет получить приближенное решение ОКЗ, основанное, например, на предположении о том, что искомая скорость V и показатель преломления n представляется в виде

$$V = V_0 + V_1, \;\; n = V^{-1}, \;\; n = n_0 + n_1, \;\; n_0 = \frac{1}{V_0},$$

$$V_0 \gg |V_1|, \;\; n_0 \gg |n_1|.$$

Кроме того, среда предполагается регулярной, т. е. паре точек «источник – приемник» соответствует одна соединяющая их траектория Г. Такая информация о свойствах V позволяет применить для решения ОКЗ метод линеаризации.

Предположим, что n_0 – известная функция, а определению подлежит функция n_1. Тогда можно «эмпирически» простой заменой в функционале Ферма геодезической Г, соответствующей $n = n_0 + n_1$, на Γ_0, соответствующую n_0, получить для n_1 следующее выражение:

$$t = \int_{\Gamma_0} (n_0 + n_1)ds = \int_{\Gamma_0} n_0 ds + \int_{\Gamma_0} n_1 ds. \qquad (3)$$

Поскольку $t_0 = \int_{\Gamma_0} n_0 ds$, то $t_1 = t - t_0 = \int_{\Gamma_0} n_1 ds$, $t_1 = An_1 \Rightarrow A^{-1}t_1 = A^{-1}An_1 = n_1$.

Таким образом получили реализацию выражения (2).

В настоящей работе рассматривается постановка обратной кинематической задачи для вертикально-неоднородных сред, в которых скорость распространения волн зависит только от глубины: $V = V(z)$. Предполагается, что функция $V(z)$ допускает линеаризацию в окрестности некоторого постоянного значения, что упрощает последующий анализ. В качестве основного сигнала используются рефрагированные акустические волны, рассеянные на внутренних неоднородностях среды, т. е. не прямые волны от источника, а именно те, что претерпели изменение траектории в результате взаимодействия с неоднородностями. Актуальность обусловлена потребностью в устойчивых и точных методах диагностики неоднородных сред, особенно в условиях зашумленных данных, что делает подход перспективным для сейсмики, акустического зондирования и неразрушающего контроля. Такой подход расширяет область применимости метода, позволяя работать со средами, в которых проявляются волноводные эффекты.

1. ПОСТАНОВКА ЗАДАЧИ И ПОСТРОЕНИЕ РЕШЕНИЯ

В рамках предложенного подхода, основанного на анализе рассеянных волн, далее рассматривается математическая постановка обратной кинематической задачи для вертикально-неоднородных сред. В отличие от классических постановок обратных кинематических задач [6, 7], где используются только прямые или отраженные волны, ключевым объектом анализа выступают рассеянные волны [8], в которых траектория была изменена в результате взаимодействия с локальными неоднородностями среды.

Эту постановку естественно назвать обратной кинематической задачей рассеяния по аналогии с обратной задачей рассеяния в квантовой механике или акустике, где по наблюдаемым характеристикам рассеянного поля восстанавливается структура рассеивателя. В данном случае по временным задержкам рассеянных волн восстанавливается профиль скорости $V(z)$.

Система наблюдений задается следующим образом: в плоскости xz источник сигнала расположен в точке $(0, 0)$, а приемник – в точке $(l, 0)$. Известны два угла: a_1 – угол выхода луча из источника, a_2 – угол прихода регистрируемого луча (рис. 1). Цель – восстановить функцию скорости $V(z)$, описывающую среду на основе измеренного времени прихода рассеянных волн.

Важным уточнением является то, что эта постановка применима в условиях, когда возможно генерировать узконаправленные волновые пучки – «шнуровые сигналы», которые являются решением волнового уравнения функционала Ферма, представляя собой лучи, аналогично работе [12]. Такие решения характеризуются высокой локализацией в пространстве и позволяют избежать неоднозначности при восстановлении волноводных структур, а также лучи являются более простыми объектами, чем волны. Хотя требования к генерации таких лучей накладывают определенные ограничения [13] на экспериментальную реализацию, они компенсируются однозначностью восстановления (исходя из принципа Ферма), в отличие от классических обратных задач, где часто возникает неоднозначность, связанная с симметрией или недостатком данных.

Математически задача сводится к решению нелинейного интегрального уравнения первого рода [9–11], связывающего время регистрации сигналов t с искомым скоростным распределением $V(z)$ в заданной среде. В линеаризованной постановке, когда $V(z)$ близка к константе, задача упрощается и может быть решена аналитически или численно:

$$A(c) = \int\limits_{\Gamma(x)} \frac{dS}{V(z)} = t(x), \quad \Gamma(x) = \Gamma_1(x) \cup \Gamma_2(x). \tag{4}$$

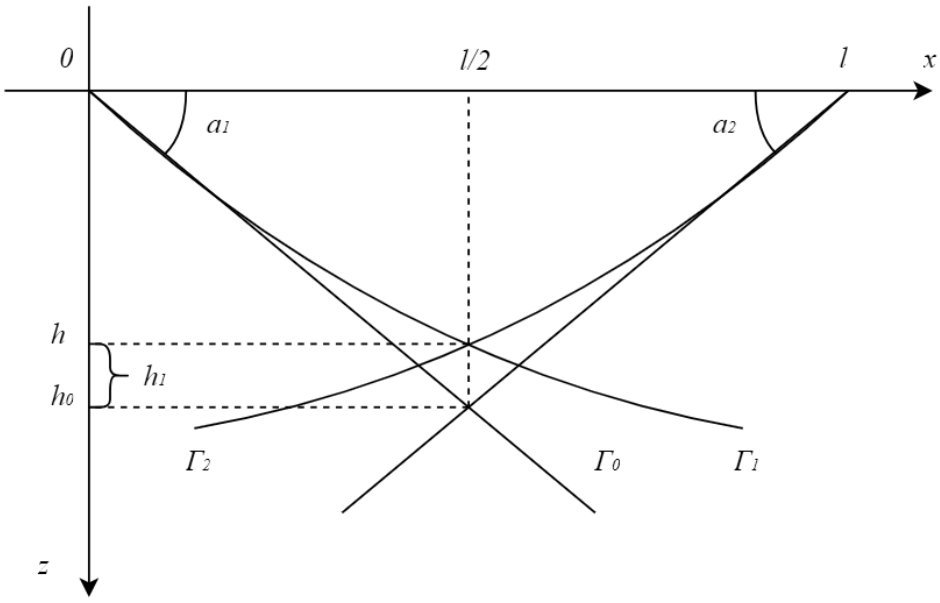

Рис. 1. Геометрическая схема выполнения измерений в ОКЗ рассеяния:
$\Gamma_0(\bar{x})$ – луч для среды $n = n_0$ при отсутствии неоднородности; $\Gamma_1(\bar{x})$ – луч для $n = n_0 + n_1$; $\Gamma_2(\bar{x})$ – луч сигнала, рассеянного на неоднородности среды

Fig. 1. Geometric diagram of measurement execution in the inverse kinematic scattering problem:
$\Gamma_0(\bar{x})$ is a ray for $n = n_0$ in the absence of heterogeneity; $\Gamma_1(\bar{x})$ is a ray for $n = n_0 + n_1$; $\Gamma_2(\bar{x})$ is a ray of a signal scattered by a medium inhomogeneity

В точке $l/2$ Γ_1 и Γ_2 пересекаются, $(\bar{x}) = (x,\ z)$.

Рассматривается частный случай, когда углы выхода и прихода лучей совпадают: $\alpha_1 = \alpha_2 = \alpha$. В рамках линеаризованной постановки предполагается, что показатель преломления среды задается константой n_0. При этих условиях справедливо следующее соотношение:

$$t - t_0 = 2 \int\limits_{\Gamma_0} n_1(z) ds, \tag{5}$$

где t_0 – время регистрации рефрагированного сигнала при отсутствии неоднородности; t – фактическое время; Γ_0 – луч (лучевая траектория), который строится исходя из уравнения

$$x = z \operatorname{ctg}(\alpha).$$

Тогда

$$ds = dz \sqrt{1 + \left(\frac{dx}{dz}\right)^2}, \tag{6}$$

$$t(\alpha) - t_0(\alpha) = 2 \int\limits_0^{\frac{1}{2}\operatorname{tg}\alpha} n_1(z) \sqrt{1 + x_0^2 dz}. \tag{7}$$

Отсюда

$$n_1\left(\frac{l}{2}\operatorname{tg}\alpha\right) = \frac{\cos^2\alpha}{2l} \frac{d}{d\alpha}\left[\sin\alpha\left(t(\alpha) - t_0(\alpha)\right)\right]. \tag{8}$$

Таким образом, меняя угол α, можно определить показатель преломления в исследуемой среде.

Прямая кинематическая задача сейсмики рассматривается в рамках лучевого метода. В результате ее решения по заданному изменению скорости распространения акустической волны в среде $V(x, y, z)$ строятся лучевые траектории, вдоль которых распространяется исследуемая волна, вычисляется время пробега волны вдоль луча.

Далее рассматривается прямая задача для модели градиентной среды с линейным законом изменения скорости от глубины:

$$V(z) = 1 + Bz, \quad B = \text{const}, \quad B > 0.$$

Соответственно, коэффициент преломления также меняется линейно:

$$n_1(z) = 1 + n(z), \text{ где } n_1(z) = Bz.$$

Согласно рис. 2

$$R^2 = \left(z + \frac{1}{B}\right)^2 + \left(x - \frac{l}{2}\right)^2 \Rightarrow z = -\frac{1}{B} + \sqrt{R^2 - \left(x - \frac{l}{2}\right)^2}. \tag{9}$$

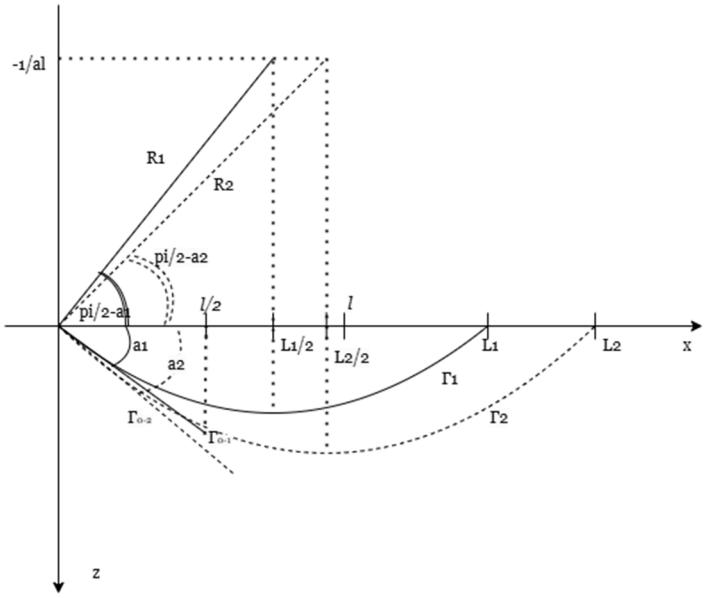

Рис. 2. Дополнительная геометрическая схема для вычислений (9)

Fig. 2. An additional geometric diagram for calculations (9)

Время прихода рассеянной и зарегистрированной в точке поверхности $x_1 = l$ волны, испущенной из источника, расположенного в точке $x_0 = 0$, рассчитывается по формуле, учитывающей соотношение (9). В результате получаем

$$t\left(0,\ \frac{l}{2}\right) = \frac{1}{2B} \ln \frac{R + \frac{l}{2}}{R - \frac{l}{2}}. \tag{10}$$

Для расчета t_0 используется формула

$$t_0(\alpha) = \frac{l}{\cos\alpha}. \tag{11}$$

Таким образом, на основе результатов вычислений (10) и (11) при разных значениях угла решается обратная кинематическая задача через формулу обращения (8). Важным уточнением является то, что подобное решение применимо для регулярных вертикально-неоднородных сред при линеаризованной постановке обратной кинематической задачи.

2. РЕЗУЛЬТАТЫ ВЫЧИСЛЕНИЙ

Далее выстраивается решение прямой кинематической задачи для генерации данных согласно постановке. В соответствии с построенным решением разработан вычислительный алгоритм и его программная реализация на языке Python, проведены численные компьютерные эксперименты с целью опреде-

ления параметра B, решения прямой кинематической задачи для линеаризованного случая и для точного. Результаты приведены на рис. 3 и 4.

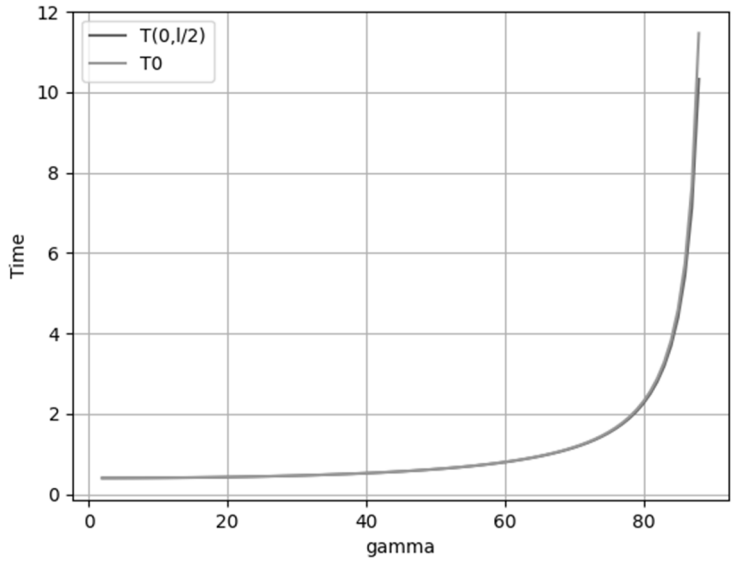

Рис. 3. Сравнение графиков времени прихода волны

Fig. 3. Comparison of wave arrival times

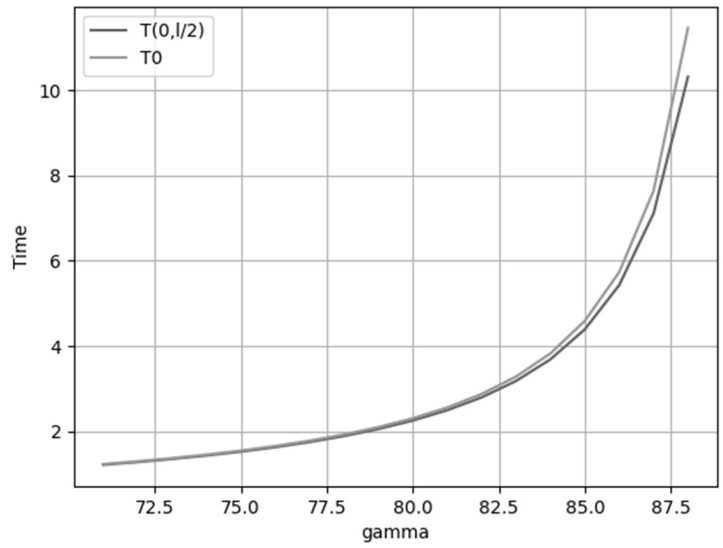

Рис. 4. Сравнение графиков времени прихода волны (приближение 70…90 градусов)

Fig. 4. Comparison of the arrival time graphs of the wave (zoom of 70…90 degrees)

Рассчитывается абсолютная и относительная ошибка и строятся графики зависимости ошибки от угла входа сигнала в исследуемую среду.

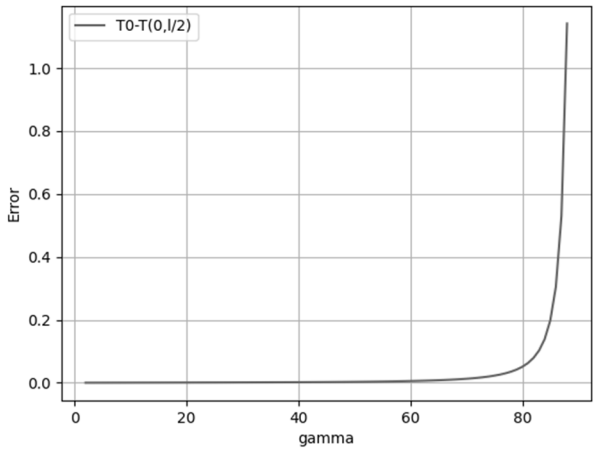

Рис. 5. График абсолютной ошибки

Fig. 5. An absolute error graph

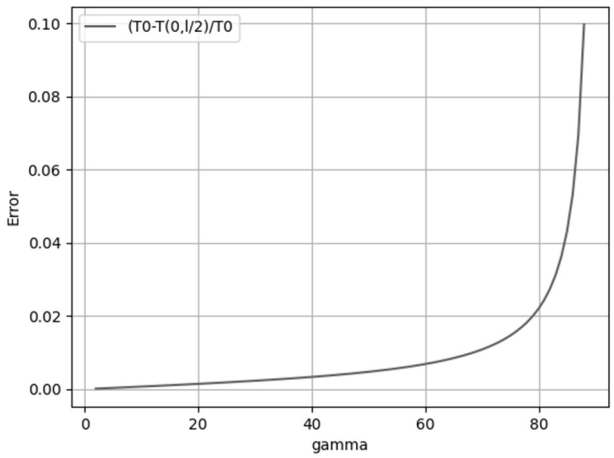

Рис. 6. График относительной ошибки

Fig. 6. Graph of the relative error

На рис. 5 и 6 виден рост абсолютной и относительной ошибки при решении прямой задачи, который может быть связан с несколькими возможными причинами: математическая ошибка из-за неточности модели решения прямой задачи или неточности, связанной с типом данных с плавающей точкой.

3. РЕШЕНИЕ ОБРАТНОЙ КИНЕМАТИЧЕСКОЙ ЗАДАЧИ

При решении обратной кинематической задачи для подобных сред справедлива следующая формула обращения (12) [14], связывающая $n_1(z)$, t, α и l [15]:

$$n_1(z(\alpha)) = \frac{\cos^2\alpha}{2l}\,\frac{d\left(\sin\alpha\,(t(\alpha)-t_0(\alpha))\right)}{d\alpha}. \qquad (12)$$

После проведения математических расчетов строится график изменения показателя коэффициента преломления с изменением глубины, затем график (9) зависимости коэффициента преломления от глубины для линеаризованного и точного решения. Коэффициент B был выбран равным –0,08.

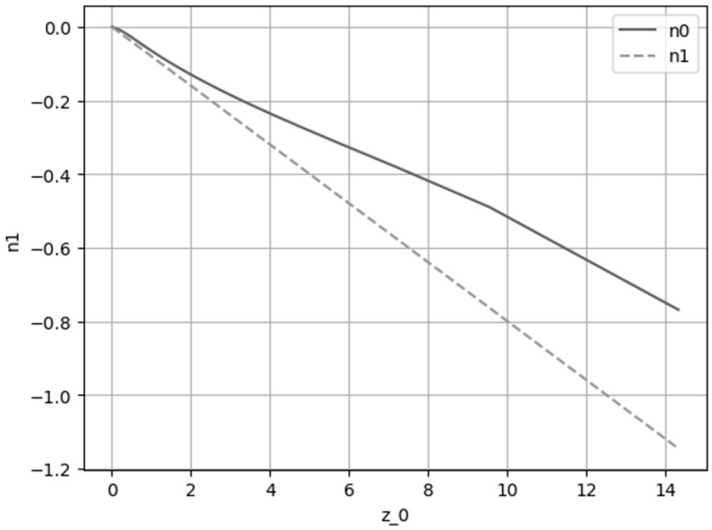

Рис. 7. Зависимость коэффициента преломления от глубины

Fig. 7. Dependence of the refraction coefficient on depth

Показано, что линеаризованное решение отличается от точного вследствие приближения линеаризации. Для оценки качества решения рассчитывается абсолютная ошибка линеаризации при изменении глубины на следующем графике (рис. 8).

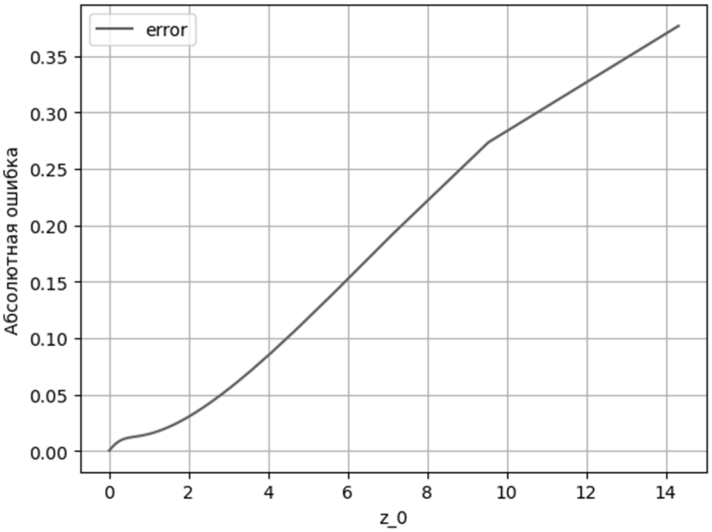

Рис. 8. Абсолютная ошибка линеаризованного решения

*Fig. 8. An a*bsolute error of the linearized solution

Как показано на рис. 8, ошибка начинает расти сильнее с ростом глубины, следовательно, необходимо внести дополнительное регуляризирующее воздействие (например, сглаживание шумов при помощи сплайнов).

4. ПОСТАНОВКА С НЕЛИНЕЙНЫМ ИЗМЕНЕНИЕМ СКОРОСТИ

Рассматривается постановка с нелинейным изменением скорости распространения сигнала от глубины. Изменение скорости сигнала будет выражаться следующей формулой:

$$V(z) = 1 + 0,02z + \sqrt{0,03z^2}. \tag{13}$$

При решении прямой кинематической задачи для нахождения времени пробега сигнала по лучевой траектории использовали численный метод Рунге – Кутты 4-го порядка в решении интегральных уравнений. На рис. 9 показаны лучевые траектории пробега сигнала при различных углах входа сигнала в среду. На данном этапе необходимо выбрать угол входа сигнала в среду (на графиках обозначается как gamma).

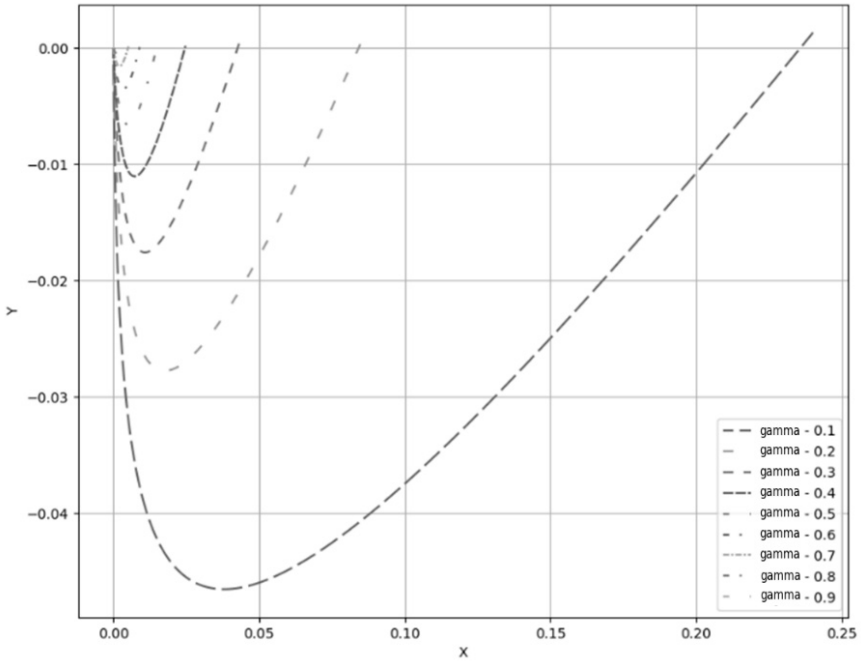

Рис. 9. Лучевые траектории

Fig. 9. Ray trajectories

Для решения ОКЗ угол входа сигнала в среду выбирается равным 0,4. Значение угла взято из соображений наглядности и управляемости, так как при таком значении угла решение получается достаточно устойчивым, но с наличием отклонений, что для подобного рода подходов нормально, особенно без

использования регуляризации. Полученная зависимость коэффициента преломления от глубины представлена на рис. 10.

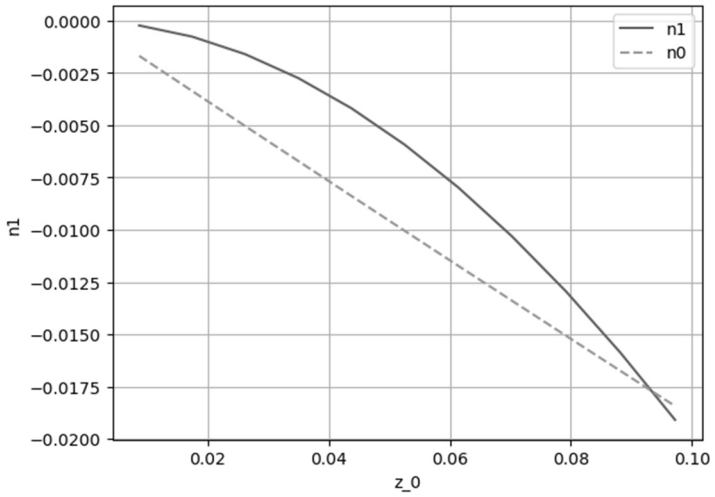

Рис. 10. Зависимость коэффициента преломления от глубины в постановке (13)

Fig. 10. Dependence of the refraction coefficient on depth in formulation (13)

Приведенные на рис. 10 кривые зависимости соотносятся, как и на рис. 7, но сказывается нелинейность скорости распространения, приведенная в (13). Как и в примере с линейной средой, в примере с нелинейной существует ошибка линеаризации, которая накладывается на несоответствие реального скоростного распределения и линеаризованной модели. На рис. 11 приведена зависимость абсолютной ошибки от глубины проникновения сигнала в среду.

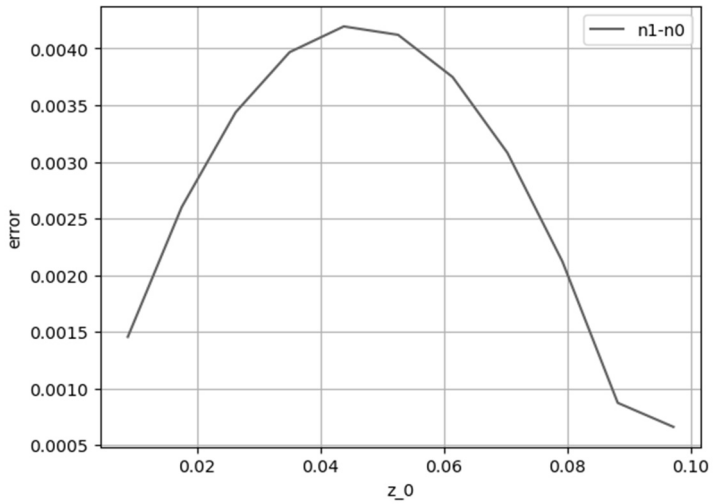

Рис. 11. Абсолютная ошибка

Fig. 11. An absolute error

Как и на рис. 5 и 6, наблюдается ошибка, но с более сложным поведением, причиной которого является группа факторов: изначально заложенная небольшая ошибка при решении прямой задачи, естественная ошибка линеаризации, также возможно влияние типа данных с плавающей точкой (в данном примере при величине ошибки 10^{-4} подобное влияние ощутимо).

ЗАКЛЮЧЕНИЕ

В настоящей работе применен метод линеаризации для решения обратной кинематической задачи с использованием рефрагированных и рассеянных волновых сигналов. В отличие от классических подходов, основанных на прямых или отраженных волнах, использование рассеянных волн расширяет применимость метода на вертикально-неоднородные среды, включая волноводные структуры, где волны «запираются» в слое и распространяются с минимальными потерями. Для этой постановки предложена линеаризация функционала Ферма, сопровождаемая анализом ошибок. Показано, что они возрастают с глубиной, что обусловливает необходимость введения регуляризирующих процедур.

На рис. 7, 8 и 10, 11 показано, что линеаризованное решение имеет отклонение от точного решения при фиксированном угле падения α, что связано с ошибкой линеаризации, которая может быть снижена за счет введения регуляризации.

Таким образом, выстроенный с учетом вышесказанного эвристический подход включает два блока.

1. Формулировка постановки задачи с использованием знаний об обратных задачах и физических основах метода решения, позволяющих построить ее приближенное решение на уровне математических представлений и вывести необходимые формулы.

2. Построение вычислительного алгоритма для реализации конкретного решения и создание программного обеспечения.

Задачи неразрушающего контроля естественным образом опираются на данные физических измерений: в зависимости от схемы эксперимента это могут быть проекции объекта либо сигналы, регистрируемые при изменяющихся условиях зондирования. Подобный класс задач встречается повсеместно – от фундаментальных исследований до инженерной практики, что делает их особенно значимыми. И в этой связи особый интерес представляют обширные классы решений – как широко используемых в приложениях, так и перспективных систем наблюдения с учетом их геометрических особенностей и коллективных свойств, отвечающих за векторы (матрицы) измерений.

Полученные результаты оправдывают продолжение этого направления исследований. Особый интерес представляет распространение метода на более сложные формы функции скоростного распределения, анализ устойчивости к помехам и пределов применимости подхода линеаризации (например, через возможность гибридизации данного метода с моделями машинного обучения и использование таких гибридов в задачах с более неоднородными средами).

СПИСОК ЛИТЕРАТУРЫ

1. *Машошин А.И.* Практические задачи гидроакустики, решаемые с использованием алгоритмов обработки сигналов, согласованных со средой их распространения (обзор) // Фундаментальная и прикладная гидрофизика. – 2017. – Т. 10, № 1. – С. 37–48.

2. *Кузнецова А.В.* Определение скоростной характеристики верхней части сейсмического разреза // Геофизические методы исследования Земли и ее недр: материалы X Международной научно-практической конкурс-конференции молодых специалистов «Геофизика-2015». – СПб., 2016. – С. 68–75.

3. *Смирнов В.И.* Курс высшей математики. Т. 1. – М.: Наука, 1968. – 802 с.

4. *Lavrentev M.M., Zerkal S.M., Trofimov O.E.* Computer modeling in tomography and ill-posed problems. – Utrecht; Boston: VSP, 2001. – 128 p.

5. *Лаврентьев М.М., Романов В.Г., Шишатский С.П.* Некорректные задачи математической физики и анализа. – Новосибирск: Наука, 1980. – 286 с.

6. *Белоносова А.В. Алексеев А.С.* Об одной постановке обратной кинематической задачи сейсмики для двумерной непрерывно-неоднородной среды // Некоторые методы и алгоритмы интерпретации геофизических данных. – М.: Наука, 1967. – С. 137–154.

7. *Машошин А.И., Мельканович В.С., Тимофеев В.Н.* Алгоритм определения глубины погружения источника звука с использованием донной линейной антенны // Гироскопия и навигация. – 2025. – Т. 33, № 1 (128). – С. 146–158.

8. *Trofimov O.E.* On the problem of reconstruction of a function of three variables from its integrals along straight lines crossing a specified curve // Optoelectronics, Instrumentation and Data Processing. – 1991. – N 5. – P. 56–62.

9. *Белоусова О.Н., Зеркаль С.М., Шапошникова Е.В.* Вычислительная томография при нетрадиционных схемах наблюдений. – Новосибирск: Сибстрин, 2010. – 160 с.

10. *Шапошникова Е.В.* Вычислительные методы трехмерной томографической реконструкции в конусе лучей: автореф. дис. … канд. техн. наук. – Новосибирск, 2006. – 18 с.

11. *Бурмин В.Ю.* Обратные кинематические задачи для неоднородных сред // Численные методы в сейсмических исследованиях. – Новосибирск: Наука, 1983. – С. 106–111.

12. *Бурмин В.Ю.* Методы численного решения обратных кинематических задач сейсмики для вертикально-неоднородных сред: автореф. дис. … д-ра физ.-мат. наук. – М., 1987.

13. *Брехорских Л.М., Лысанов Ю.П.* Теоретические основы акустики океана. – М.: Наука, 2007. – 369 с.

14. *Романов В.Г.* Обратные задачи для дифференциальных уравнений. – Новосибирск: НГУ, 1973. – 252 с.

15. *Berkovich L.M.* The method of an exact linearization of n-order ordinary differential equations // Journal of Nonlinear Mathematical Physics. – 1996. – Vol. 3 (3–4). – P. 341–350.

Зеркаль Сергей Михайлович, доктор технических наук, профессор кафедры вычислительной техники, Новосибирский государственный технический университет. Основное направление научных исследований – численные методы и методы искусственного интеллекта в прикладных задачах физики, акустики, сейсмики и неразрушающего контроля. Имеет более 50 печатных работ. E-mail: zerkal@corp.nstu.ru

Путята Евгений Александрович, аспирант кафедры вычислительной техники, Новосибирский государственный технический университет. Основное направление научных исследований – методы искусственного интеллект для решения прикладных задач. E-mail: putyata.2018@stud.nstu.ru

Zerkal Sergey M., D.Sc. (Eng.), Professor at the Department of Computer Engineering, Novosibirsk State Technical University. His main research interests focus on numerical methods and artificial intelligence techniques applied to problems in physics, acoustics, seismology, and non-destructive testing. He has authored more than 50 publications. E-mail: zerkal@corp.nstu.ru

Putyata Evgeny A., Ph.D. Student, Department of Computer Engineering, Novosibirsk State Technical University. His main research interests focus on artificial intelligence methods for solving applied problems. E-mail: putyata.2018@stud.nstu.ru

DOI: 10.17212/2782-2001-2026-3-23-38

Numerical solution of the inverse kinematic problem in a linearized formulation for inhomogeneous media using refractive and scattered waves[*]

S.M. ZERKAL[a], *E.A. PUTYATA*[b]

Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation

[a] *zerkal@corp.nstu.ru* [b] *putyata.2018@stud.nstu.ru*

Abstract

The inverse kinematic problem for vertically inhomogeneous media, in which wave propagation velocity depends only on depth, has been investigated using refracted and scattered wave signals. Unlike classical formulations of inverse kinematic problems, which primarily focus on direct or reflected waves, this study centers on scattered waves – those whose trajectories have been altered due to interaction with local medium inhomogeneities. This approach extends the applicability of the method to media exhibiting waveguide effects, where waves are "trapped" within a layer and propagate over long distances with minimal losses. A computational algorithm based on the linearization of the Fermat functional has been developed, enabling determination of the velocity distribution $V(z)$ in such media. For the direct problem, both analytical solutions (in the linear case $V(z) = 1 + Bz$) and numerical methods (for nonlinear cases), including the 4th-order Runge-Kutta method, have been applied. The implemented solution demonstrated the feasibility of reconstructing the medium velocity distribution from measured wave arrival times, while also revealing errors associated with the inaccuracy of linearization. The analysis showed that linearization error increases with depth, which, in turn, may be compensated by introducing additional regularization procedures – for example, noise smoothing using splines. The obtained results may be of interest for solving problems in seismology, acoustic sounding, and non-destructive testing. Prospects for further research are primarily associated with transitioning to more complex medium models, applying more advanced regularization techniques, and hybridizing this approach with artificial intelligence models – all of which open prospects for expanding the method's applicability and resolution capability.

Keywords: inverse kinematic problem, refractive waves, scattering, computer modeling, linearization of the Fermat functional, waveguide media, numerical methods, vertically inhomogeneous media

REFERENCES

1. Mashoshin A.I. Prakticheskie zadachi gidroakustiki, reshaemye s ispol'zovaniem algoritmov obrabotki signalov, soglasovannykh so sredoi ikh rasprostraneniya (obzor) [Underwater acoustics problems solving with using matched field processing]. *Fundamental'naya i prikladnaya gidrofizika = Fundamental and Applied Hydrophysics*, 2017, vol. 10, no. 1, pp. 37–48.

2. Kuznetsova A.V. [Determination of velocity characteristics of a near surface seismic section by diving waves method]. *Geofizicheskie metody issledovaniya Zemli i ee nedr* [Geophysical methods of survey the Earth and its subsoil]. Materials of the X International scientific and technical conference-competition of young specialists "Geophysics-2015", St. Petersburg, 2016, pp. 68–75. (In Russian).

3. Smirnov V.I. *Kurs vysshei matematiki*. T. 1 [A course of higher mathematics. Vol. 1]. Moscow, Nauka Publ., 1968. 802 p.

4. Lavrentev M.M., Zerkal S.M., Trofimov O.E. *Computer modeling in tomography and ill-posed problems*. Utrecht, Boston, VSP, 2001. 128 p.

5. Lavrent'ev M.M., Romanov V.G., Shishatskii S.P. *Nekorrektnye zadachi matematicheskoi fiziki i analiza* [Ill-posed problems of mathematical physics and analysis]. Novosibirsk, Nauka Publ., 1980. 286 p.

[*] *Received 10 March 2026.*

6. Belonosova A.V. Alekseev A.S. Ob odnoi postanovke obratnoi kinematicheskoi zadachi seismiki dlya dvumernoi nepreryvno-neodnorodnoi sredy [On one formulation of the inverse kinematic problem of seismology for a two-dimensional continuously inhomogeneous medium]. *Nekotorye metody i algoritmy interpretatsii geofizicheskikh dannykh* [Some methods and algorithms for interpretation of geophysical data]. Moscow, Nauka Publ., 1967, pp. 137–154.

7. Mashoshin A.I., Mel'kanovich V.S., Timofeev V.N. Algoritm opredeleniya glubiny pogruzheniya istochnika zvuka s ispol'zovaniem donnoi lineinoi antenny [Algorithm for determining the depth of a sound source using a bottom linear antenna]. *Giroskopiya i navigatsiya = Gyroscope and Navigation*, 2025, vol. 33, no.1 (128), pp. 146–158.

8. Trofimov O.E. On the problem of reconstruction of a function of three variables from its integrals along straight lines crossing a specified curve. *Optoelectronics, Instrumentation and Data Processing*, 1991, no. 5, pp. 56–62.

9. Belousova O.N., Zerkal S.M., Shaposhnikova E.V. *Vychislitel'naya tomografiya pri netraditsionnykh skhemakh nablyudenii* [Computational tomography under non-traditional observation schemes]. Novosibirsk, Sibstrin Publ., 2010. 160 p.

10. Shaposhnikova E.V. *Vychislitel'nye metody trekhmernoi tomograficheskoi rekonstruktsii v konuse luchei*. Avtoref. diss. kand. tekhn. nauk [Computational methods of three-dimensional tomographic reconstruction in a cone of rays]. Novosibirsk, 2006. 18 p.

11. Burmin V.Yu. Obratnye kinematicheskie zadachi dlya neodnorodnykh sred [Inverse kinematic problems for inhomogeneous media]. *Chislennye metody v seismicheskikh issledovaniyakh* [Numerical methods in seismic research]. Novosibirsk, Nauka Publ., 1983, pp. 106–111.

12. Burmin V.Yu. *Metody chislennogo resheniya obratnykh kinematicheskikh zadach seismiki dlya vertikal'no-neodnorodnykh sred*. Avtoref. diss. dokt. fiz.-mat. nauk [Methods of numerical solution of inverse kinematic problems of seismology for vertically inhomogeneous media. Author's abstract of Dr. phys. and math. sci. diss.]. Moscow, 1987.

13. Brekhorskikh L.M., Lysanov Yu.P. *Teoreticheskie osnovy akustiki okeana* [Theoretical foundations of ocean acoustics]. Moscow, Nauka Publ., 2007. 369 p.

14. Romanov V.G. *Obratnye zadachi dlya differentsial'nykh uravnenii* [Inverse problems for differential equations]. Novosibirsk, NSU Publ., 1973. 252 p.

15. Berkovich L.M. The method of an exact linearization of n-order ordinary differential equations. *Journal of Nonlinear Mathematical Physics*, 1996, vol. 3 (3–4), pp. 341–350.

Для цитирования:

Зеркаль С.М., Путята Е.А. Численное решение обратной кинематической задачи в линеаризованной постановке для неоднородных сред с использованием рефрагированных и рассеянных волн // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 23–38. – DOI: 10.17212/2782-2001-2026-3-23-38.

For citation:

Zerkal S.M., Putyata E.A. Chislennoe reshenie obratnoi kinematicheskoi zadachi v linearizovannoi postanovke dlya neodnorodnykh sred s ispol'zovaniem refragirovannykh i rasseyannykh voln [Numerical solution of the inverse kinematic problem in a linearized formulation for inhomogeneous media using refractive and scattered waves]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 23–38. DOI: 10.17212/2782-2001-2026-3-23-38.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 39–56

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 39–56

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 517.977.5

DOI: 10.17212/2782-2001-2026-3-39-56

Схема придания автоматическим системам с широтно-импульсной модуляцией 1–5-го рода свойства суперпозиции*

В.В. КУЛИКОВ[1,a], Н.Н. КУЦЫЙ[2,b], Е.А. ОСИПОВА[2,c]

[1] *664074, г. Иркутск, ул. Чернышевского, 15, Иркутский государственный университет путей сообщения*
[2] *664074, г. Иркутск, ул. Лермонтова, 83, Иркутский национальный исследовательский технический университет*

[a] *godefire@mail.ru* [b] *kucyinn@mail.ru* [c] *osipovaelizaveta@yandex.ru*

В работе рассматривается схема придания свойства суперпозиции автоматической системе, использующей регулятор с переменной структурой и содержащей импульсный элемент (ИЭ) с широтно-импульсной модуляцией 1–5-го рода. Представлены результаты численных экспериментов для широтно-импульсной модуляции (ШИМ) 1–5-го рода, подтверждающие наличие свойства суперпозиции. Применение логических элементов для переключения параметров в типовых ПИ-регуляторах в автоматических системах с запаздыванием позволяет повысить скорость протекания переходного процесса, используя при этом в законе переключения только знак производной от ошибки системы. Ограничением такой схемы является то, что она применима только к линейному объекту управления и регулятору с переменной структурой, в котором выполняется принцип суперпозиции и содержится интегральный компонент, необходимый для реализации последовательности импульсов, которая должна исчезать при нулевой ошибке системы. Рассматривается случай, когда значение периода повторения импульсов не изменяется. В основе исследования лежит адаптация по задающему воздействию, связанная с реализацией бинарных систем, описанных в работах С.В. Емельянова и С.К. Коровина. Использование только критериев оптимизации, учитывающих специфику нелинейного ИЭ, не позволяет решить проблему сохранения времени регулирования при разных величинах задающего воздействия. Положительным моментом предложенного технического решения является то, что оно не содержит настраиваемых параметров. В качестве модельного примера выбрана автоматическая система поддержания температуры в туннельной печи для обжига кирпича. На ней показано, что при изменении величины постоянного задающего воздействия время нарастания и время регулирования переходного процесса автоматической системы практическим не изменяется.

Ключевые слова: транспортное запаздывание, автоматическая система, суперпозиция, широтно-импульсная модуляция, регулятор с переменной структурой, адаптация, импульсный элемент, управление, алгоритм, структурный синтез

---

* *Статья получена 01 ноября 2025 г.*

ВВЕДЕНИЕ

При синтезе автоматических систем возникают ситуации, когда в линейную систему автоматического управления добавляются элементы с ШИМ для повышения ее помехозащищенности [1, 2]. Вместе с тем появление такой нелинейности в системе приводит к нарушению свойства суперпозиции, и, следовательно, в этой автоматической системе нельзя сделать так, чтобы время окончания переходного процесса не зависело от величины задающего воздействия [2]. В работах [3–5] описан способ линеаризации автоматической системы с ШИМ-элементом через эквивалентность сигналов с линейным амплитудно-импульсным элементом для исследования динамических режимов таких систем. В настоящей работе сделан акцент на сохранении свойства суперпозиции нелинейной дискретной автоматической системы при изменении задающего воздействия.

Цель настоящей работы заключается в наделении свойством суперпозиции автоматических систем, содержащих дискретные элементы с импульсным принципом действия и переключением параметров [6–10], путем использования адаптации по задающему воздействию в ИЭ.

1. ПОСТАНОВКА ЗАДАЧИ

Структурная схема рассматриваемой автоматической системы представлена на рис. 1.

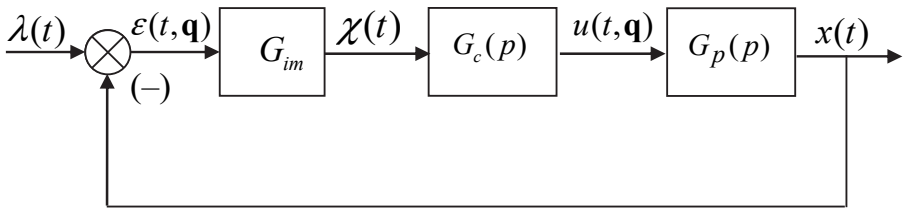

Рис. 1. Структурная схема автоматической системы

Fig. 1. Block diagram of the automatic system

Описание процессов в данной автоматической системе с ПИ-регулятором с переменной структурой [6, 7] и нелинейным ИЭ имеет следующий вид:

$$\varepsilon(t,\mathbf{q}) = \lambda(t) - x(t),$$

$$\chi(t) = G_{im}(\varepsilon(t,\mathbf{q})), \tag{1}$$

$$u(t,\mathbf{q}) = \chi(t)\left(\sum_{i=0}^{n-1} q_{2i+1}\varphi_{i+1}(t)\right) + \frac{\chi(t)}{p}\left(\sum_{i=1}^{n} q_{2i}\varphi_i(t)\right),$$

$$(\varphi_1(t)=1,\ \ \varphi_2(t)=0,\ \ \varphi_3(t)=0) \ \text{ при } \ \Psi^{+}=\text{true},$$

$$(\varphi_1(t)=0,\ \ \varphi_2(t)=1,\ \ \varphi_3(t)=0) \ \text{ при } \ \Psi^{-}=\text{true},$$

$$(\varphi_1(t)=0,\ \ \varphi_2(t)=0,\ \ \varphi_3(t)=0)\ \text{ при }\ \Psi^0=\text{true},$$

$$\Psi^+=((\Psi(t,\ \varepsilon(t,\mathbf{q}))>0)\vee(0<t<\tau_{ob}))\wedge(|\varepsilon(t,\mathbf{q})|>\alpha\lambda(t)),$$

$$\Psi^-=(\Psi(t,\ \varepsilon(t,\mathbf{q}))<0)\wedge(|\varepsilon(t,\mathbf{q})|>\alpha\lambda(t)),$$

$$\Psi^0=(|\varepsilon(t,\mathbf{q})|<\alpha\lambda(t)),$$

$$x(t)=G_p(p)u(t,\mathbf{q}),$$

где $\mathbf{q}=(q_1,...,\ q_n)$ – вектор настраиваемых параметров пропорционально-интегрального регулятора с переменной структурой (ПИ – РПС); $\varepsilon(t,\mathbf{q})$ – ошибка системы; $\varphi_1(t),...,\ \varphi_n(t)$ – функции включения, принимающие значение «0» или «1» в зависимости от значений логических выражений Ψ^-, Ψ^+, Ψ^0; $\Psi(t,\ \varepsilon(t,\mathbf{q}))=\varepsilon(t,\mathbf{q})\dfrac{p\varepsilon(t,\mathbf{q})}{T_f p+1}$ – функция переключения с реальным дифференцирующим звеном; T_f – параметр реального дифференцирующего звена, определяется после получения $\mathbf{q}$ путем увеличения его значения до тех пор, пока не изменятся моменты переключения структур ПИ – РПС; n = 3 – количество структур ПИ – РПС; $\lambda(t)$ – ступенчатое задающее воздействие; $\chi(t)$ – выходная координата ИЭ; $u(t,\mathbf{q})$ – выходная координата ПИ – РПС; $x(t)$ – выходная координата автоматической системы; $G_p(p)$ – оператор объекта; $G_c(p)$ – оператор ПИ – РПС; G_{im} – оператор ИЭ (ШИМ с 1-го по 5-й род); $p=d/dt$ – оператор дифференцирования; τ_{ob} – время запаздывания объекта; λ – параметр, определяющий зону нечувствительности ПИ – РПС; $\vee$, $\wedge$ – логические операции дизъюнкции и конъюнкции соответственно.

Выражения, позволяющие моделировать G_{im} при используемых видах модуляций с 1-го по 5-й род [1], можно представить в следующем виде:

ШИМ-1

$$\chi(t)=\begin{cases}\text{sgn}(\varepsilon(t,\mathbf{q})),\ |\varepsilon(t,\mathbf{q})|>\delta_1(t);\\ 0,\ |\varepsilon(t,\mathbf{q})|<\delta_1(t);\end{cases}\qquad(2)$$

ШИМ-2

$$\chi(t)=\begin{cases}\text{sgn}(\varepsilon(t,\mathbf{q})),\ |\varepsilon[kT,\mathbf{q}]|>\delta_1(t);\\ 0,\ |\varepsilon[kT,\mathbf{q}]|<\delta_1(t);\end{cases}\qquad(3)$$

ШИМ-3

$$\chi(t)=\begin{cases}\text{sgn}(\varepsilon(t,\mathbf{q})),\ f(t,\ \varepsilon(t,\mathbf{q}))=\text{true};\\ 0,\ f(t,\ \varepsilon(t,\mathbf{q}))=\text{false};\end{cases}\qquad(4)$$

$$f(t,\ \varepsilon(t,\mathbf{q})) = (|\varepsilon(t,\mathbf{q})| > \delta_2(t)) \vee$$

$$\vee\left(\left(|\varepsilon[t,\mathbf{q}]|_{\delta_2(t)=|\varepsilon(t,\mathbf{q})|}\right| > \delta_1(t)\right) \wedge \left(\left|\varepsilon[t,\mathbf{q}]|_{\delta_2(t)=|\varepsilon(t,\mathbf{q})|}\right| < \delta_2(t)\right)\right).$$

ШИМ-4

$$\chi(t) = \begin{cases} \operatorname{sgn}(\varepsilon(t,\mathbf{q})), & \dfrac{\int\limits_{kT}^{kT+t_k} |\varepsilon(t,\mathbf{q})|\,dt}{t_k} > \delta_1(t); \\ 0, & \dfrac{\int\limits_{kT}^{kT+t_k} |\varepsilon(t,\mathbf{q})|\,dt}{t_k} < \delta_1(t); \end{cases} \tag{5}$$

ШИМ-5

$$\chi(t) = \begin{cases} \operatorname{sgn}(\varepsilon(t,\mathbf{q})), & \dfrac{\int\limits_{kT}^{(k+1)T} |\varepsilon(t,\mathbf{q})|\,dt}{T} > \delta_1(t); \\ 0, & \dfrac{\int\limits_{kT}^{(k+1)T} |\varepsilon(t,\mathbf{q})|\,dt}{T} < \delta_1(t), \end{cases} \tag{6}$$

где T – период повторения ИЭ; $\delta_1(t)$ – функция, реализующая пилообразные импульсы с периодом повторения T и амплитудой A_1; $\delta_2(t)$ – функция, реализующая пилообразные импульсы с периодом повторения T и амплитудой A_2 (далее $A_2 = 2A_1$); $\vee, \wedge$ – логические операции дизъюнкции и конъюнкции соответственно; t_k – длительность импульса на k-м такте.

Выше в описании импульсных элементов используются опорные функции $\delta_1(t)$ и $\delta_2(t)$ с постоянными максимальными значениями. Этот факт препятствует сохранению свойства суперпозиции в обозначенных системах при изменении величины задающего воздействия $\lambda(t)$, влияющего на последовательность значений ошибки системы $\varepsilon(t,\mathbf{q})$. Ставится задача использовать вычислительную схему, которая позволит в ШИМ-1, ШИМ-2, ШИМ-3, ШИМ-4, ШИМ-5 сохранять последовательность значений ошибки системы $\varepsilon(t,\mathbf{q})$, подаваемых в ИЭ. Предлагаемое решение не гарантирует сохранения свойства суперпозиции в автоматической системе при изменении задающего воздействия $\lambda(t)$ с периодом повторения T и требует проведения отдельного исследования для этого случая.

2. СХЕМА ПРИДАНИЯ СВОЙСТВА СУПЕРПОЗИИ НЕЛИНЕЙНОЙ ИМПУЛЬСНОЙ АВТОМАТИЧЕСКОЙ СИСТЕМЕ

Так как используется построение импульсного элемента на основе сравнения модуля ошибки системы $\varepsilon(t, \mathbf{q})$, зависящей от задающего воздействия $\lambda(t)$ (1), с опорной функцией, то прослеживается следующая закономерность: с увеличением задающего воздействия $\lambda(t)$ увеличивается ошибка системы $\varepsilon(t, \mathbf{q})$ и количество тактов, когда ширина импульса совпадает с периодом *T*. Это вносит нелинейность типа насыщения в автоматическую систему (1). Вместе с тем стоит отметить, что согласно работе [10] переключения параметров в регуляторе не влияют на свойство суперпозиции.

Для соблюдения свойства суперпозиции в автоматической системе (1) предлагается использовать следующее выражение, которое основывается на делении и умножении сигналов при построении бинарных автоматических систем [11] (рис. 2) и заключается в том, что суммарная величина насыщения рассматриваемых ИЭ не меняется в зависимости от величины задающего воздействия:

$$\chi(t) = \begin{cases} |\lambda(t)|\,\mathrm{sgn}\,(\varepsilon(t, \mathbf{q})),\ f\left(\dfrac{|\varepsilon(t, \mathbf{q})|}{|\lambda(t)|},\ t\right) = \mathrm{true}; \\ 0,\ f\left(\dfrac{|\varepsilon(t, \mathbf{q})|}{|\lambda(t)|},\ t\right) = \mathrm{false}; \end{cases} \tag{7}$$

$$(k = 0,\ 1,\ 2,...),$$

где $f(t)$ – логическое выражение, при истинности которого действуют импульсы.

В качестве примера рассмотрим выражение (7) для сохранения последовательностей значений ошибки системы $\varepsilon(t, \mathbf{q})$ на временном интервале $t* \in [0,\ t_{ob}]$ при задающих воздействиях $\lambda_1(t)$ и $\lambda_2(t)$:

$$\frac{\left|\varepsilon(t*, \mathbf{q})\,|_{\lambda_1(t*)} = \lambda_1(t*) - x(t*)\,|_{\lambda_1(t*)}\right|}{\left|\lambda_1(t*)\right|} = \frac{\left|\varepsilon(t*, \mathbf{q})\,|_{\lambda_2(t*)} = \lambda_2(t*) - x(t*)\,|_{\lambda_2(t*)}\right|}{\left|\lambda_2(t*)\right|},$$

$$x(t*)\,|_{\lambda_1(t*)} = 0, \quad x(t*)\,|_{\lambda_2(t*)} = 0,$$

$$\frac{\left|\varepsilon(t*, \mathbf{q})\,|_{\lambda_1(t*)} = \lambda_1(t*)\right|}{\left|\lambda_1(t*)\right|} = \frac{\left|\varepsilon(t*, \mathbf{q})\,|_{\lambda_2(t*)} = \lambda_2(t*)\right|}{\left|\lambda_2(t*)\right|}, \tag{8}$$

$$\left|\varepsilon(t*, \mathbf{q})\,|_{\lambda_1(t*)}\right| = \left|\varepsilon(t*, \mathbf{q})\,|_{\lambda_2(t*)}\right| = 1.$$

Так как используется линейный оператор объекта $G_p(p)$ и ПИ – РПС, не нарушающий свойства суперпозиции, то временной интервал $[0,t_{ob}]$ в (8) может быть увеличен до времени регулирования в автоматической системе (1).

В выражении (7) при $\lambda(t)=0$ появляется неопределенность, которая исключается путем отключения схемы (7) для этого значения $\lambda(t)$:

$$\chi(t)=\begin{cases}\operatorname{sgn}(\varepsilon(t,\mathbf{q})),\ f(|\varepsilon(t,\mathbf{q})|,\ t)=\text{true};\\ 0,\ f(|\varepsilon(t,\mathbf{q})|,\ t)=\text{false};\end{cases} \tag{9}$$

$$(k=0,\ 1,\ 2,...).$$

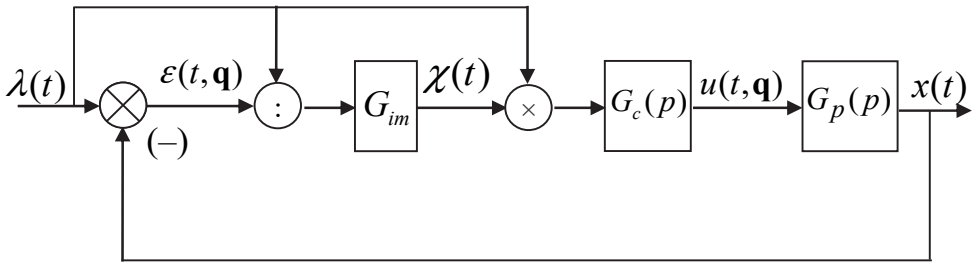

Рис. 2. Схема адаптации входного/выходного воздействия импульсного элемента G_{im} по задающему воздействию $\lambda(t)$ для одноканальной автоматической системы

Fig. 2. Scheme of adaptation of the input/output effects of the G_{im} pulse element according to the master effect of $\lambda(t)$ for a single-channel automatic system

Представим ниже вид выражений (2) – (6) с учетом (7):
ШИМ-1

$$\chi(t)=\begin{cases}|\lambda(t)|\operatorname{sgn}(\varepsilon(t,\mathbf{q})),\ \dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}>\delta_1(t);\\ 0,\ \dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}<\delta_1(t);\end{cases} \tag{10}$$

ШИМ-2

$$\chi(t)=\begin{cases}|\lambda(t)|\operatorname{sgn}(\varepsilon(t,\mathbf{q})),\ \dfrac{|\varepsilon[kT,\mathbf{q}]|}{|\lambda(t)|}>\delta_1(t);\\ 0,\ \dfrac{|\varepsilon[kT,\mathbf{q}]|}{|\lambda(t)|}<\delta_1(t);\end{cases} \tag{11}$$

ШИМ-3

$$\chi(t)=\begin{cases}|\lambda(t)|\operatorname{sgn}(\varepsilon(t,\mathbf{q})),\ f\left(t,\ \dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}\right)=\text{true};\\ 0,\ f\left(t,\ \dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}\right)=\text{false};\end{cases} \tag{12}$$

$$f\left(t,\ \frac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}\right)=\left(\frac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}>\delta_2(t)\right)\vee$$

$$\vee\left(\left(\frac{\left|\varepsilon[t,\mathbf{q}]|_{\delta_2(t)=|\varepsilon(t,\mathbf{q})|}\right|}{|\lambda(t)|}>\delta_1(t)\right)\wedge\left(\frac{\left|\varepsilon[t,\mathbf{q}]|_{\delta_2(t)=|\varepsilon(t,\mathbf{q})|}\right|}{|\lambda(t)|}<\delta_2(t)\right)\right);$$

ШИМ-4

$$\chi(t)=\begin{cases}|\lambda(t)|\operatorname{sgn}(\varepsilon(t,\mathbf{q})),\ \dfrac{\int\limits_{kT}^{kT+t_k}\dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}dt}{t_k}>\delta_1(t);\\ 0,\ \dfrac{\int\limits_{kT}^{kT+t_k}\dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}dt}{t_k}<\delta_1(t);\end{cases} \tag{13}$$

ШИМ-5

$$\chi(t)=\begin{cases}|\lambda(t)|\operatorname{sgn}(\varepsilon(t,\mathbf{q})),\ \dfrac{\int\limits_{kT}^{(k+1)T}\dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}dt}{T}>\delta_1(t);\\ 0,\ \dfrac{\int\limits_{kT}^{(k+1)T}\dfrac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|}dt}{T}<\delta_1(t).\end{cases} \tag{14}$$

Отметим сложность реализации импульсной модуляции 5-го рода, которая заключается в том, что тем или иным образом необходимо использовать экстраполяцию ошибки системы для получения ее среднего значения на текущем такте работы импульсного элемента. Так как указанная задача выходит за рамки настоящей работы, предлагается использовать среднее значение ошибки системы на предыдущем такте $(k-1)T$, $k=1, 2,\ldots$. С учетом сделанного

замечания выражение (14) изменится, и дальше будем его использовать в численных экспериментах:

$$\Phi_{k-1} = \frac{\int\limits_{(k-1)T}^{kT} \frac{|\varepsilon(t,\mathbf{q})|}{|\lambda(t)|} dt}{T}; \tag{15}$$

ШИМ-5

$$\chi(t) = \begin{cases} |\lambda(t)| \operatorname{sgn}(\varepsilon(t,\mathbf{q})), & \Phi_{k-1} > \delta_1(t); \\ 0, & \Phi_{k-1} < \delta_1(t). \end{cases} \tag{16}$$

Предложенная выше схема придания свойства суперпозиции может быть применена к таким ШИМ-элементам, в которых ширина импульсов определяется полиномом *n*-й степени от модуля ошибки системы $\varepsilon(t,\mathbf{q})$ с дополнительными настраиваемыми параметрами, если при изменении задающего воздействия $\lambda(t)$ площадь импульсов ШИМ-элемента будет пропорционально изменяться по отношению к $\lambda(t)$.

3. ЧИСЛЕННЫЙ ПРИМЕР

Далее показано применение схемы придания свойства суперпозиции автоматической системе (1) с оператором объекта $G_p(p)$, который описывает процесс нагрева в туннельной печи для обжига кирпича [12–15]:

$$\varepsilon(t,\mathbf{q}) = \lambda(t) - x(t),$$

$$\chi(t) = G_{im}(\varepsilon(t,\mathbf{q})), \tag{17}$$

$$u(t,\mathbf{q}) = \chi(t)\left(\sum_{i=0}^{2} q_{2i+1}\varphi_{i+1}(t)\right) + \frac{\chi(t)}{p}\left(\sum_{i=1}^{3} q_{2i}\varphi_i(t)\right),$$

$$x(t) = G_p(p)u(t,\mathbf{q}),$$

$$G_p(p) = \frac{k_{ob}}{T_{ob1}T_{ob2}p^2 + (T_{ob1}+T_{ob2})p + 1} e^{-p\tau_{ob}},$$

$$T_{ob1} = 5{,}2\text{ с}, \quad T_{ob2} = 0{,}64\text{ с}, \quad k_{ob} = 0{,}75, \quad \tau_{ob} = 1{,}8\text{ с}.$$

Параметры ПИ – РПС получены с помощью алгоритма из работы [16] и штрафной функции [17]:

$$\mathbf{q} = (0{,}017340;\ 0{,}144287;\ 0{,}081546;\ 0{,}043054). \tag{18}$$

На примере ШИМ-1 демонстрируется (рис. 3) влияние нелинейности типа насыщения на время первого пересечения кривой разгона автоматической

системы (17) и кривой задающего воздействия без схемы придания свойства суперпозиции.

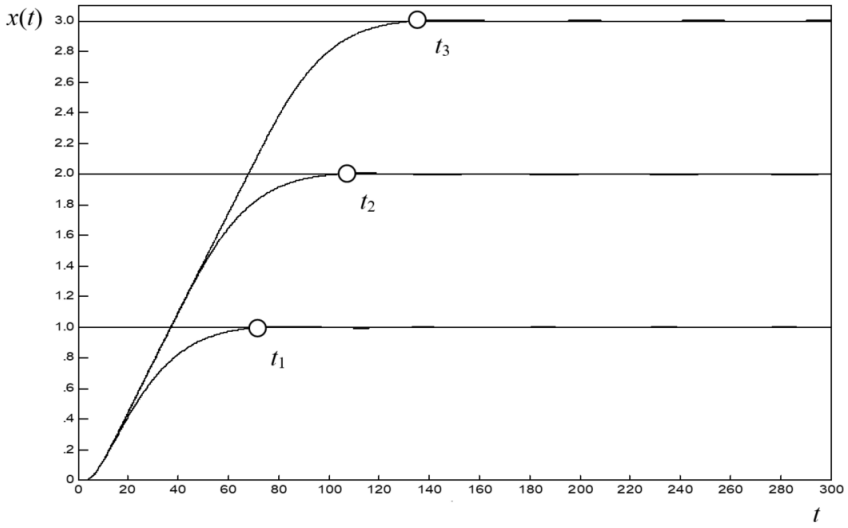

Рис. 3 Переходные процессы в автоматической системе (17) при использовании ШИМ-1 без схемы придания свойства суперпозиции (период квантования $T = 4$ с). Задающее воздействие $\lambda(t)$: $1(t)$, $2(t)$, $3(t)$

Fig. 3 Transients in the automatic system (17) when using PWM-1 without a superposition property scheme (quantization period $T = 4$ s). Setting effect $\lambda(t)$: $1(t)$, $2(t)$, $3(t)$

Анализ временных точек, отмеченных как t_1, t_2, t_3 на графике рис. 3, показывает, что при увеличении величины задающего воздействия с $1(t)$ до $3(t)$ время t_3 увеличивается почти в два раза по отношению к t_1.

На рис. 4–7 показаны графики внутренних процессов автоматической системы (17) с задающим воздействием $1(t)$.

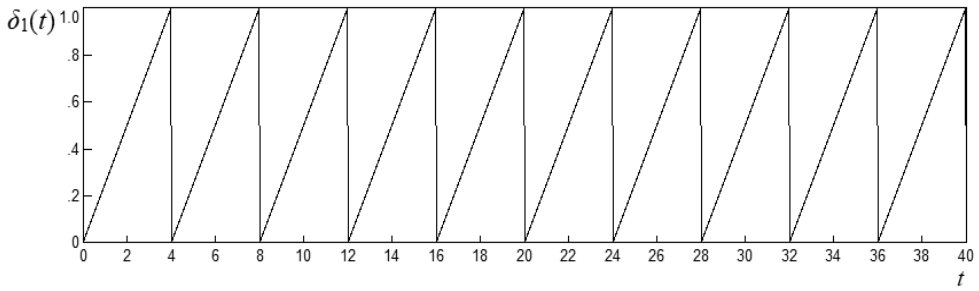

Рис. 4. График $\delta_1(t)$ функции, реализующей пилообразные импульсы с периодом повторения $T = 4$ с и амплитудой $A = 1$

Fig. 4. Graph $\delta_1(t)$ of a function implementing sawtooth pulses with a repetition period of $T = 4$ s and an amplitude of $A = 1$

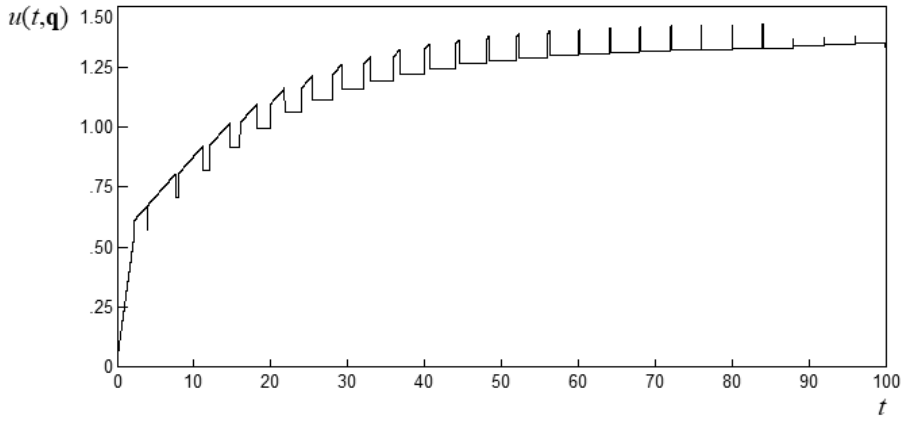

Рис. 5. График выходной координаты ПИ – РПС $u(t, \mathbf{q})$ (17)

Fig. 5. Graph of the output coordinate of the PI – CVS $u(t, \mathbf{q})$ (17)

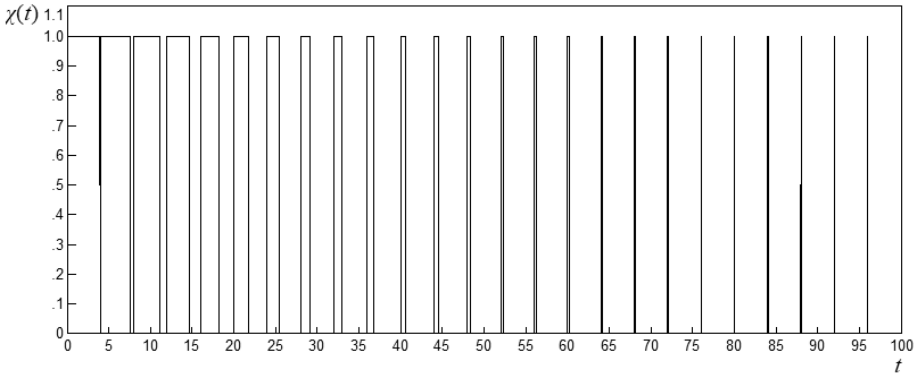

Рис. 6. График выходной координаты импульсного элемента $\chi(t)$ (17)

Fig. 6. Graph of the output coordinate of the pulse element $\chi(t)$ (17)

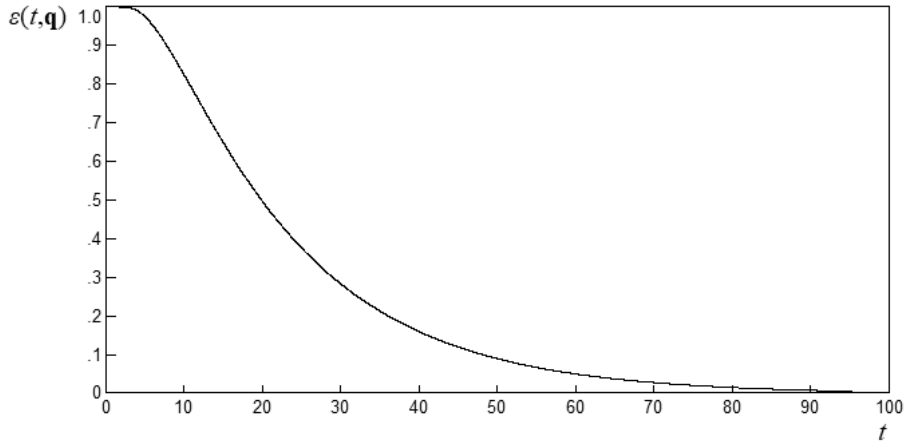

Рис. 7. График ошибки системы $\varepsilon(t, \mathbf{q})$ (17)

Fig. 7. Graph of the system error $\varepsilon(t, \mathbf{q})$ (17)

Далее демонстрируется предложенная схема придания системе (17) свойства суперпозиции при использовании ШИМ 1–5-го рода при различных задающих воздействиях $\lambda(t)$: $1(t)$, $2(t)$, $3(t)$.

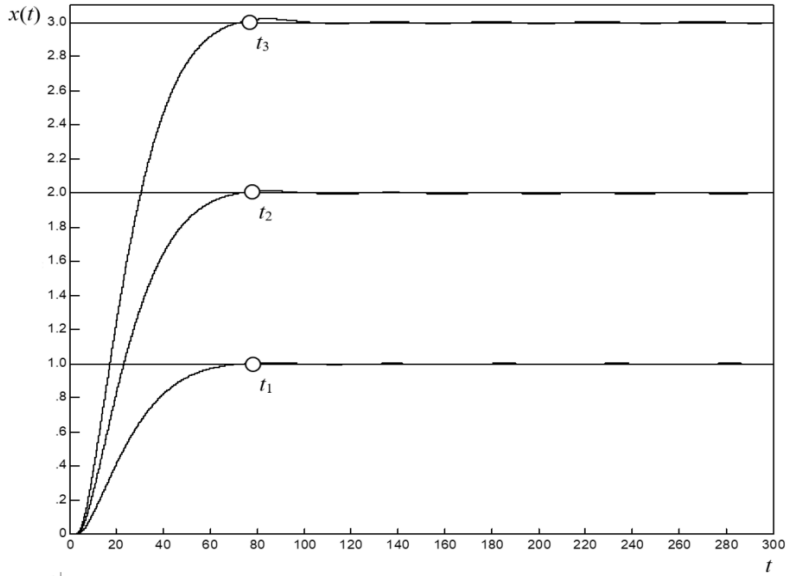

Рис. 8. Переходные процессы в автоматической системе (17) при использовании ШИМ-1 со схемой придания свойства суперпозиции (период квантования $T = 4$ с)

Fig. 8. Transients in the automatic system (17) using PWM-1 with a scheme for imparting superposition properties (quantization period $T = 4$ s)

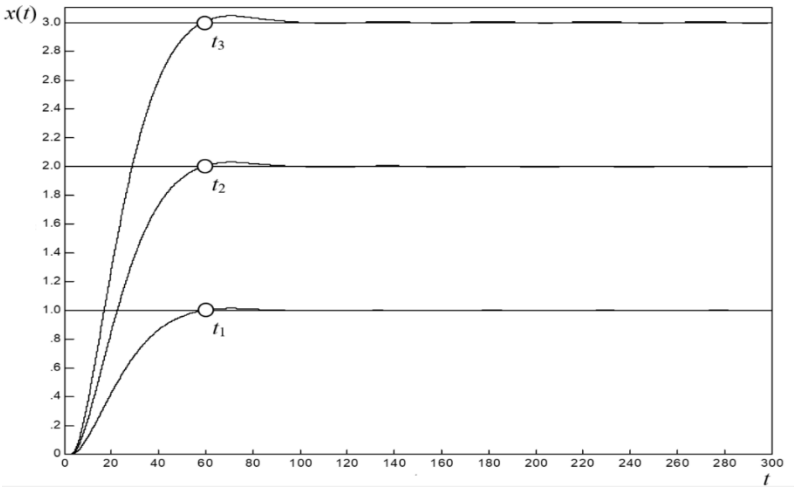

Рис. 9. Переходные процессы в автоматической системе (17) при использовании ШИМ-2 со схемой придания свойства суперпозиции (период квантования $T = 4$ с)

Fig. 9. Transients in the automatic system (17) using PWM-2 with a scheme for imparting superposition properties (quantization period $T = 4$ s)

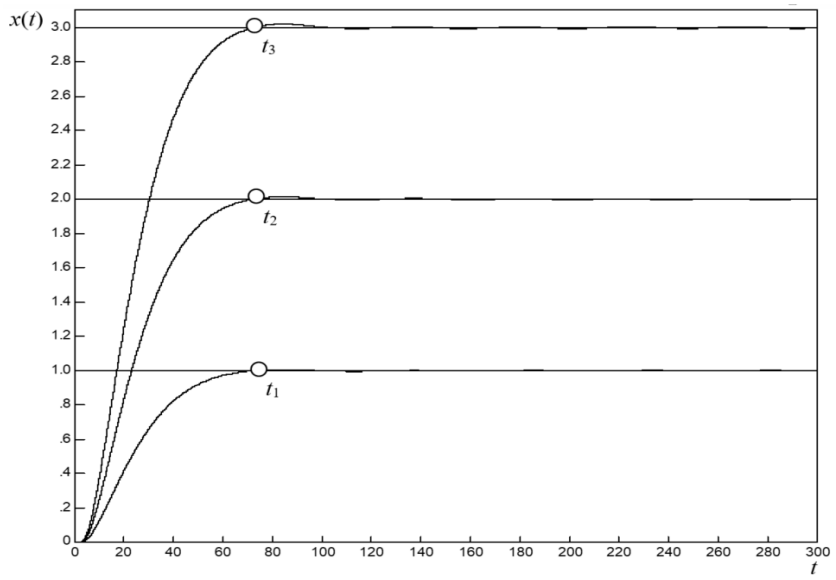

Рис. 10. Переходные процессы в автоматической системе (17) при использовании ШИМ-3 со схемой придания свойства суперпозиции (период квантования $T = 4$ с)

Fig. 10. Transients in the automatic system (17) using PWM-3 with a scheme for imparting superposition properties (quantization period $T = 4$ s)

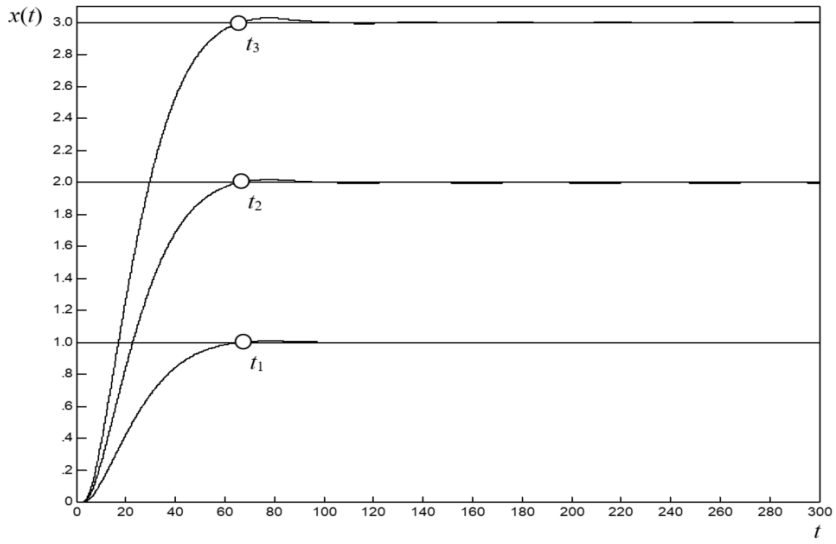

Рис. 11. Переходные процессы в автоматической системе (17) при использовании ШИМ-4 со схемой придания свойства суперпозиции (период квантования $T = 4$ с)

Fig. 11. Transients in the automatic system (17) using PWM-4 with a scheme for imparting superposition properties (quantization period $T = 4$ s)

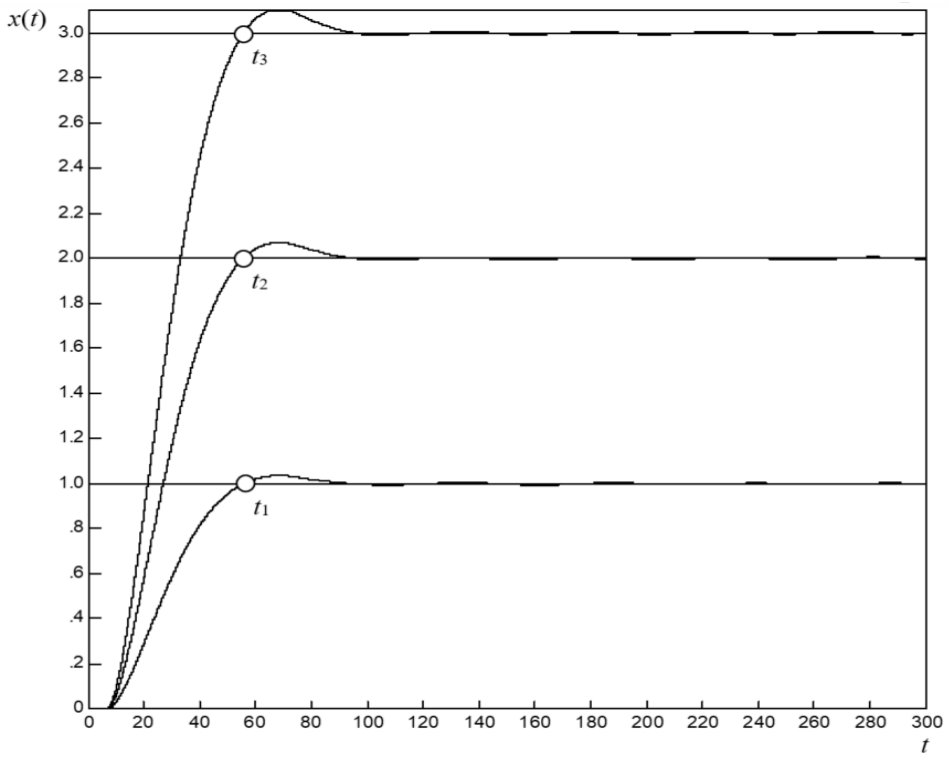

Рис. 12. Переходные процессы в автоматической системе (17) при использовании ШИМ-5 со схемой придания свойства суперпозиции (период квантования $T = 4$ с)

Fig. 12. Transients in the automatic system (17) using PWM-5 with a scheme for impartingsuperposition properties (quantization period $T = 4$ s)

Приведенные на рис. 8–12 графики переходных процессов системы регулирования температуры в туннельной печи для обжига кирпича (17) свидетельствуют о выполнении принципа суперпозиции в рассматриваемой системе с нелинейным ИЭ и регулятором с переключаемыми параметрами. В частности, видно, что моменты времени первого пересечения кривой разгона с соответствующей кривой задающего воздействия совпадают, т. е. $t_1 = t_2 = t_3$, и время окончания переходного процесса не зависит от величины задающего воздействия. Заметим, что в установившемся режиме присутствуют периодические колебания, которые обусловлены наличием двух нелинейностей: насыщением в ИЭ и зоной нечувствительности в ПИ – РПС.

Для демонстрации свойства устойчивости автоматической системы (17) на рис. 13 показаны переходные процессы с ШИМ с 1-го по 5-й род при последовательной смене задающего воздействия $\lambda(t)$ с разным знаком.

Анализируя семейство переходных процессов на рис. 13, можно увидеть, что переходный процесс при ШИМ-3 существенно отличается от остальных переходных процессов потому, что при реализации ШИМ-3 используются две опорные функции, в отличие от других рассматриваемых ШИМ.

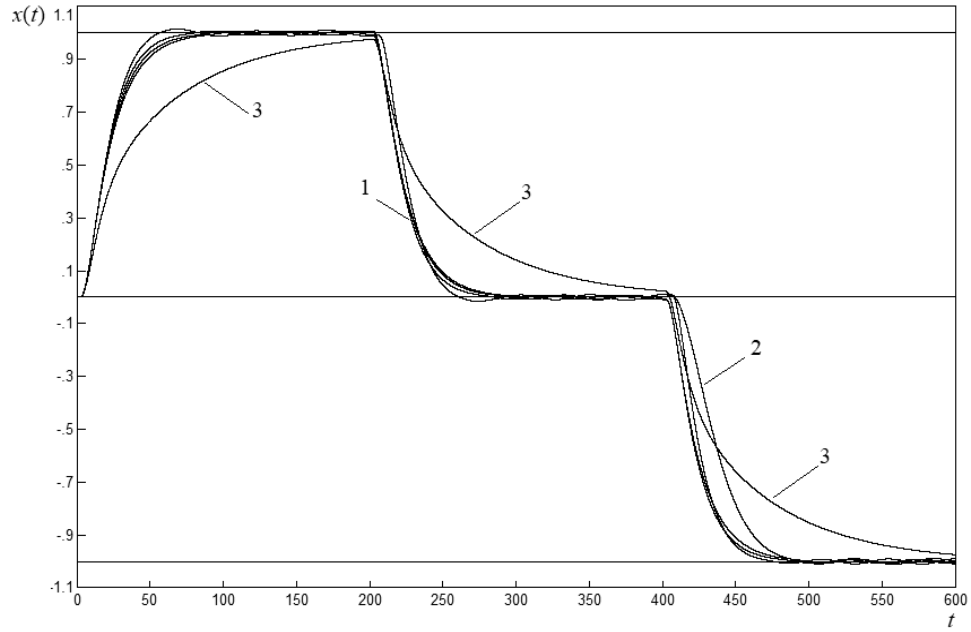

Рис. 13. Переходные процессы в автоматической системе (17) (*1* – ШИМ-1, ШИМ-2, ШИМ-4, ШИМ-5; *2* – ШИМ-5; *3* – ШИМ-3) при последовательной смене задающего воздействия $\lambda(t)$: $1(t) \to 0(t) \to -1(t)$ (период квантования $T = 4$ с)

Fig. 13. Transient processes in the automatic system (17) (*1* – PWM-1, PWM-2, PWM-4, PWM-5; *2* – PWM-5; *3* – PWM-3) with successive changes in the reference signal $\lambda(t)$: $1(t) \to 0(t) \to -1(t)$ (quantization period $T = 4$ s)

Изменение максимальных значений опорных функции в ШИМ-3 влияет на качество переходного процесса.

ЗАКЛЮЧЕНИЕ

Научная новизна заключается в том, что предлагаемая схема придания автоматической системе свойства суперпозиции не зависит от рода модуляции, используемой в ШИМ, и позволяет не осуществлять подстройку настраиваемых параметров автоматической системы, содержащей нелинейный элемент, при изменении задающего воздействия. Результаты численного моделирования автоматической системы поддержания температуры в туннельной печи для обжига кирпича показали, что схема адаптации по задающему воздействию в ШИМ позволяет сохранять время нарастания и время регулирования переходного процесса при увеличении задания. Предложенная схема придания свойства суперпозиции импульсной автоматической системе без дополнительных преобразований может быть применена к таким ШИМ-элементам, в которых ширина импульсов определяется полиномом n-й степени от модуля ошибки системы $\varepsilon(t, \mathbf{q})$ с дополнительными настраиваемыми параметрами, если площадь импульсов ИЭ находится в прямой зависимости от задающего воздействия. В рассматриваемой автоматической системе относительно небольшую погрешность в получении суперпозиции вносит зона нечувствительности, используемая в регуляторе с переменной структурой для реализации

режима конечных переключений структур такого регулятора. Эту погрешность можно компенсировать, если величину зоны нечувствительности определять в процентах от задающего воздействия в автоматической системе.

СПИСОК ЛИТЕРАТУРЫ

1. *Слепов Н.Н., Дроздов Б.В.* Широтно-импульсная модуляция: (анализ и применение в магнитной записи). – М.: Энергия, 1978. – 191 с.

2. *Попков Ю.С.*, *Цыпкин Я.З.* Теория нелинейных импульсных систем. – М.: Наука, 1973. – 414 с.

3. *Лурье М.С.*, *Лурье О.М.*, *Фролов А.С.* Исследование динамических режимов систем стабилизации тока мощных электромагнитов с широтно-импульсной модуляцией // Мехатроника, автоматизация, управление. – 2021. – Т. 22, № 6. – С. 313–320. – DOI: 10.17587/mau.22.313-320.

4. *Andeen P.E.* Analysis of pulse duration sampled data systems with linear elements // IRE Transactions on Automatic Control. – 1960. – Vol. 5 (4). – P. 306–313.

5. *Nease R.E.* Analysis and design of nonlinear sampled date control systems // Massachusetts Institute of Technology, Servomechanisms Laboratory. Wright Air Development Center (WADC) Technical Note. – Cambridge, June 1957. – P. 57–162.

6. *Шигин Е.К.* Автоматическое регулирование объекта с чистым запаздыванием регулятором с переключаемыми параметрами. II // Автоматика и телемеханика. – 1966. – № 6. – С. 72–81.

7. *Говоров А.А.* Зонные регуляторы с переменной структурой. Принципы построения, методы анализа и способы реализации. – Саарбрюккен: LAP Lambert Academic Publishing, 2014. – 124 с.

8. *Sun* Z., *Ge S.S.* Switched linear systems: Control and design. – Springer, 2005. – (Communications and control engineering).

9. *Емельянов С.В.* Системы автоматического управления с переменной структурой. – М.: Наука, 1967. – 336 с.

10. *Жильцов К.К.* Приближенные методы расчета систем с переменной структурой. – М.: Энергия, 1974. – 224 с.

11. *Емельянов С.В.*, *Фурсов А.С.* Координатно-операторная обратная связь. Свойства. Особенности. Перспективы // Автоматика и телемеханика. – 2015. – № 10. – С. 3–39.

12. *Васильков Ю.В.*, *Василькова Н.Н.* Математическое моделирование объектов и систем автоматического управления: учебное пособие. – М.; Вологда: Инфра-Инженерия, 2020. – 428 с.

13. *Митришкин Ю.В.* Линейные математические модели динамических систем с управлением. – М.: Ленанд, 2019. – 464 с.

14. *Сабанчин В.Р.*, *Тугов В.В.* Автоматизация процесса обжига керамического кирпича: монография. – Оренбург: ОГУ, 2019. – 150 с.

15. *Tseligorov N.A.*, *Chubukin A.V.*, *Tseligorova E.N.* Investigation of the robust absolute stability of the tunnel kiln control system with delay // WSEAS Transactions on Systems. – 2021. – Vol. 20. – P. 25–30. – DOI: 10.37394/23202.2021.20.4.

16. *Куликов В.В.*, *Куцый Н.Н.*, *Куцый А.П.* Градиентный алгоритм параметрической оптимизации ПИ-регулятора с переменной структурой с зоной нечувствительности // Мехатроника, автоматизация, управление. – 2020. – Т. 21, № 9. – С. 530–534. – DOI: 10.17587/mau.21.530-534.

17. *Куликов В.В.*, *Куцый Н.Н.*, *Осипова Е.А.* Параметрическая оптимизация ПИД-регулятора с ограничением на основе метода сопряженных градиентов Полака – Поляка – Рибьера // Мехатроника, автоматизация, управление. – 2023. – Т. 24, № 5. – С. 240–248. – DOI: 10.17587/mau.24.240-248.

Куликов Владимир Валерьевич, соискатель кафедры «Автоматизация производственных процессов» Иркутского государственного университета путей сообщения. Основное направление научных исследований – параметрическая оптимизация автоматических

систем с переключаемыми параметрами. Имеет более 15 печатных работ. E-mail: godefire@mail.ru

Куцый Николай Николаевич, доктор технических наук, профессор Института информационных технологий и анализа данных Иркутского национального исследовательского технического университета. Основное направление научных исследований – автоматическая параметрическая оптимизация сложных дискретных систем, линейное программирование, многоиндексные и многокритериальные задачи. Имеет более 200 печатных работ. E-mail: kucyinn@mail.ru

Осипова Елизавета Алексеевна, кандидат технических наук, доцент Института информационных технологий и анализа данных Иркутского национального исследовательского технического университета. Основное направление научных исследований – параметрическая оптимизация дискретных систем, методы анализа данных и нейросетевое моделирование. Имеет более 30 печатных работ. E-mail: osipovaelizaveta@yandex.ru

Kulikov Vladimir V., degree ccandidate at the Department of Automation of Industrial Processes, Irkutsk State University of Railway Transport. The main area of his scientific research is parametric optimization of automatic systems with switchable parameters. He has more than 15 publications. E-mail: godefire@mail.ru

Kutsyy Nikolay N., D.Sc. (Eng.), professor at the Institute of Information Technology and Data Analysis, Irkutsk National Research Technical University. The main research area is automatic parametric optimization of complex discrete systems, linear programming, multi–index and multi-criteria tasks. He has more than 200 printed works. E-mail: kucyinn@mail.ru

Osipova Elizaveta A., PhD (Eng.), associate professor in the Institute of Information Technology and Data Analysis, Irkutsk National Research Technical University. The main area of her scientific research is parametric optimization of discrete systems, data analysis methods and neural network modeling. She has more than 30 printed works. E-mail: osipovaelizaveta@yandex.ru

DOI: 10.17212/2782-2001-2026-3-39-56

Scheme for imparting superposition property to automatic systems with pulse-width modulation of the 1st – 5th kind[*]

V.V. KULIKOV[1,a], *N.N. KUTSYY*[2,b], *E.A. OSIPOVA*[2,c]

[1] *Irkutsk State Transport University, 15, Chernyshevsky Street, Irkutsk, 664074, Russian Federation*

[2] *Irkutsk National Research Technical University, 83, Lermontov Street, Irkutsk, 664074, Russian Federation*

[a] *godefire@mail.ru* [b] *kucyinn@mail.ru* [c] *osipovaelizaveta@yandex.ru*

Abstract

The paper proposes and considers a scheme for imparting the superposition property to an automatic system using a regulator with a variable structure and containing a pulse element (IE) with pulse width modulation of the 1st-5th kind. The results of numerical experiments for pulse width modulation (PWM) of the 1st-5th kind are presented confirming the presence of the superposition property. The use of logic gates for parameter switching in typical PI controllers in automatic systems with a delay makes it possible to increase the speed of the transient process, while using only the sign of the derivative of the system error in the switching law. The limitation of this scheme is that it is applicable only to a linear control object and a regulator with a variable structure which fulfills the principle of superposition and contains an integral component necessary to implement a pulse sequence that should disappear with a zero-system error. The case is

[*] *Received 01 November 2025.*

considered when the value of the pulse repetition period does not change. The study is based on adaptation according to the setting effect which is ideologically related to the implementation of binary systems described in the works by S.V. Yemelyanov and S.K. Korovin. Using only optimization criteria that take into account the specifics of nonlinear IE does not solve the problem of saving control time with different values of the setting effect. The positive aspect of the proposed technical solution is that it does not contain configurable parameters. An automatic temperature control system in a tunnel kiln for firing bricks is chosen as a model example. It shows that when the magnitude of the constant setting effect changes, the rise time and the time for regulating the transient process of the automatic system practically do not change.

Keywords: transport delay, automatic system, superposition, pulse width modulation, variable structure regulator, adaptation, pulse element, control, algorithm, structural synthesis

REFERENCES

1. Slepov N.N., Drozdov B.V. *Shirotno-impul'snaya modulyatsiya: (analiz i primenenie v magnitnoi zapisi)* [Pulse width modulation: (analysis and application in magnetic recording)]. Moscow, Energiya Publ., 1978. 191 p.

2. Popkov Yu.S., Tsypkin Ya.Z. *Teoriya nelineinykh impul'snykh sistem* [Theory of nonlinear impulse systems]. Moscow, Nauka Publ., 1978. 414 p.

3. Lur'e M.S., Lur'e O.M., Frolov A.S. Issledovanie dinamicheskikh rezhimov sistem stabilizatsii toka moshchnykh elektromagnitov s shirotno-impul'snoi modulyatsiei [Study of dynamic modes of current stabilization systems of powerful electromagnets with pulse-width modulation]. *Mekhatronika, avtomatizatsiya, upravlenie = Mechatronics, Automation, Control*, 2021, vol. 22 (6), pp. 313–320. DOI: 10.17587/mau.22.313-320.

4. Andeen P.E. Analysis of pulse duration sampled data systems with linear elements. *IRE Transactions on Automatic Control*, 1960, vol. 5 (4), pp. 306–313.

5. Nease R.E. Analysis and design of nonlinear sampled date control systems. Massachusetts Institute of Technology, Servomechanisms Laboratory. *Wright Air Development Center (WADC) Technical Note*. Cambridge, June 1957, pp. 57–162.

6. Shigin E.K. Avtomaticheskoe regulirovanie ob"ekta s chistym zapazdyvaniem regulyatorom s pereklyuchaemymi parametrami. II [Automatic control of a plant with pure time delay and a regulator with switching parameters. II]. *Avtomatika i telemekhanika = Automation and Remote Control*, 1966, no. 6, pp. 72–81. (In Russian).

7. Govorov A.A. *Zonnye regulyatory s peremennoi strukturoi. Printsipy postroeniya, metody analiza i sposoby realizatsii* [Zone regulators with variable structure. Principles of construction, methods of analysis and methods of implementation]. Saarbryukken, LAP Lambert Academic Publishing, 2014. 124 p.

8. Sun Z., Ge S.S. *Switched linear systems: Control and design*. *Communications and control engineering*. Springer, 2005.

9. Emel'yanov S.V. *Sistemy avtomaticheskogo upravleniya s peremennoi strukturoi* [Automatic control systems with variable structure]. Moscow, Nauka Publ., 1967. 336 p.

10. Zhil'tsov K.K. *Priblizhennye metody rascheta sistem s peremennoi strukturoi* [Approximate methods for calculating systems with variable structure]. Moscow, Energiya Publ., 1974. 224 p.

11. Emel'yanov S.V., Fursov A.S. Coordinate-to-operator feedback: Properties, features, outlooks. *Automation and Remote Control*, 2015, vol. 76 (10), pp. 1713–1742. DOI: 10.1134/S000511791510001X. Translated from *Avtomatika i telemekhanika*, 2015, no. 10, pp. 3–39.

12. Vasil'kov Yu.V., Vasil'kova N.N. *Matematicheskoe modelirovanie ob"ektov i sistem avtomaticheskogo upravleniya* [Mathematical modeling of objects and automatic control systems]. Moscow, Vologda, Infra-Inzheneriya Publ., 1974. 428 p.

13. Mitrishkin Yu.V. *Lineinye matematicheskie modeli dinamicheskikh sistem s upravleniem* [Linear mathematical models of dynamic systems with control]. Moscow, Lenand Publ., 2019. 464 p.

14. Sabanchin V.R., Tugov V.V. *Avtomatizatsiya protsessa obzhiga keramicheskogo kirpicha* [Automation of the ceramic brick firing process: monograph]. Orenburg, OGU Publ., 2019. 150 p.

15. Tseligorov N.A., Chubukin A.V., Tseligorova E.N. Investigation of the robust absolute stability of the tunnel kiln control system with delay. *WSEAS Transactions on Systems*, 2021, vol. 20, pp. 25–30. DOI: 10.37394/23202.2021.20.4.

16. Kulikov V.V., Kutsyi N.N., Kutsyi A.P. Gradientnyi algoritm parametricheskoi optimizatsii PI-regulyatora s peremennoi strukturoi s zonoi nechuvstvitel'nosti [The gradient-based algorithm for parametric optimization of a variable-structure PI controller with dead band]. *Mekhatronika, avtomatizatsiya, upravlenie = Mechatronics, Automation, Control*, 2020, vol. 21 (9), pp. 530–534. DOI: 10.17587/mau.21.530-534.

17. Kulikov V.V., Kutsyi N.N., Osipova E.A. Parametricheskaya optimizatsiya PID regulyatora s ogranicheniem na osnove metoda sopryazhennykh gradientov Polaka – Polyaka – Rib'era [Parametric optimization of the PID controller with restriction based on the method of conjugate gradients of Polak – Polak – Ribier]. *Mekhatronika, avtomatizatsiya, upravlenie = Mechatronics, Automation, Control*, 2023, vol. 24 (5), pp. 240–248. DOI: 10.17587/mau.24.240-248.

Для цитирования:

Куликов В.В., Куцый Н.Н., Осипова Е.А. Схема придания автоматическим системам с широтно-импульсной модуляцией 1–5-го рода свойства суперпозиции // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 39–56. – DOI: 10.17212/2782-2001-2026-3-39-56.

For citation:

Kulikov V.V., Kutsyy N.N., Osipova E.A. Skhema pridaniya avtomaticheskim sistemam s shirotno-impul'snoi modulyatsiei 1–5 roda svoistva superpozitsii [Scheme for imparting superposition property to automatic systems with pulse-width modulation of the 1st – 5th kind]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 39–56. DOI: 10.17212/2782-2001-2026-3-39-56.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 57–74

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 57–74

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.246.8

DOI: 10.17212/2782-2001-2026-3-57-74

Последовательное применение метода Прони для повышения точности прогнозирования временных рядов*

И.А. КУЗЬМИНЫХ[1,a], А.Б. ГАВРИЛОВ[1,b], А.С. ТОЛСТИКОВ[2,c], В.М. ТИССЕН[2,d]

[1] *РФ, 630073, г. Новосибирск, пр. К. Маркса, 20, Новосибирский государственный технический университет*
[2] *РФ, 630004, г. Новосибирск, пр. Димитрова, 4, Западно-Сибирский филиал ФГУП ВНИИФТРИ*

[a] *i.kuzminyx.2019@stud.nstu.ru* [b] *a.gavrilov@corp.nstu.ru* [c] *tolstikov@sniim.ru*
[d] *tissen.ksn@yandex.ru*

Работа посвящена модификации метода Прони для повышения точности прогнозирования колебательных временных рядов. В основе предложенного алгоритма лежит многоэтапный подход: модель последовательно аппроксимирует остаточные компоненты ряда, которые затем интегрируются в итоговую функцию. Модификация была практически применена на задаче прогнозирования временных рядов, в структуре которых доминируют сложные колебательные процессы и гармонические составляющие: параметров вращения Земли, а именно отклонений оси вращения x_p и y_p, по архиву данных Международной службы вращения Земли и систем отсчета (IERS, International Earth Rotation Service) за 2025 год. Сравнительный анализ с исходным методом Прони, а также с прогнозами IERS и ЗСФ ВНИИФТРИ на интервалах 7 и 90 суток показал, что эффективность модификации существенно зависит от заданных условий. Предложенный подход не во всех случаях превосходит существующие решения IERS и ЗСФ ВНИИФТРИ, однако для определенных горизонтов прогнозирования последовательное уточнение модели обеспечивает локальное повышение точности и снижение погрешности прогнозов. Выявленная неоднородность результатов указывает на то, что эффективность последовательного применения метода Прони существенно зависит от заданных параметров модели. Полученные результаты открывают возможности для применения метода в системах автономной навигации спутников. Проведенный анализ позволил обозначить границы применимости текущей модели и сформулировать рекомендации по ее развитию. В частности, обосновывается необходимость перехода от использования статически заданных параметров к разработке механизмов их адаптивной настройки, что позволит стабилизировать результаты метода при изменении характера исходных данных.

* *Статья получена 25 января 2026 г.*

Ключевые слова: прогнозирование, временные ряды, цифровая обработка измерительной информации, авторегрессионные модели, метод Прони, оптимизация параметров алгоритма, гармонический анализ, параметры вращения Земли, отклонение оси вращения Земли, точность прогноза

ВВЕДЕНИЕ

Прогнозирование представляет собой методологическую основу для принятия обоснованных решений в условиях неопределенности. Использование методов прогнозирования в таких областях, как энергетика, транспорт, машиностроение или телекоммуникации, помогает своевременно выявлять потенциальные проблемы, строить производственные планы и рационально распределять ресурсы. Широкий спектр практических задач требует применения различных подходов к прогнозированию, чтобы обеспечить инструменты для работы в самых разных сферах. Во многих прикладных задачах исходные данные представлены в виде временных рядов (ВР), используемых для дальнейшего анализа и прогнозирования.

Для прогнозирования ВР применяется ряд методов: метод наилучшей прямой аппроксимации, метод авторегрессии и модели со скользящим средним (Autoregressive Integrated Moving Average, ARIMA) [1], метод моделей с изменением дисперсии [2], метод моделей с изменяющимся параметром и сезонными компонентами (Seasonal Auto-Regressive Integrated Moving Average, SARIMA) [3], метод сингулярного спектрального анализа (Singular Spectrum Analysis, SSA) [4], фильтрация Калмана, вейвлет-преобразование [5], каждый из которых имеет свои преимущества и ограничения. Несмотря на разнообразие доступных инструментов, выбор подходящего метода зависит от специфики исходных данных и поставленных задач, что делает изучение их особенностей особенно важным для достижения точных прогнозов.

В настоящей статье рассматривается возможность использования метода гармонического разложения Прони как альтернативный подход к прогнозированию. Этот метод предназначен для анализа ВР с колебательными явлениями и предполагает наличие в исходных данных экспоненциально затухающих гармоник и случайных компонент [6]. Реализация метода состоит в определении параметров модели Прони, наилучшим образом аппроксимирующей исходный ВР. После определения параметров модель Прони может быть использована для исследования спектральной структуры ВР, а также для прогнозирования его поведения в будущем. Также метод Прони предоставляет возможность получать и использовать информацию о составляющих ВР в частотной и временной областях.

Возможности подхода исследуются на примере прогнозирования параметров вращения Земли (ПВЗ). Наличие периодических составляющих в сочетании с нестационарным поведением делает задачу прогнозирования ПВЗ особенно сложной. Точные прогнозы ПВЗ критически важны для обеспечения работы спутниковых навигационных систем, преобразования координат между небесной и земной системами отсчета, а также управления орбитальными группировками в периоды отсутствия связи с наземными станциями.

1. ПАРАМЕТРЫ ВРАЩЕНИЯ ЗЕМЛИ

Параметры вращения Земли представляют собой набор математических величин, определяющих ориентацию мгновенного положения оси вращения Земли в инерциальном звездном пространстве. К ним относятся всемирное время *UT1*, отклонения оси вращения Земли x_p и y_p, параметры прецессии и нутации. В случаях, когда прямые определения всемирного времени невозможны, вычисляется эксцесс длительности суток [7]. Следует отметить, что знание параметров отклонения оси вращения позволяет определить координаты полюса вращения на поверхности Земли. Совокупность данных ПВЗ обеспечивает точную связь между земной и небесной системами отсчета. Объектом исследования настоящей статьи являются ВР, вызываемые перемещениями углового положения северного полюса относительно Международного условного начала (МУН). Временные ряды отклонений оси вращения демонстрируют выраженные колебательные компоненты различной природы: от регулярных годового периода и группы Чандлеровских колебаний с наиболее значимым периодом порядка 1,18 года до высокочастотных суточных и лунных периодов [8].

Изменения ПВЗ, в том числе отклонения оси вращения, формируются под воздействием сложного комплекса геофизических процессов, включающих гравитационное влияние Луны и Солнца, атмосферно-океанические взаимодействия, тектонические движения и перераспределение масс внутри Земли. Существенный вклад также вносят климатические факторы, такие как таяние ледников и глобальный гидрологический цикл, что делает ВР ПВЗ сложными нестационарными сигналами с выраженными колебательными компонентами. Данные о ПВЗ собираются с помощью наземных и спутниковых наблюдений и аккумулируются Международной службой вращения Земли и систем отсчета (International Earth Rotation and Reference Systems Service, IERS) [9]. В базе данных IERS представлены значения полученных из обработки спутниковых и других средств измерений данных ПВЗ с 1962 года с шагом в один день, а также в ежедневном режиме прогнозы со сроками от недели до трех месяцев и до одного года, обновляемые раз в неделю.

Прогнозирование ПВЗ в зависимости от характера решаемых задач проводится на различные глубины. Краткосрочный прогноз с глубиной в 7 суток применяется для решения оперативных задач высокоточной навигации, в которых требуется корректное преобразование координат между небесной и земной системами отсчета. Среднесрочный прогноз, охватывающий период до 90 суток, направлен на обеспечение автономного функционирования космических аппаратов в условиях отсутствия связи с наземными станциями управления. Долгосрочные прогнозы от года и более используются при построении согласующих моделей изменений климата и динамики внутриземных геофизических процессов [10].

2. СУЩЕСТВУЮЩИЕ МЕТОДЫ ПРОГНОЗИРОВАНИЯ

В указанных выше источниках, а также в [1, 11] представлен широкий спектр методов прогнозирования, однако в контексте практического применения для ПВЗ целесообразно сосредоточиться на подходах, используемых

профильными организациями, прежде всего IERS [12]. Сравнительный анализ методов прогнозирования ПВЗ традиционно проводится в рамках международных кампаний, координируемых IERS. В ходе указанных кампаний методика, разработанная Западно-Сибирским филиалом Всероссийского научно-исследовательского института физико-технических и радиотехнических измерений (ЗСФ ВНИИФТРИ) [13], продемонстрировала конкурентоспособные результаты по сравнению с альтернативными подходами [14]. В связи с этим для обеспечения объективной оценки качества предлагаемых прогнозов в настоящей работе их результаты сопоставляются с данными, полученными IERS и ЗСФ ВНИИФТРИ.

2.1. МЕТОД АВТОРЕГРЕССИИ И МОДЕЛИ СО СКОЛЬЗЯЩИМ СРЕДНИМ

Для прогнозирования IERS использует комбинированные методы, основанные на статистических моделях авторегрессии (АР) и скользящего среднего (СС), а также на аналитических моделях вращения и полярного движения [12].

Метод авторегрессии скользящего среднего (АРСС) является одним из классических подходов к анализу и прогнозированию ВР. Он сочетает в себе две ключевые компоненты: авторегрессию и скользящее среднее. Математически модель АРСС (p, q) состоит из АР модели порядка p (т. е. количество предыдущих значений ВР, используемых для прогнозирования текущего значения). А также модель АРСС (p, q) представляет СС-модель порядка q (т. е. количество предыдущих ошибок прогнозирования, учитываемых в модели).

Объединяя эти две части, модель АРСС (p, q) принимает вид

$$x_t = c + a_1 x_{t-1} + \ \dots \ + a_p x_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \ \dots \ + b_q \varepsilon_{t-q}, \qquad (1)$$

где x_t – текущее значение ряда в момент времени t; c – константа смещения; $a_1, \dots, a_p$ – коэффициенты АР; $b_1, \dots, b_p$ – коэффициенты СС; ε_t – текущая ошибка.

Кроме того, IERS комбинируют модель АРСС с теоретическими моделями колебаний, океанических приливов, свободной нутации ядра и среднее положение ядра. На этой основе строится прогностическая модель IERS. Именно такой подход лежит в основе оперативных прогнозов ПВЗ, выполняемых центрами IERS.

2.2. ПРОГНОЗИРОВАНИЕ С ПОМОЩЬЮ АДАПТИВНЫХ ГАРМОНИЧЕСКИХ МОДЕЛЕЙ

Ключевое отличие подхода заключается в том, что для построения прогностической модели используются временные интервалы разной протяженности, что позволяет учитывать как медленные, так и быстрые вариации исследуемого процесса. Для трендовых изменений параметры модели оцениваются по данным длиной до 100 лет, тогда как для быстропеременных и сезонных составляющих применяются интервалы от 6 до 40 лет [7]. Такое разделение делает модель адаптивной, поскольку каждая составляющая подстраивается под характер динамики соответствующего диапазона временных

масштабов. Оставшиеся после аппроксимации ВР тригонометрическими моделями остаточные колебания моделируются АР-моделью, что позволяет учитывать компоненты ВР, связанные с атмосферными и океаническими процессами. Модель авторегрессии выглядит следующим образом:

$$X_{pr,i} = X_{rel,0} - X_{pra,0} + X_{pra,i} + T_{pp,i} + T_{trn,i} + T_{sez,i}\,, \qquad (2)$$

где $X_{pr,i}$ – прогноз X на i-й момент времени; $X_{rel,0}$ и $X_{pra,0}$ – начальное наблюдение и его АР-аппроксимация для коррекции смещения; $X_{pra,i}$ – прогноз псевдослучайной составляющей (авторегрессия); $T_{pp,i}$, $T_{trn,i}$ и $T_{sez,i}$ – поправки на океанические приливы, тренд и сезонность соответственно.

Для учета псевдослучайных колебаний $X_{pra,i}$ используется авторегрессионный анализ.

3. МЕТОД ПРОНИ

Метод Прони – численный метод, предложенный французским математиком Гаспаром де Прони в 1795 году для аппроксимации сигнала суммой экспоненциальных функций по нескольким измерениям при решении практической задачи об испарении смеси жидкостей [15]. Важной особенностью исходной формулировки метода было равенство количества гармонических компонент числу доступных отсчетов сигнала, что существенно ограничивало его практическое применение при работе с реальными данными. В последнее время практическое использование метода развивается быстрыми темпами в связи с совершенствованием средств вычислительной техники [16]. Современные модификации этого подхода распространены на модель, в которой слагаемыми ряда являются затухающие синусоиды. Именно такая модель и рассмотрена в настоящей статье.

В основе метода лежит представление исходного сигнала как суммы экспоненциально затухающих синусоид с определенными параметрами: амплитудами, частотами, коэффициентами затухания и фазами:

$$\hat{x}(t) = \sum_{i=1}^{p} A_i \sin\left(f_i t + \theta_i\right) e^{\alpha_i t}\,, \qquad (3)$$

где временная последовательность $\hat{x}(t)$ характеризуется суммой компонент p (часто используется термин «порядок модели p»), параметров A_i, α_i, θ_i и f_i. В выражении (1) A_i – начальные амплитуды в тех же единицах измерения, что и $\hat{x}(t)$; α_i – коэффициенты затухания, с^{-1}; θ_i – начальные фазы, рад; f_i – частоты, Гц.

Аппроксимация отсчетов ВР происходит путем нахождения значений этих параметров. Применение метода основано на предположении, что многие ВР могут быть представлены в виде (3). Такой подход особенно полезен при обработке данных, имеющих выраженную периодическую структуру, в том числе ВР ПВЗ.

Метод Прони обладает высокой точностью при оптимальном задании его параметров. Это свойство особенно актуально в тех областях науки и техники, где требуется точное восстановление функций по экспериментальным данным.

В общем случае метод Прони представляет исходные данные в виде дискретной последовательности

$$x[n] \ \ (n = 1,..., \ N) \tag{4}$$

в виде (3), что является крайне важным свойством метода, состоящего в аппроксимации дискретного ряда непрерывной аналитической функцией. Для последующего анализа и прогнозирования (3) легко преобразуется в дискретную форму подстановкой шага дискретизации Δt:

$$\hat{x}[n] = \sum_{i=1}^{p} A_i e^{\alpha_i (n-1)\Delta t} \sin\left(f_i (n-1)\Delta t + \theta_i\right). \tag{5}$$

Параметры модели Прони определяются последовательным применением математических приемов, включающих синтез АР-модели, определение корней полинома и решение системы линейных алгебраических уравнений с помощью метода наименьших квадратов (МНК).

Метод состоит из трех этапов:

1) определение параметров АР-модели путем решения системы линейных уравнений;

2) формирование из найденных коэффициентов характеристического полинома, определение его корней и расчет по ним частот и коэффициентов затухания соответствующих гармонических компонент;

3) оценка амплитуд и начальных фаз гармонических компонент на основе комплексного представления модели Прони.

В отличие от первоначальной формулировки метода Прони, где количество исходных точек данных лишь незначительно превышало порядок модели, современные реализации метода способны эффективно обрабатывать объемы данных, существенно превосходящие задаваемый порядок системы.

4. ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ МЕТОДОМ ПРОНИ

Метод позволяет разложить сложные ВР на простые гармонические составляющие, что значительно упрощает процесс анализа физических процессов, а также их поведение в будущем. Основные этапы анализа.

1. ***Выбор исходных данных***. Определение источника данных и содержательные характеристики ВР (например, финансовые показатели, метеорологические данные и т. д.).

2. ***Выбор параметров метода Прони***. Определение ключевых параметров, которые будут использоваться: длина N исходного сигнала (выбор размера окна для анализа), порядок модели p Прони (определение числа гармонических составляющих, которые будут извлечены из ВР, так как это важно для

обеспечения баланса между точностью и сложностью модели), глубина прогноза M (длина прогнозируемого отрезка).

3. ***Применение метода***. После подготовки данных и выбора параметров можно переходить к непосредственному применению метода. Прогнозируемое значение в момент времени $N+M$ вычисляется как сумма значений гармонических компонент:

$$\hat{x}[N+M] = \sum_{i=1}^{p} A_i e^{\alpha_i (N+M)\Delta t} \sin\left(f_i (N+M)\Delta t + \theta_i\right). \quad (6)$$

Прогнозирование ВР методом Прони включает в себя решение задачи определения длины временной выборки исходной функции, подлежащей анализу. При отсутствии априорных данных об оптимальной длине выборки ее значение определяется опытным путем. Наилучший порядок модели также находится на основе вычислительных экспериментов.

5. ПОСЛЕДОВАТЕЛЬНОЕ ПРИМЕНЕНИЕ МЕТОДА ПРОНИ

В стремлении повысить точность прогнозирования авторами был предложен подход, основанный на концепции последовательного применения метода Прони (ППМП). После выполнения стандартных этапов метода Прони с рекомендуемыми параметрами [17] в конце исходного ряда выделяются короткие сегменты и соответствующие участки модели. Затем определяется разница между этими фрагментами, и разностный ряд аппроксимируется новой моделью с помощью метода Прони. Полученная последовательность интегрируется в ранее построенную модель путем суммирования, что повышает качество приближения.

Эта процедура показана на рис. 1, где выделен короткий интервал $N1 - N2$ и соответствующая разность между исходным ВР и его аппроксимацией.

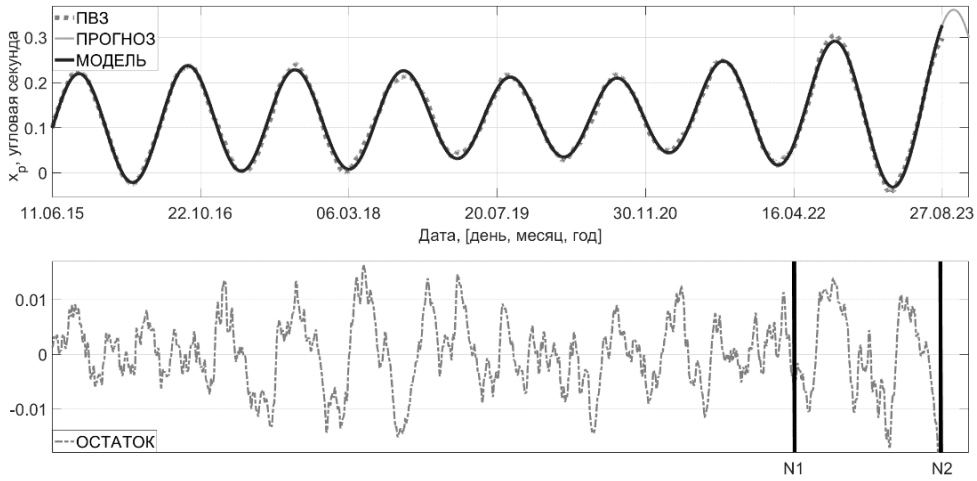

Рис. 1. Первичный этап обработки и выделение короткого фрагмента $N1 - N2$

Fig. 1. Primary stage of processing and selection of a short fragment $N1 - N2$

На рис. 2 показан остаток на интервале $N1 - N2$ и результат его аппроксимации, построенной повторным применением метода Прони для экстраполяции короткого сегмента.

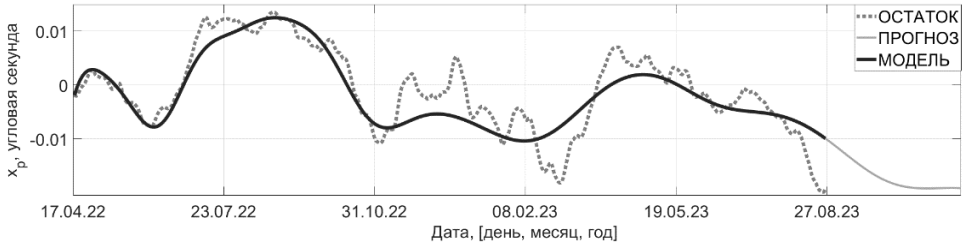

Рис. 2. Полученный остаток $N1 - N2$ и его аппроксимация

Fig. 2. Resulting remainder $N1 - N2$ and its approximation

Рисунок 3 отражает результат применения ППМП. На графике видно, что исходный прогноз, полученный методом Прони (см. рис. 1), уже в начале интервала прогнозирования демонстрирует заметное расхождение с фактическими значениями. С помощью ППМП удалось существенно увеличить точность прогноза по сравнению с результатами после выполнения стандартных этапов метода Прони.

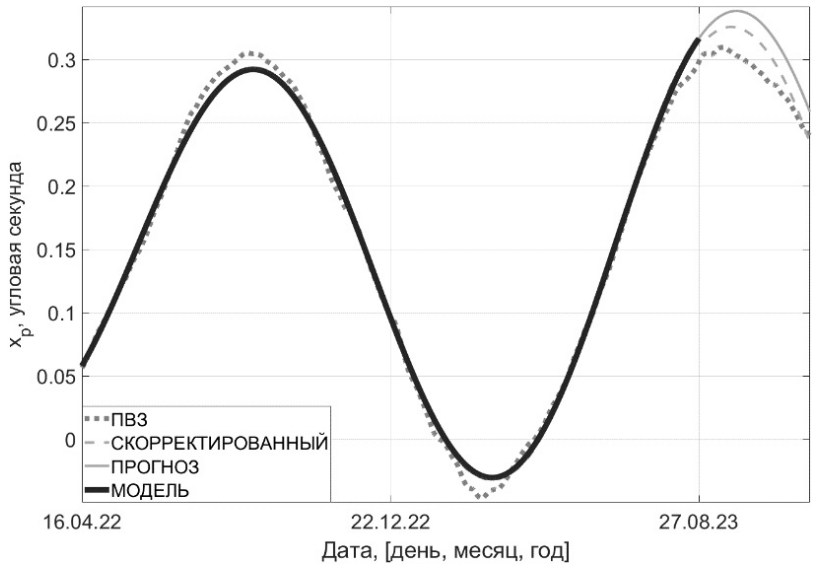

Рис. 3. Результат повторной обработки

Fig. 3. Result of repeated processing

Для подтверждения работоспособности метода ППМП необходимо убедиться в сохранении его свойств для разных временных отрезков на протяжении всего временного ряда, поведение которого в будущем представляет интерес.

6. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

В исследовании использованы данные вариации отклонений оси вращения Земли x_p и y_p за период с 03.01.2025 по 27.06.2025. Прогноз проводился через каждые 7 дней. Параметры алгоритма были заданы в соответствии с рекомендациями на основе многочисленных экспериментов и исследований [17]. Порядок модели Прони (6) задается отдельно для каждого ВР: $p = 13$ для x_p и $p = 8$ для y_p. Размер исходной выборки $N = 3000$ суток. Длина разностного ряда $N_1 - N_2 = 500$ суток. Глубина прогнозирования $M = 7$ и 90 суток (краткосрочные и долгосрочные прогнозы).

В исследованиях по прогнозированию ВР применяются различные метрики точности [18], включая среднюю абсолютную ошибку (САО), среднюю абсолютную процентную ошибку (САПО), среднеквадратическую ошибку (СКО) и другие. Поскольку априори нет оснований для выбора какой-либо конкретной метрики, выбрана одна, которая широко использовалась предыдущими исследователями в этой области.

Среднеквадратическая погрешность (СКП) определяется как квадратный корень из среднего значения квадратов разностей между прогнозируемыми и фактическими значениями:

$$\text{СКП} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \hat{x}_i\right)^2}\ , \tag{7}$$

где x_i – фактическое значение, $\hat{x}_i$ – прогнозируемое значение, n – количество прогнозируемых отсчетов.

7. СРАВНЕНИЕ РЕЗУЛЬТАТОВ

В табл. 1 и 2 представлены сравнительные характеристики точности прогнозирования ПВЗ классическим методом Прони, модифицированным методом Прони и методом авторегрессии – модели со скользящим средним (IERS). В табл. 1 представлены показатели СКП при среднесрочном прогнозировании. Оценка проводилась по всей длине прогноза: по отсчетам с 1-го по 90-й на основе формулы (7).

Таблица 1

Table 1

Глубина прогноза *M* = 90

Forecasting depth *M* = 90

№ п/п	Дата начала прогноза	Метод прогнозирования	x_p	y_p
			СКП, угловая миллисекунда	
1	03.01.25	Метод Прони	20,5	6,9
2		ППМП	20,5	4,9
3		IERS	0,3	6,6
4	10.01.25	Метод Прони	13,5	8,1
5		ППМП	13,1	6,2
6		IERS	19,0	6,6

Окончание табл. 1

End of the Tab. 1

№ п/п	Дата начала прогноза	Метод прогнозирования	x_p	y_p
			СКП, угловая миллисекунда	
7	17.01.25	Метод Прони	8,4	10,1
8		ППМП	7,7	8,2
9		IERS	16,2	7,2
10	24.01.25	Метод Прони	8,2	12,4
11		ППМП	7,4	10,2
12		IERS	20,0	11,6
13	31.01.25	Метод Прони	8,0	14,7
14		ППМП	7,0	12,3
15		IERS	16,6	10,9
…				
40	04.04.25	Метод Прони	62,4	6,2
41		ППМП	8,2	6,0
42		IERS	4,5	7,2
43	11.04.25	Метод Прони	70,6	4,2
44		ППМП	9,0	4,0
45		IERS	6,0	8,3
46	18.04.25	Метод Прони	83,3	3,2
47		ППМП	11,1	2,6
48		IERS	4,8	7,7
…				
64	30.05.25	Метод Прони	11,5	8,0
65		ППМП	9,9	4,9
66		IERS	11,6	3,9
67	06.06.25	Метод Прони	21,4	7,2
68	06.06.25	ППМП	15,4	4,7
69		IERS	12,6	2,4
70	13.06.25	Метод Прони	21,1	7,1
71		ППМП	17,9	5,0
72		IERS	13,4	2,9
73	20.06.25	Метод Прони	26,2	6,6
74		ППМП	20,9	4,4
75		IERS	13,7	4,0
76	27.06.25	Метод Прони	28,5	5,7
77		ППМП	23,0	2,8
78		IERS	14,7	5,3
	Среднее значение	Метод Прони	10,8	11,1
		ППМП	9,6	8,8
		IERS	9,3	8,6

В табл. 2 приводится сравнение СКП краткосрочных прогнозов и полученных ранее указанными способами. Оценка выполнялась по всей длине краткосрочного прогноза (по отсчетам с 1-го по 7-й прогноз).

Таблица 2

Table 2

Глубина прогноза $M = 7$

Forecasting depth $M = 7$

№ п/п	Дата начала прогноза	Метод прогнозирования	x_p	y_p
			СКП, угловая миллисекунда	
1	03.01.25	Метод Прони	7,3	2,2
2		ППМП	6,4	0,6
3		IERS	1,3	0,2
4		ЗСФ ВНИИФТРИ	1,2	1,7
5	10.01.25	Метод Прони	6,8	1,4
6		ППМП	4,9	1,0
7		IERS	2,9	1,5
8		ЗСФ ВНИИФТРИ	4,3	2,5
9	17.01.25	Метод Прони	3,5	1,7
10		ППМП	2,9	1,3
11		IERS	3,8	0,8
12		ЗСФ ВНИИФТРИ	7,2	2,5
13	24.01.25	Метод Прони	3,6	1,0
14		ППМП	2,5	1,0
15		IERS	3,9	0,7
16		ЗСФ ВНИИФТРИ	8,1	0,3
17	31.01.25	Метод Прони	2,9	1,4
18		ППМП	2,2	1,4
19		IERS	3,6	2,1
20	31.01.25	ЗСФ ВНИИФТРИ	0,7	0,9
…				
65	25.04.25	Метод Прони	2,8	3,1
66		ППМП	1,0	2,0
67		IERS	0,6	0,9
68		ЗСФ ВНИИФТРИ	1,4	1,2
69	02.05.25	Метод Прони	4,7	1,4
70		ППМП	3,2	0,4
71		IERS	0,7	0,8
72		ЗСФ ВНИИФТРИ	0,7	0,6
73	09.05.25	Метод Прони	4,7	3,0
74		ППМП	4,1	1,7
75		IERS	0,5	0,4
76		ЗСФ ВНИИФТРИ	1,6	1,0
…				
89	06.06.25	Метод Прони	8,3	5,3
90		ППМП	6,1	2,4
91		IERS	0,8	0,5
92		ЗСФ ВНИИФТРИ	1,2	1,0
93	13.06.25	Метод Прони	4,7	5,7
94		ППМП	2,2	3,3
95		IERS	2,0	0,8
96		ЗСФ ВНИИФТРИ	1,9	1,5

Окончание табл. 2

End of the Tab. 2

№ п/п	Дата начала прогноза	Метод прогнозирования	x_p	y_p
			СКП, угловая миллисекунда	
97	20.06.25	Метод Прони	2,1	5,7
98		ППМП	1,7	4,2
99		IERS	1,1	0,8
100		ЗСФ ВНИИФТРИ	1,2	1,4
101	27.06.25	Метод Прони	5,1	3,9
102		ППМП	1,4	3,5
103		IERS	1,6	0,8
104		ЗСФ ВНИИФТРИ	0,5	1,6
	Среднее значение	Метод Прони	4,6	4,7
		ППМП	3,2	2,3
		IERS	1,5	0,8
		ЗСФ ВНИИФТРИ	2,0	1,3

Как следует из данных табл. 2, ППМП привело к снижению СКП относительно исходной реализации метода Прони. На рис. 4–7 визуализирована динамика прогнозирования отклонений оси вращения x_p и y_p, рассчитанных как разность между последними точками прогноза (90 и 7) и фактическими значениями, в соответствии с выражением

$$\delta x = \left| \hat{x} - x \right| . \tag{8}$$

Оценки выполнены для интервалов прогнозирования 7 и 90 суток за период с 03.01.25 по 27.06.25.

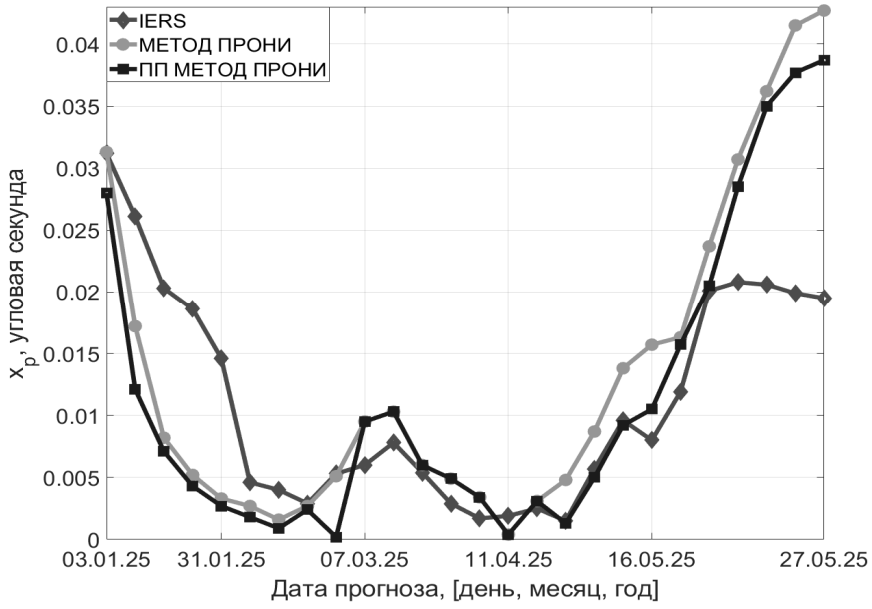

Рис. 4. Сравнение отклонений от истинных значений x_p при глубине $M = 90$

Fig. 4. Comparison of deviations from true values x_p depth $M = 90$

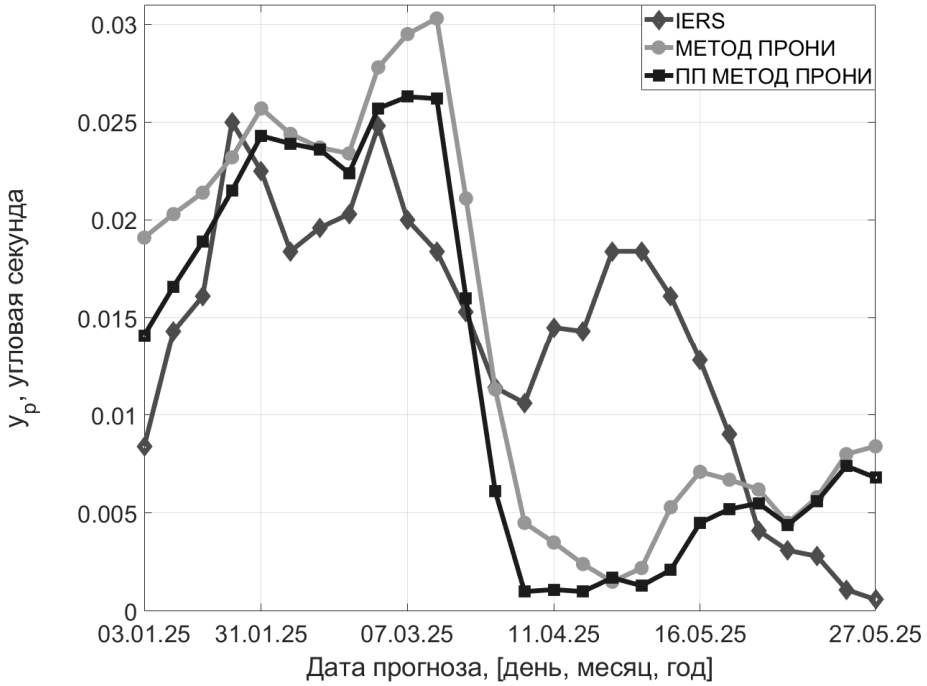

Рис. 5. Сравнение отклонений от истинных значений y_p при глубине $M = 90$

Fig. 5. Comparison of deviations from true values y_p depth $M = 90$

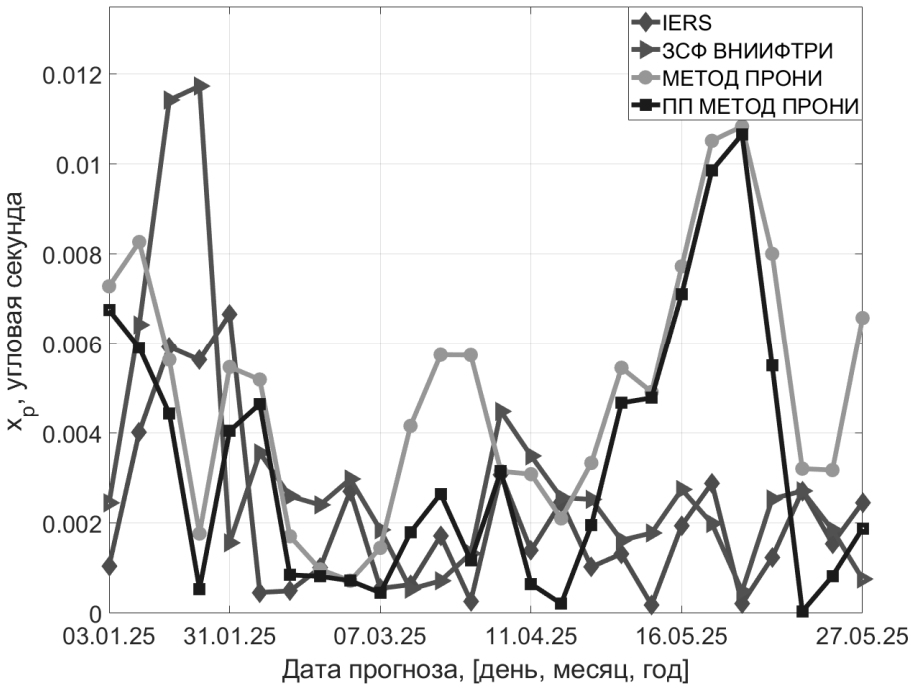

Рис. 6. Сравнение отклонений от истинных значений x_p при глубине $M = 7$

Fig. 6. Comparison of deviations from true values x_p depth $M = 7$

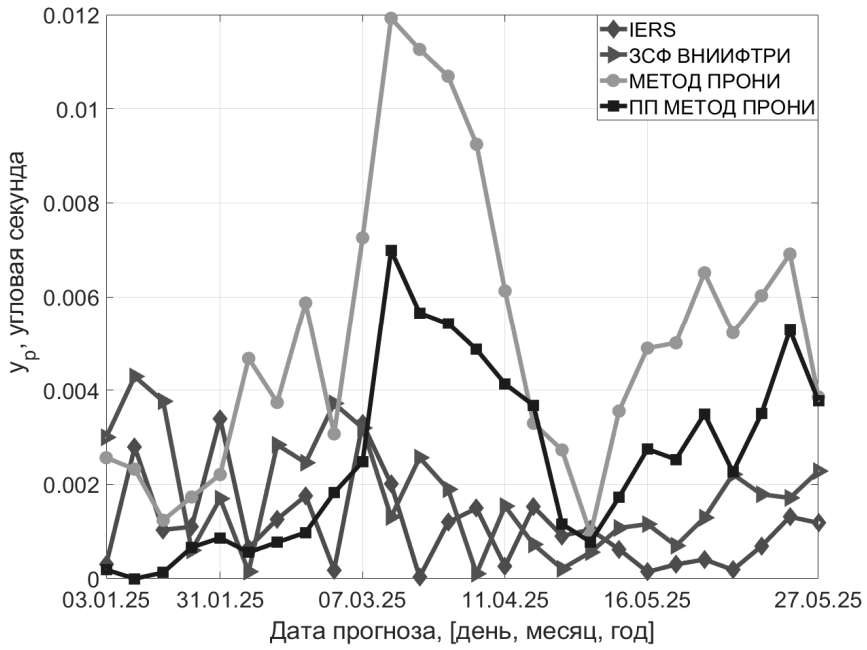

Рис. 7. Сравнение отклонений от истинных значений y_p при глубине $M = 7$

Fig. 7. Comparison of deviations from true values y_p depth $M = 7$

Результаты исследования подтверждают эффективность предложенной модификации метода: разработанный на ее основе алгоритм демонстрирует существенное повышение точности прогнозирования по сравнению с исходной реализацией метода Прони. Однако сравнение с прогнозами профильных организаций показывает нестабильность результатов применения метода ППМП на протяжении всего временного ряда.

ЗАКЛЮЧЕНИЕ

В результате проведенного исследования разработан и апробирован алгоритм последовательного применения метода Прони для прогнозирования временных рядов.

Экспериментальная проверка эффективности реализации представленного метода выполнена на временных рядах вариаций отклонений оси вращения x_p и y_p в сравнении с исходным методом Прони и прогнозами профильных организаций (IERS, ЗСФ ВНИИФТРИ). Установлено, что предложенная модификация позволяет достичь СКП на 50 % ниже по сравнению с классической реализацией метода Прони для x_p и на 30 % ниже для y_p при долгосрочных прогнозах (см. табл. 1, рис. 4, 5), а также на 60 % ниже по сравнению с классической реализацией метода Прони для x_p и на 30 % ниже для y_p при краткосрочных прогнозах (см. табл. 2, рис. 6). Однако проведенное сравнение выявило более высокую точность прогнозов, предоставляемых IERS, особенно при глубине прогноза $M = 7$, из чего можно сделать вывод, что долгосрочные прогнозы ППМП более конкурентоспособны, что важно для автономной

работы спутниковых группировок. Следует отметить, что в исходном временном ряду были выявлены отрезки, для которых наблюдается значительное снижение погрешности прогнозов ППМП отклонений оси вращения x_p (рис. 4) в интервале с 10.01.25 до 14.02.25, а для отклонений оси вращения y_p (рис. 5) в интервале с 04.04.25 до 16.05.25 в сравнении с другими методами.

Проведенное исследование раскрыло определенные ограничения применения статически заданных параметров алгоритма ППМП: на участках с выраженной сменой тенденций развития временного ряда, а именно при переходе от восходящего тренда к нисходящему и наоборот, наблюдается систематическое снижение точности прогнозов.

Предложенный подход может быть применен для анализа временных рядов различной физической природы, содержащих колебательные компоненты. Однако для каждого конкретного прогноза требуется отдельное исследование сочетания параметров алгоритма, включая порядок модели. Использование фиксированных параметров алгоритма не позволяет достичь оптимальных результатов прогнозирования на всех временных горизонтах. Наблюдаемые отклонения в точности прогнозов свидетельствуют о том, что текущие параметры модели не являются универсальными для всего спектра решаемых задач. Полученные результаты указывают на необходимость разработки адаптивных механизмов настройки параметров алгоритма. Дальнейшее развитие метода будет сосредоточено на разработке и апробации версий алгоритма, способных автоматически настраивать его параметры в соответствии с изменяющимся содержанием исходных данных.

СПИСОК ЛИТЕРАТУРЫ

1. *Бокс Дж.*, *Дженкинс Г.* Анализ временных рядов: прогноз и управление. Вып. 1. – М.: Мир, 1974. – 406 с.

2. *Буре В.М., Плахотник С.В.* Адаптивные методы прогнозирования временных рядов в среде MATLAB // Проектирование научных и инженерных приложений в среде MATLAB: труды Всероссийской научной конференции. – СПб., 2007. – С. 1363–1370.

3. SARIMA modelling approach for forecasting of traffic accidents / N. Deretić, D. Stanimirović, M.A. Awadh, N. Vujanović, A. Djukić // Sustainability. – 2022. – Vol. 14, N 8. – P. 4403.

4. *Wang C., Zhang H., Ma P.* Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network // Applied Energy. – 2020. – Vol. 259. – P. 114139.

5. *Li X., Tang P.* Stock index prediction based on wavelet transform and FCD-MLGRU // Journal of Forecasting. – 2020. – Vol. 39 (8). – P. 1229–1237.

6. *Gong J., Liu S.* A harmonic and interharmonic detection method for power systems based on enhanced SVD and the Prony algorithm // Applied Sciences. – 2023. – Vol. 13 (13). – P. 7558.

7. High precision prediction of universal time based on 100-year data / V.M. Tissen, A.S. Tolstikov, A.Y. Balakhnenko, Z.M. Malkin // Measurement Techniques. – 2009. – Vol. 52 (12). – P. 1249–1255.

8. Исследование основных составляющих в движении полюса Земли / М.В. Воротков, В.Л. Горшков, Н.О. Миллер, Е.Я. Прудникова // Известия Главной астрономической обсерватории в Пулкове. – 2002. – № 216. – С. 406–414.

9. International Earth Rotation and Reference Systems Service. – URL: http://www.iers.org/iers/products (accessed: 29.10.2025).

10. *Тиссен В.М.* Долгосрочное прогнозирование изменений климата по вариациям скорости вращения Земли // Метрология времени и пространства: материалы XI Международного симпозиума, Менделеево, 27–29 сентября 2023 г. – Менделеево, 2023. – С. 170–173.

11. *Montgomery D.C., Jennings C.L., Kulahci M.* Introduction to time series analysis and forecasting. – John Wiley & Sons, 2015.

12. IERS Annual Report 2018 / ed. by W.R. Dick, D. Thaller; International earth rotation and reference systems service, Central Bureau. – Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie, 2020. – 207 p. – ISBN 978-3-86482-136-3.

13. *Тиссен В.М., Толстиков А.С., Симонова Г.В.* Прогнозирование параметров вращения Земли с помощью адаптивных гармонических моделей // Вестник СГУГиТ. – 2020. – Т. 25, № 4. – С. 238–245.

14. *Petit G, Luzum B.* IERS conventions 2010. – Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie, 2010. – 179 p. – (IERS Technical Note; 36).

15. *Марпл С.Л. (мл.).* Цифровой спектральный анализ и его приложения: пер. с англ. – М.: Мир, 1990. – 584 с.

16. Advanced computational methods for modeling, prediction and optimization – a review / J. Krzywanski, M. Sosnowski, K. Grabowska, A. Zylka, L. Lasek, A. Kijo-Kleczkowska // Materials. – 2024. – Vol. 17 (14). – P. 3521.

17. *Gavrilov A.B., Ipatiev V.A., Kuzminykh I.A.* Application of Prony's method for predicting the earth rotation parameters // 2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM). – IEEE, 2024. – P. 1890–1893.

18. *Малкин З.М., Тиссен В.М.* Сравнение различных оценок точности прогноза параметров вращения земли // Астрономический журнал. – 2022. – Т. 99, № 1. – С. 70–74.

Кузьминых Илья Александрович, аспирант кафедры автоматизированных систем управления Новосибирского государственного технического университета. Основное направление научных исследований – анализ и прогнозирование временных рядов. E-mail: i.kuzminyx.2019@stud.nstu.ru

Гаврилов Алексей Борисович, кандидат технических наук, доцент кафедры автоматики Новосибирского государственного технического университета. Основное направление научных исследований – метрология, навигация, анализ и прогнозирование временных рядов. E-mail: a.gavrilov@corp.nstu.ru

Толстиков Александр Сергеевич, доктор технических наук, ведущий научный сотрудник отдела «Государственная служба времени, частоты и определения параметров вращения Земли» Западно-Сибирского филиала ФГУП ВНИИФТРИ. Профессор кафедры специальных устройств, инноватики и метрологии Сибирского государственного университета геосистем и технологий. Профессор кафедры систем сбора и обработки данных Новосибирского государственного технического университета. Имеет более 100 печатных работ. Основное направление научных исследований – метрология, навигация. E-mail: tolstikov@sniim.ru

Виктор Мартынович Тиссен, кандидат технических наук, старший научный сотрудник Западно-Сибирского филиала ФГУП ВНИИФТРИ. Имеет более 70 печатных работ. Основное направление научных исследований – метрология, навигация. E-mail: tissen.ksn@yandex.ru

Kuzminykh Ilya A., postgraduate student at the Department of Automated Control Systems, Novosibirsk State Technical University. His research interests are currently focused on analysis and forecasting of time series. E-mail: i.kuzminyx.2019@stud.nstu.ru

Gavrilov Alexey B., PhD (Eng.), assistant professor of the Department of Automation at Novosibirsk State Technical University. His research interests are currently focused on metrology, navigation, analysis and forecasting of time series. E-mail: a.gavrilov@corp.nstu.ru

Tolstikov Alexander S., D.Sc. (Eng.), leading researcher at the Department "State Service of Time, Frequency and Determination of the Parameters of the Earth's Rotation", West Siberian Branch of the Federal State Unitary Enterprise VNIIFTRI. Professor at the Department of Special Devices, Innovation and Metrology, Siberian State University of Geosystems and Technologies. Professor at the Department of Data Collection and Processing Systems, Novosibirsk State Technical University. He has more than 100 publications. His research interests are currently focused on metrology, satellite navigation. E-mail: tolstikov@sniim.ru

Tissen Viktor M., PhD (Eng.), Senior Researcher, West Siberian Branch of the Federal State Unitary Enterprise VNIIFTRI. He has more than 70 publications. His research interests are currently focused on metrology, satellite navigation. E-mail: tissen.ksn@yandex.ru

DOI: 10.17212/2782-2001-2026-3-57-74

Sequential application of the Prony's method for improved time series forecasting accuracy[*]

I.A. KUZMINYKH[1,a], *A.B. GAVRILOV*[1,b], *A.S. TOLSTIKOV*[2,c], *V.M. TISSEN*[2,d]

[1] *Novosibirsk State Technical University, 20 K. Marks Prospekt, Novosibirsk, 630073, Russian Federation*

[2] *West Siberian Branch of the Russian Metrological Institute of Technical Physics and Radio Engineering, 4 Dimitrova Prospekt, Novosibirsk, 630004, Russian Federation*

[a] *i.kuzminyx.2019@stud.nstu.ru* [b] *a.gavrilov@corp.nstu.ru* [c] *tolstikov@sniim.ru*
[d] *tissen.ksn@yandex.ru*

Abstract

This work is devoted to modifying the Prony's method to improve the accuracy of forecasting oscillatory time series. The proposed algorithm is based on a multi-stage approach: the model sequentially approximates the residual components of the series, which are then integrated into the final function. The modification was applied in practice to the task of forecasting time series dominated by complex oscillatory processes and harmonic components: Earth rotation parameters, namely the pole coordinates x_p and y_p, according to the International Earth Rotation and Reference Systems Service (IERS) data archive for 2025. A comparative analysis with the original Prony's method, as well as with the forecasts of the IERS and the West Siberian branch FSUE VNIIFTRI at intervals of 7 and 90 days, showed that the effectiveness of the modification significantly depends on the specified conditions. The proposed approach does not always outperform the existing solutions of the IERS and VNIIFTRI, but for certain forecast horizons, sequential refinement of the model provides a local increase in accuracy and a reduction in forecast error. The identified heterogeneity of the results indicates that the effectiveness of the sequential application of the Prony's method significantly depends on the specified model parameters. The results obtained open up opportunities for applying the method in autonomous satellite navigation systems. The analysis allowed us to identify the limits of the current model's applicability and formulate recommendations for its development. In particular, it justifies the need to move from the use of statically specified parameters to the development of mechanisms for their adaptive adjustment, which will stabilize the results of the method when the nature of the initial data changes.

Keywords: forecasting, time series, digital processing of measurement data, autoregressive models, Prony's method, optimization of algorithm parameters, harmonic analysis, Earth rotation parameters, tilt of the Earth's axis, forecast accuracy

REFERENCES

1. Box G., Jenkins G. *Analiz vremennykh ryadov: prognoz i upravlenie*. Vyp. 1 [Time series analysis: forecasting and control]. Moscow, Mir Publ., 1974. 406 p. (In Russian).

2. Bure V.M., Plakhotnik S.V. [Adaptive methods of forecasting time series in the MATLAB environment]. *Proektirovanie nauchnykh i inzhenernykh prilozhenii v srede MATLAB* [Designing scientific and engineering applications in the MATLAB environment]. Proceedings of the All‑Russian scientific conference, St. Petersburg, 2007, pp. 1363–1370. (In Russian).

3. Deretić N., Stanimirović D., Awadh M.A., Vujanović N., Djukić A. SARIMA modelling approach for forecasting of traffic accidents. *Sustainability*, 2022, vol. 14 (8), p. 4403.

4. Wang C., Zhang H., Ma P. Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network. *Applied Energy*, 2020, vol. 259, p. 114139.

[*] *Received 25 January 2026.*

5. Li X., Tang P. Stock index prediction based on wavelet transform and FCD-MLGRU. *Journal of Forecasting*, 2020, vol. 39 (8), p. 1229–1237.

6. Gong J., Liu S. A harmonic and interharmonic detection method for power systems based on enhanced SVD and the Prony algorithm. *Applied Sciences*, 2023, vol. 13 (13), p. 7558.

7. Tissen V.M., Tolstikov A.S., Balakhnenko A.Y., Malkin Z.M. High precision prediction of universal time based on 100-year data. *Measurement Techniques*, 2009, vol. 52 (12), pp. 1249–1255.

8. Vorotnikov M.V., Gorshkov V.L., Miller N.O., Prudnikova E.Ya. Issledovanie osnovnykh sostavlyayushchikh v dvizhenii polyusa Zemli [Research of the Basic Components in the Earth's Pole Motion]. *Izvestiya Glavnoi astronomicheskoi observatorii v Pulkove*, 2002, no. 216, pp. 406–414. (In Russian).

9. *International Earth Rotation and Reference Systems Service*. Available at: http://www.iers.org/iers/products (accessed 29.10.2025).

10. Tissen V.M. [Long-term forecasting of climate change based on variations in the speed of the Earth's rotation]. *Metrologiya vremeni i prostranstva* [Metrology of time and space]. Proceedings of the XI International symposium, Mendeleevo, 2023, pp. 170–173. (In Russian).

11. Montgomery D.C., Jennings C.L., Kulahci M. *Introduction to time series analysis and forecasting*. John Wiley & Sons, 2015.

12. Dick W.R., Thaller D., eds. *IERS Annual Report 2018*. International earth rotation and reference systems service, Central Bureau. Frankfurt am Main, Verlag des Bundesamts für Kartographie und Geodäsie, 2020. 207 p. ISBN 978-3-86482-136-3.

13. Tissen V.M., Tolstikov A.S., Simonova G.V. Prognozirovanie parametrov vrashcheniya Zemli s pomoshch'yu adaptivnykh garmonicheskikh modelei [Prediction of Earth rotation parameters with using adaptive harmonic patterns]. *Vestnik SGUGiT = Vestnik SSUGiT*, 2020, vol. 25, no. 4, pp. 238–245.

14. Petit G, Luzum B. *IERS conventions 2010*. *IERS Technical Note*, 36. Frankfurt am Main, Verlag des Bundesamts für Kartographie und Geodäsie, 2010. 179 p.

15. Marple Jr S.L. *Tsifrovoi spektral'nyi analiz i ego prilozheniya* [Digital spectral analysis with applications]. Moscow, Mir Publ., 1990. 584 p. (In Russian).

16. Krzywanski J., Sosnowski M., Grabowska K., Zylka A., Lasek L., Kijo-Kleczkowska A. Advanced computational methods for modeling, prediction and optimization – a review. *Materials*, 2024, vol. 17 (14), p. 3521.

17. Gavrilov A.B., Ipatiev V.A., Kuzminykh I.A. Application of Prony's method for predicting the earth rotation parameters. *2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM)*. IEEE, 2024, pp. 1890–1893.

18. Malkin Z.M., Tissen V.M. Sravnenie razlichnykh otsenok tochnosti prognoza parametrov vrashcheniya zemli [Comparison of different estimates of the accuracy of smart earth parameter prediction]. *Astronomicheskii zhurnal = Astronomy Reports*, 2022, vol. 99, no. 1, pp. 70–74.

Для цитирования:

Последовательное применение метода Прони для повышения точности прогнозирования временных рядов / И.А. Кузьминых, А.Б. Гаврилов, А.С. Толстиков, В.М. Тиссен // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 57–74. – DOI: 10.17212/2782-2001-2026-3-57-74.

For citation:

Kuzminykh I.A., Gavrilov A.B., Tolstikov A.S., Tissen V.M. Posledovatel'noe primenenie metoda Proni dlya povysheniya tochnosti prognozirovaniya vremennykh ryadov [Sequential application of the Prony's method for improved time series forecasting accuracy]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 57–74. DOI: 10.17212/2782-2001-2026-3-57-74.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 75–96

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 75–96

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.85

DOI: 10.17212/2782-2001-2026-3-75-96

Обработка текстовых данных в медицинских диалоговых системах на примере виртуального ассистента эндокринолога*

М.С. МАМИКЬЯН[a], Я.А. БЕКЕНЕВА[b]

РФ, 197022, г. Санкт-Петербург, ул. Профессора Попова, 5, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

[a] *mmamikyan2001@gmail.com* [b] *yabekeneva@etu.ru*

Доля больных с патологиями эндокринной системы постоянно увеличивается во всех экономически развитых странах мира. В России за последние 6 лет распространенность заболеваний эндокринной системы выросла на 26 %, а показатели первичной заболеваемости – на 36 %. Одной из главных проблем, усугубляющих эпидемиологическую ситуацию, является дефицит местных терапевтов и неравномерное распределение медицинских специалистов между сельскими и городскими регионами. Жители малонаселенных пунктов вынуждены ездить в соседние города для получения необходимой медицинской помощи, что влечет дополнительные временные и финансовые затраты.

Также стоит отметить, что определение необходимых анализов для сдачи может занять немало времени, а спешка при приеме может привести к сдаче избыточных анализов. Для оптимизации этого процесса важно заранее определить, какие именно анализы нужны пациенту, что требует постановки предварительного диагноза. В этом может помочь система распознавания заболеваний на основе жалоб пациента, которая сможет автоматизировать постановку предварительного диагноза и избежать ненужных исследований. Основу представленного программного решения составляет комбинация использования распознавания именованных сущностей для извлечения симптомов из жалоб пациентов в свободной форме и методов машинного обучения для определения диагноза. Для обучения использован набор данных RuDRec в совокупности с синтетическими данными. Описан алгоритм диалога с пользователем для уточнения симптомов и формирования перечня необходимых анализов.

Такая система поможет сократить случаи самолечения, когда пациенты, не имея точных рекомендаций, следуют советам знакомых. Таким образом, формирование направлений на анализы на основе качественной первичной консультации пациентов, особенно в малонаселенных пунктах, является актуальной задачей. Практическая значимость работы заключается в оптимизации нагрузки на медицинские учреждения, в снижении рисков самолечения и повышении диагностики в удаленных регионах.

* *Статья получена 06 июня 2025 г.*

Ключевые слова: виртуальный ассистент эндокринолога, методы обработки медицинских данных, большие языковые модели, чат-бот, диалоговый интерфейс, диагностика, эндокринологическое заболевание, телемедицина, именованные сущности, распознавание текстов, метрики точности, интеллектуальная система машинного обучения

ВВЕДЕНИЕ

Постановка эндокринного диагноза представляет собой комплексную проблему. В монографиях по эндокринным заболеваниям, как правило, методы исследования эндокринной системы описаны схематично [1]. Диагностика пациента включает в себя следующее [2].

1. Анамнез. Жалобы у больных с эндокринной патологией чрезвычайно многообразны, поэтому опрос пациента должен включать в себя несколько этапов, которые позволят врачу собрать информацию о симптомах и их продолжительности. Врач тщательно выслушивает все жалобы пациента и задает уточняющие вопросы, чтобы сузить круг возможных заболеваний. Затем уточняется история жизни больного, особое внимание уделяется опросу о родственниках, так как ко многим эндокринопатиям имеется наследственная предрасположенность.

2. Осмотр. При общем осмотре пациента врач обращает внимание на несколько ключевых аспектов: внешний вид, состояние кожи и волос, особенности лица.

3. Пальпация – физический метод медицинской диагностики, проводимый путем ощупывания тела пациента. При пальпации щитовидной железы врач ориентируется на ее размер, консистенцию и наличие узлов.

После сбора анамнеза, исходя из жалоб и состояния пациента, врач может поставить предполагаемый диагноз и назначить определенные анализы для постановки точного диагноза. Перечень анализов подбирается индивидуально врачом-эндокринологом в зависимости от симптомов, предварительного диагноза и состояния пациента. Автоматизация процесса сбора анамнеза и подбора необходимых анализов может позволить сократить количество приемов пациента у медицинского специалиста. Такое решение может стать особенно актуальным для отдаленных населенных пунктов, где могут отсутствовать профильные специалисты. В некоторых регионах России время ожидания приема профильного специалиста может достигать нескольких месяцев.

Технологии телемедицины сегодня являются важным инструментом для обеспечения равного доступа к медицинской помощи по всему миру. Их активное внедрение уже стало стандартом во многих странах, включая Россию. Электронная запись к врачу, система электронной очереди и другие цифровые сервисы помогают повысить качество обслуживания пациентов, оптимизировать работу медицинского персонала и снизить затраты на оказание медицинских услуг. Современные телемедицинские технологии охватывают широкий спектр задач – от диагностики и лечения заболеваний до дистанционного биомониторинга [3]. Перспективы развития телемедицины связаны с дальнейшим внедрением цифровых технологий. Уже сейчас телемедицинские платформы активно используются для удаленных консультаций с узкими специалистами, для мониторинга хронических заболеваний и контроля жизненно важных показателей, таких как уровень сахара в крови, артериальное давление, сердечный ритм. Несмотря на активное развитие, имеются факторы, ограничивающие

широкомасштабное внедрение телемедицины. Тем не менее мировая практика показывает, что в ближайшие годы телемедицина станет неотъемлемой частью системы здравоохранения, обеспечивая доступность и качество медицинских услуг независимо от географического положения пациента [4].

В первом разделе статьи рассматриваются и анализируются существующие постановки задачи обработки текстов и создания диалоговых систем. Во втором разделе представлен подход к обработке жалоб, вводимых пациентом в текстовой форме. В третьем разделе представлены результаты экспериментальных исследований предлагаемой диалоговой системы.

1. ПОСТАНОВКА ЗАДАЧИ

Задача автоматизации процесса первичной диагностики пациента может быть сведена к задаче автоматизации сбора анамнеза. Как правило, на приеме у врача-эндокринолога пациент перечисляет жалобы в свободной форме, а врач задает уточняющие вопросы. Автоматизация этого процесса может быть реализована в виде диалоговой системы. Чат-боты, например, становятся неотъемлемой частью цифровой медицины, помогают автоматизировать взаимодействие с пациентами и снижать нагрузку на медицинские учреждения. Они представляют собой программных ассистентов, способных анализировать жалобы, задавать уточняющие вопросы и предоставлять рекомендации. Первые подобные системы появились еще в 1960-х годах, однако только в последние годы они активно внедряются в сферу здравоохранения.

1.1. ЧАТ-БОТЫ КАК СРЕДСТВО ОРГАНИЗАЦИИ ДИАЛОГА С ПАЦИЕНТОМ

Современные медицинские диалоговые системы, в частности чат-боты, выполняют различные функции – от поиска ближайших медицинских учреждений и записи на прием к врачу до сбора первичной информации о симптомах пациента. Такие сервисы доступны бесплатно и могут оценивать необходимость визита к врачу. Они помогают экономить время, могут напоминать пациентам о приеме лекарств, предоставлять рекомендации и минимизировать число излишних обращений в медицинские учреждения. Чат-бот может проконсультировать пациента в экстренной ситуации, предоставив рекомендации по оказанию первой помощи до приезда специалистов [5, 6].

Babylon [7] – медицинский чат-бот, предлагающий консультацию как на основе искусственного интеллекта, так и непосредственно врачом. Эта система включает в себя технологии обработки естественного языка и алгоритмы машинного обучения. Пользователь в свободной форме вводит симптомы, затем чат-бот на основе введенных жалоб задает уточняющие вопросы. После завершения цикла опроса пациента система сравнивает полученные данные с медицинскими базами знаний и выдает вероятные диагнозы.

Ada [8] – популярный чат-бот, ориентированный на самодиагностику. Принцип работы пользователя с системой аналогичен *Babylon*. Пользователь в свободной форме вводит жалобу. Далее система предлагает список подходящих жалоб с их подробным описанием, чтобы пользователь мог выбрать наиболее соответствующую проблему. После ввода пациентом всех жалоб

система задает ряд встречных вопросов с вариантами ответов. После завершения цикла опроса пациента система выводит описание предположительного заболевания. В описание входит информация о том, сколько по статистике человек сталкивались с этой проблемой, также подробно описываются риски, симптомы, требуемый уход, профилактика и прогнозы. Стоит отметить, что *Ada* имеет возможность интегрирования в медицинские системы, предоставляя обширный функционал. Например, с его помощью можно эффективно направлять пациентов к нужным врачам, тем самым снижая нагрузку на медицинское учреждение.

Buoy Health [9] – система, обученная на основе большого массива клинических данных. Проблема, вокруг которой была сформирована эта система, заключалась в том, что люди неправильно интерпретируют свои симптомы, основываясь на информации, которую они нашли в интернете. Принцип работы аналогичен *Ada*. Существенным отличием является разная последовательность ввода жалоб. В *Ada* пользователь в первую очередь заносит все жалобы, которые его беспокоят, и только затем система задает встречные вопросы. В *Buoy Health* пользователь может ввести следующий симптом только после ответа на ряд вопросов о первом симптоме. Также стоит отметить, что в рамках каждого нового диалога *Buoy Health* задает ряд вопросов, не связанных с жалобой пациента, чтобы исключить экстренные ситуации, например приступ инсульта. В процессе опроса ассистент уточняет геолокацию, чтобы в дальнейшем порекомендовать ближайшее медицинское учреждение. После завершения цикла опроса пациента система выводит названия предположительных диагнозов, сопровождая их статьями, с которыми пользователь может ознакомиться.

Florence [10] – надстройка мессенджера, которая работает как ассистент пациента. В отличие от рассмотренных ранее систем, *Florence* не предоставляет диагностику заболеваний. *Florence* помогает пациентам соблюдать режим приема лекарств, напоминая им о необходимости принимать медикаменты, а также предоставляет информацию о назначенных препаратах; мотивирует пользователей придерживаться назначенного режима, что снижает риск ухудшения здоровья из-за неправильного или нерегулярного приема лекарств. Кроме того, *Florence* может отслеживать состояние здоровья пользователя и помогает достигать поставленных целей. Например, с помощью этого чатбота можно контролировать параметры веса. *Florence* также предоставляет информацию о различных заболеваниях и может помочь в поисках врача или аптеки.

Symptomate [11] – система быстрой диагностики симптомов. Работает аналогично *Ada* и *Buoy Health*, но имеет меньший дополнительный функционал, в открытом доступе. В отличие от рассмотренных выше систем, *Symptomate* предоставляет удобный функционал выбора зоны, беспокоящей пациента. На основе выбранной части тела система выдает список возможных жалоб, относящихся к ней. После завершения диалога с пациентом *Symptomate* выводит список возможных заболеваний. Каждому заболеванию соответствует панель инструментов, отражающая вероятность наличия этого заболевания у пациента.

Sensely [12] – виртуальный медицинский ассистент, использующий ИИ для взаимодействия с пациентами. На основе диалога с пациентом система

помогает выявить, требуется ли немедленное вмешательство врача. Интегрируясь в медицинские учреждения, *Sensely* предоставляет дополнительный функционал: помощь в записи пациентов на прием к врачу, возможность видеоконсультаций с реальными врачами, трекер состояния здоровья.

Все рассмотренные системы представляют собой инновационные инструменты, использующие искусственный интеллект для упрощения получения первичной медицинской диагностики, мониторинга здоровья и в целом повышения доступности медицинской помощи. Каждая из этих систем имеет свои преимущества и недостатки. Среди преимуществ можно выделить легкодоступность для жителей тех стран, где эти системы распространены. Пользователь имеет круглосуточный доступ к медицинской помощи, при этом снижается нагрузка на медицинские учреждения, так как система помогает эффективно распределить потоки пациентов, направляя их к нужным специалистам. Использование такой системы способствует уменьшению случаев ошибочной самодиагностики путем выявления диагноза через интернет. Однако рассмотренные системы в настоящее время недоступны для использования в России. Кроме того, они характеризуются отсутствием локализации языка, что значительно ограничивает их использование среди пациентов, не владеющих английским языком; имеют ограниченный функционал, многие платформы не предоставляют в открытый доступ полный функционал. Среди существующих проблем выделяют ограничения, связанные с данными, на которых обучаются модели [13], этические аспекты использования подобных систем, возможное неправильное толкование полученной информации пациентами. Отмечается также, что популярные чат-боты дают общие рекомендации без учета состояния пользователя и особенностей его показателей здоровья [14].

Чат-боты не могут самостоятельно оказывать медицинскую помощь, они всё еще нуждаются в контроле со стороны человека для гарантии действительности предоставляемой информации. Тем не менее они представляют собой важный шаг в направлении цифровизации и улучшения доступности медицинской помощи. Важно продолжать развитие и совершенствование этих систем, тем самым увеличивая их точность и безопасность. Это позволит сделать чат-боты полноценными инструментами как для пациентов, так и для медицинских работников.

1.2. РАСПОЗНАВАНИЕ ИМЕНОВАННЫХ СУЩНОСТЕЙ ДЛЯ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ

Диалог пациента с врачом, как правило, не имеет какой-то четкой структуры. Поэтому главной задачей в разработке ассистента является извлечение необходимых сведений из диалогов в формате естественного языка. Для решения этой задачи возможно применение метода распознавания именованных сущностей – *Named Entity Recognition* (NER). С помощью данного метода [15] можно реализовать поиск, выделение и классификацию ключевых фрагментов текста, относящихся к медицинским терминам, например:

1) симптомы – сонливость, боль в горле, головокружение;
2) название анализов и их значений – ТТГ, ОАМ;
3) диагнозы – сахарный диабет, гипотиреоз.

Основной способ реализации NER – это методы машинного обучения, натренированные на размеченных медицинских текстах. Из моделей машинного обучения широкое распространение получили CRF (Conditional Random Fields – условное случайное поле) и LSTM (Long Short-Term Memory – длительная кратковременная память), однако модели на основе трансформеров, например BERT, способны демонстрировать более высокую точность, свыше 80 % [16], благодаря способности учитывать контекст предложения.

2. ОПИСАНИЕ ПОДХОДА

Основной задачей разрабатываемой системы является сбор ключевых данных о пациенте, обработка и анализ полученных сведений и постановка предварительного диагноза с перечнем необходимых анализов. Предварительный диагноз (или варианты диагнозов) может быть представлен только врачу, но скрыт от пациента. Выходной информацией, получаемой пациентом, является перечень необходимых анализов, результаты которых необходимо представить на очном приеме у врача-эндокринолога.

В ходе работы от медицинских специалистов в области эндокринологии были получены данные, представляющие собой перечень распространенных заболеваний, встречающихся в их практике, таких как сахарный диабет, гипотиреоз, гипертиреоз, феохромоцитома, синдром Конна, синдром Кушинга, надпочечниковая недостаточность. Данные включают описание наиболее характерных симптомов и жалоб для конкретного заболевания, а также наводящие вопросы, помогающие в диагностировании болезни. Для каждого заболевания указан необходимый перечень лабораторных и клинических исследований перед посещением профильного специалиста.

Разрабатываемая система построена по модульному принципу. Основу системы составляют три ключевых модуля, взаимодействующих друг с другом, для последовательной обработки информации и формирования диагностического заключения. Первый модуль системы отвечает за обработку текстовых жалоб пользователя с помощью модели извлечения симптомов. Ключевой особенностью этого модуля является его способность работать с разнообразными вариантами описания симптомов, включая разговорные и неформальные выражения. Это достигается благодаря семантической нормализации, которая преобразует разнообразные формулировки симптомов в стандартизированные медицинские категории. Выделенные симптомы передаются в модуль классификации диагнозов. Полученные данные анализируются, и на основе этого вычисляется вероятность наличия определенного заболевания. Если уверенность в диагнозе оказывается недостаточно высокой, система переходит в интерактивный режим, в котором задает пользователю уточняющие вопросы. Диалоговый модуль анализирует текущее состояние классификации и определяет, дополнительная информация о каком симптоме окажет наибольшее влияние на уточнение результатов. Вопросы к пользователю формулируются на естественном языке по заранее заданным шаблонам. Ответы пользователя обрабатываются с учетом возможных синонимичных выражений. Все модули интегрированы в единый сервисный слой, который координирует их взаимодействие, управляет потоком данных и формирует итоговый ответ для пользователя. Такая архитектура обеспечивает не только точность диагностики,

но и удобство использования. Разделение системы на модули позволяет при необходимости без высоких затрат корректировать модели и классификаторы без глобальных изменений всей архитектуры.

2.1. БАЗОВЫЙ АЛГОРИТМ ФУНКЦИОНИРОВАНИЯ ДИАЛОГОВОЙ СИСТЕМЫ

Процесс распознавания эндокринологических заболеваний можно разделить на три этапа (рис. 1): сбор анамнеза, обработка входных данных и классификация диагноза. В свободной форме пациент вводит жалобы, в которых описывает, что его беспокоит.

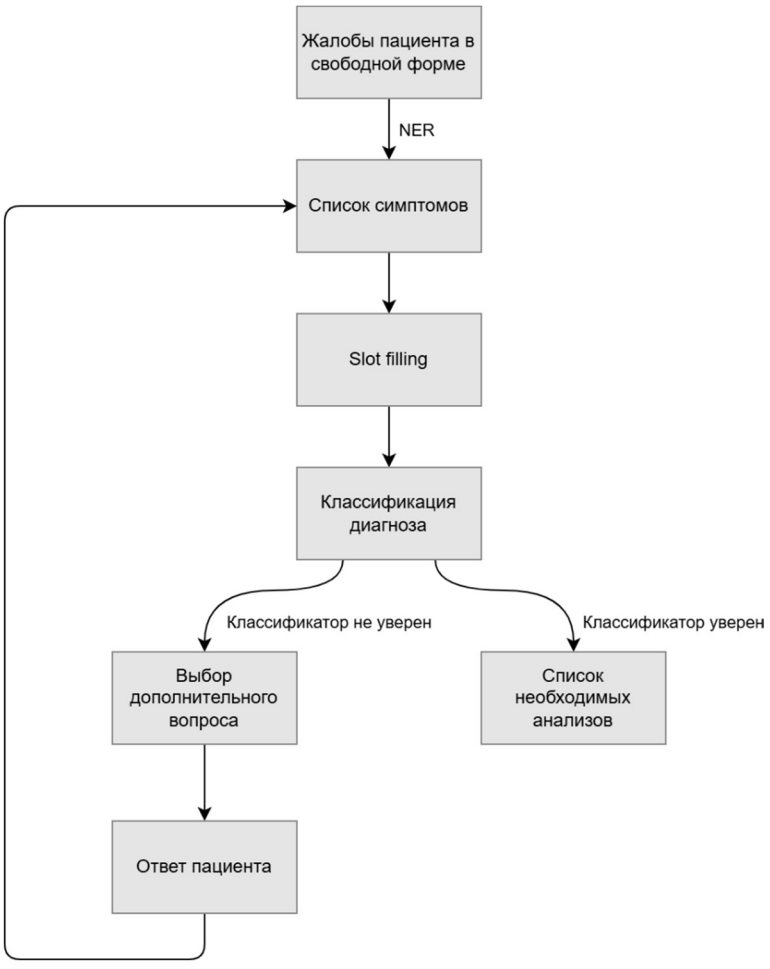

Рис. 1. Алгоритм работы системы

Fig. 1. System algorithm

Тексты уточняющих вопросов могут быть сгенерированы как заранее для каждого возможного симптома, так и во время диалога с подстановкой названия симптома в шаблоны вида «Вы испытывали {симптом} в последнее время?», «Отмечаете ли вы {симптом}?» и т. д. Однако важным аспектом является

генерация вопросов в закрытой форме, ответ на которую предполагает только два варианта – «да» или «нет». Такой подход призван снизить вычислительную нагрузку на систему поиска именованных сущностей и упростить последующую обработку ответов пользователя. Результаты ответа пользователя интерпретируются таким образом, что при получении ответа «да» признак наличия соответствующего симптома фиксируется положительным значением (единицей), при получении ответа «нет» – отрицательным значением (минус единицей). Эти значения записываются в специальную структуру, представляющую собой вектор симптомов, где каждому потенциальному симптому всегда соответствует определенное состояние: «точно имеется», «точно отсутствует» или «неизвестно».

Этап обработки начинается с применения к полученному тексту модели NER, которая обучена определять и извлекать сущности, относящиеся к симптомам. Далее выполняется *slot filling* – процесс заполнения матрицы симптомов. По полученному списку симптомов модель классификации пытается определить наиболее вероятный диагноз. Если классификатор не уверен, ему не хватает данных, то система возвращается на этап сбора анамнеза и инициирует процесс выбора уточняющего вопроса с целью получить больше данных. После ответа пациента система определяет и обновляет наличие симптомов и заново передает данные в модель классификатора. Система продолжает процесс выбора уточняющего вопроса до тех пор, пока классификатор не будет уверен.

В результате работы, если классификатор уверен и ему хватило входных данных, программа выдает возможный диагноз и список необходимых анализов, которые стоит сдать для подтверждения или опровержения поставленного диагноза. Для повышения точности определения диагноза необходимо предусмотреть механизм формирования дополнительных вопросов, который должен активироваться в случае недостаточной уверенности классификатора по результатам первичной обработки жалоб пациента в свободной форме. Основной целью такого механизма является получение дополнительной информации от пользователя для снижения неопределенности диагноза.

В качестве дополнительных вопросов подразумеваются уточняющие запросы о наличии или отсутствии симптомов, наиболее критичных для дифференцирования между несколькими потенциальными диагнозами. Поэтому для выбора вопросов следует руководствоваться принципом максимизации прироста информации и для выбора вопроса выбирать те симптомы, наличие или отсутствие которых наиболее существенно влияет на распределение вероятностей диагнозов. Таким образом предпочтение будет отдаваться симптомам, высоко специфичным для одного из возможных диагнозов, благодаря чему будет достигнут критерий минимизации числа вопросов.

Дополнительно, чтобы не перегружать пользователя, имеет смысл ограничить число задаваемых вопросов. Если модель классификации не смогла определить заболевание за ограниченный круг вопросов, то это может сигнализировать о том, что либо заболевание пользователя не имеет отношения к эндокринологии, либо случай пользователя слишком специфичный. В любом из двух возможных случаев, если модель долго остается неуверенной, лучше перенаправить пользователя к специалисту.

2.2. РЕАЛИЗАЦИЯ МОДУЛЯ ИЗВЛЕЧЕНИЯ СИМПТОМОВ

Модуль извлечения симптомов из текста является основным компонентом системы, так как именно он отвечает за обработку жалоб пациентов и выделение из них отдельных сущностей, связанных с симптомами. Основная задача этого модуля заключается в преобразовании неструктурированных текстовых данных, введенных пользователем, в структурированные данные в формате, пригодном для дальнейшего использования. Для решения этой задачи используются методы обработки естественного языка, включающие извлечение именованных сущностей и семантическую нормализацию.

Разработанный модуль включает в себя несколько этапов: подготовку данных, работу модели NER и нормализацию найденных симптомов. Подробно описываются этапы реализации модуля, такие как выбор и подготовка набора данных, обучение модели, оценка ее эффективности и методы нормализации симптомов.

Для обучения модели NER был взят открытый набор данных пользовательских отзывов о лекарственных препаратах на русском языке – *The Russian Drug Reaction Corpus* (RuDReC) [17], который предназначен для извлечения информации в области здравоохранения для обнаружения побочных эффектов. В настоящий момент набор данных состоит примерно из 5000 размеченных отзывов, собранных из различных интернет-источников. Из этого количества записей более половины содержат сущности, связанные с лекарствами и заболеваниями. Этот набор данных был выбран в работе в качестве основного, так как он является одним из немногих доступных открытых специализированных наборов данных на русском языке, содержащих качественную экспертную разметку медицинских сущностей. Также этот набор данных предоставляет достаточно репрезентативную выборку для дообучения языковых моделей.

Классы аннотированных сущностей в наборе данных представляют разнообразные аспекты, связанные с лекарствами и их воздействием, а также имеют дополнительную разметку связей на уровне выражений для идентификации тонких подтипов. Однако основные типы сущностей включают следующие классы.

1. Drug Name – упоминание наименований препаратов и активных веществ.
2. Drug Class – упоминание классов или предназначений лекарственных препаратов, как, например, «противовоспалительные», «иммуномодулирующие» и т. д.
3. Drug Form – форма выпуска лекарственного препарата.
4. Adverse Drug Reaction (ADR) – описанные пользователем побочные эффекты, связанные с приемом лекарств.
5. Drug Indication (DI) – показания к применению препарата, включающие в себя симптомы или упоминания заболеваний.
6. Findings – описанные пользователем упоминания изменений состояния здоровья, включающие в себя описания ухудшения или улучшения после приема препарата.

В рамках настоящей работы наибольшую ценность представляют последние три пункта, поскольку именно они содержат информацию о симптомах пользователей. По этой причине текстовые записи из набора данных были

отфильтрованы по критерию наличия аннотированных сущностей интересующих типов. Если запись содержала хотя бы одну сущность из классов ADR, DI или Findings, то она добавлялась в результирующий набор данных, при этом информация о разметке сущностей нецелевых типов в такой записи игнорировалась. Дополнительно отфильтрованный набор данных был дополнен записями, которые не содержат ни одной именованной сущности, чтобы эти записи составляли не более 10 % от общего размера набора данных. В результате размер итогового набора данных составил 1550 записей.

Для генерации синтетического набора данных использовалась большая языковая модель (БЯМ) *DeepSeek* [18]. В эту модель подавались подобранные эмпирическим путем задания на генерацию, составленные таким образом, чтобы модель перефразировала текст и/или заменяла упоминания симптомов в записях из RuDReC. Также модели выдавались задания на генерацию новых записей по реальным примерам из набора данных. Опытным путем было выявлено, что БЯМ плохо определяет позиции символов в тексте и не способна выдавать разметку в виде пары значений start и end. Для решения этой проблемы тексты заданий для модели были скорректированы таким образом, чтобы вместо позиций начала и конца именованной сущности модель выдавала в разметке вырезанный фрагмент текста сущности. Затем по этим фрагментам в исходной строке алгоритмическими методами определялись начальная и конечная позиция.

Также несмотря на то, что БЯМ хорошо справлялась с распознаванием упоминаний симптомов, у модели наблюдались проблемы с автоматической классификацией размеченных ею сущностей при отнесении симптомов к корректной категории. Отчасти это связано с отсутствием унифицированного стандарта наименований классов, из-за чего одни и те же симптомы могли быть отнесены к разным категориям. По этой причине наравне с тем, что и в RuDReC разметка не имеет указаний на конкретные типы симптомов, в разрабатываемой системе модель NER должна будет предсказывать только один обобщенный класс, сигнализирующий о том, что текстовая последовательность является описанием симптома, а дальнейшее уточнение его типа будет выполняться на последующих этапах обработки с использованием других методов. С помощью такого подхода удалось в короткие сроки сгенерировать набор синтетических данных, объем которого составил более 100 записей, что является достаточным для представления до модели NER информации о симптомах, специфических для эндокринологии. В совокупности сочетание реального набора данных с синтетически сгенерированными примерами позволит компенсировать нехватку релевантных в целевой предметной области данных и обеспечить большую адаптацию модели к специфике задачи, предоставив достаточно большой объем обучающих данных.

Для решения задачи NER была выбрана предварительно обученная, в том числе и на русскоязычных текстах, модель, основанная на архитектуре BERT cointegrated/LaBSE-en-ru. Исходный набор данных был разбит на две выборки: тренировочную и валидационную в соотношении 80 : 20 соответственно. Перед подачей данных в модель необходимо было провести их токенизацию. Исходный текст записи разбивался на подслова с помощью токенизатора, предоставляемого вместе с выбранной моделью, и дополнялся специальными токенами так, чтобы выровнять и привести все записи к единому размеру

в 256 токенов. Целевые метки сущностей также необходимо было привести в формат токенов с использованием BIO-нотации. Основная сложность заключается в том, что токенизатор может разбить одно слово из сущности на несколько подслов, в таком случае невозможно просто сопоставить номер слова из сущности с номером токена. Так как токенизатор предоставляет индекс начала и конца каждого токена в исходной строке, для преобразования меток из формата (start, end) в BIO-формат был использован следующий алгоритм.

1. Производится итерация по каждому интервалу из начала и конца токенов.

2. Осуществляется пропуск сущностей из разметки, конец которых меньше начала текущего интервала.

3. Если начало текущего интервала больше или равно началу сущности и конец интервала меньше или равен концу сущности из разметки, то проверяется предыдущий интервал:

а) если предыдущего интервала нет либо он не пересекается с текущей сущностью, то токену присваивается метка «B» начала сущности;

б) в ином случае токену присваивается метка «I» продолжения сущности.

4. Если интервал токена не пересекается с разметкой сущности, то такому токену присваивается метка «O».

С помощью описанного алгоритма получается разметка, в которой каждому токену из входящей последовательности сопоставлена целевая метка, которую будет пытаться предсказать модель. В цикле обучения на вход модели подается пакет токенизированных текстовых последовательностей. Для каждого входного токена модель предсказывает метку, вычисляется ошибка модели и с помощью оптимизатора корректируются веса модели. Обучение модели проводилось с помощью сервиса Google Colab, который представляет собой облачную среду для разработки и выполнения кода на языке Python. Также данный сервис предоставляет доступ к видеокартам Nvidia T4, которые, обладая видеопамятью в 24 Гб, являются незаменимыми при обучении языковых моделей. Гиперпараметры для обучения были подобраны эмпирическим путем. Основными параметрами при обучении стали оптимизатор AdamW с параметром learning_rate = 2e-5 и размером пакета 32. Обучение проводилось на 20 эпохах, после которых модель выходила на стабильное плато и переставала улучшать показатели.

При использовании модели в рабочей среде для получения предсказания необходимо выполнить аналогичную токенизацию текста и подать его на вход модели, затем по предсказанным меткам восстановить начало и конец предсказанных сущностей. В дополнение к предсказанию меток модель выдает некоторую меру уверенности в своем ответе, которая может быть использована для отсечения по заданному пороговому значению предсказаний, в которых модель слабо уверена.

2.3. МЕТРИКИ КАЧЕСТВА

Основными метриками при обучении моделей NER являются *F1-Score*, *Precision* и *Recall*. Метрика *Precision (PR)* характеризует, насколько уверенно

модель предсказывает сущности, т. е. сколько из всех предсказанных моделью сущностей было предсказано верно, и рассчитывается по формуле

$$PR = \frac{TP}{TP + FP},$$

где TP – количество истинно-положительных предсказаний, FP – количество ложноположительных предсказаний.

Метрика *Recall* (*RC*) является метрикой полноты и показывает, насколько внимательно модель находит сущности, т. е. сколько из всех реальных сущностей она нашла, и рассчитывается по формуле

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

где FN – количество ложнонегативных предсказаний.

Метрика *F1-Score (F1S)* представляет собой гармоническое среднее между точностью и полнотой и рассчитывается по формуле

$$F1S = 2\frac{PR \cdot RC}{PR + RC}.$$

В рамках данной задачи интерес представляют высокие показатели точности и в меньшей степени полнота, так как если будет пропущена какая-либо сущность, это не так критично, потому что она потом будет доопределена на последующих этапах общения с пользователем. Однако сильный дисбаланс между этими метриками является негативным маркером того, что модель обучилась плохо, поэтому метрика *F1-Score* также важна как показатель баланса между точностью и полнотой модели. Как правило, признаком дисбаланса между *Precision* и *Recall* является существенное различие значений этих метрик, при этом *F1-Score* заметно ниже обоих показателей.

В результате обучения модели были достигнуты следующие показатели метрик: $F1S$ = 0,972 и PR = 0,971. Однако полученные метрики не в полной мере отображают действительную способность модели предсказывать метки именованных сущностей, так как наблюдается колоссальный дисбаланс, смещенный в сторону метки «О», который связан с тем, что в естественных данных большая часть слов не является целевыми медицинскими именованными сущностями. Таким образом, эти метрики в большей степени показывают способность модели определять, что слово не является именованной сущностью. Поэтому для более объективной оценки использовались метрики $F1S_{clear}$ и PR_{clear}. Для их подсчета не учитывались пары, в которых целевой меткой была метрика «О», и модель предсказала метку «О». Эти метрики обозначены индексом *clear* и вычисляются по следующим формулам:

$$TP_{clear} = \sum_{i=1}^{N}\left[y_{true}^{(i)} \neq O \wedge y_{pred}^{(i)} \neq O \wedge y_{true}^{(i)} = y_{pred}^{(i)} \right],$$

где $y_{true}^{(i)}$ – истинное значение целевой переменной, $y_{pred}^{(i)}$ – спрогнозированное значение целевой переменной, $\wedge$ – логическое «И», O – метка «О»;

$$FP_{clear} = \sum_{i=1}^{N}\left[\left(y_{true}^{(i)} \neq O \wedge y_{pred}^{(i)} \neq O \wedge y_{true}^{(i)} = y_{pred}^{(i)}\right) \vee \left(y_{true}^{(i)} = O \wedge y_{pred}^{(i)} \neq O\right)\right],$$

где $\vee$ – логическое «ИЛИ»;

$$FN_{clear} = \sum_{i=1}^{N}\left[y_{true}^{(i)} \neq O \wedge y_{pred}^{(i)} = O\right],$$

$$PR_{clear} = \frac{TP_{clear}}{TP_{clear} + FP_{clear}},$$

$$RC_{clear} = \frac{TP_{clear}}{TP_{clear} + FN_{clear}},$$

$$F1S_{clear} = 2\frac{PR_{clear} \cdot RC_{clear}}{PR_{clear} + RC_{clear}}.$$

Эти метрики помогают более достоверно оценить качество распознавания именно сущностей и показывают, как модель справляется только с именованными сущностями, без искажения из-за дисбаланса меток «O». Для обученной модели метрики составили: $F1S_{clear} = 0{,}951$ и $PR_{clear} = 0{,}950$.

Также для наглядной визуализации оценки обученной модели были построены матрицы ошибок (*confusion matrix*) в нескольких вариантах. Матрица ошибок широко используется для оценки качества классификации в задачах машинного обучения и представляет собой квадратную матрицу, элементы которой отражают частоту соответствия между истинными и предсказанными метками классов. Строки матрицы соответствуют истинному классу, а столбцы – предсказанному моделью, при этом элементы на главной диагонали характеризуют корректно классифицированные объекты, а остальные элементы описывают случаи ошибочной классификации.

На рис. 2 представлена стандартная матрица ошибок, где видно преобладание диагональных элементов для класса «O», что подтверждает его доминирование в данных. Для того чтобы более детально показать качество распознавания именно целевых медицинских сущностей, была построена сфокусированная матрица, в которой аналогично подсчету метрик выше были исключены случаи верно предсказанных метрик «O».

На рис. 3 представлена сфокусированная версия матрицы ошибок, где видно, что модель с достаточной точностью выделяет медицинские сущности из текста с минимальными перекрестными ошибками между метками начала и продолжения сущности.

Дополнительно для комплексной оценки производительности модели NER целесообразно проанализировать не только раздельные метки в BIO-формате, но и агрегированный результат, в котором метки начала и продолжения сущности объединены в общий класс симптома *SYM*. Такая агрегированная матрица (рис. 4) позволяет оценить, насколько корректно модель идентифицирует сущности в целом, так как при использовании модели в режиме предсказания некоторые ошибки в разметке начальных и внутренних меток возможно исправить алгоритмическими методами.

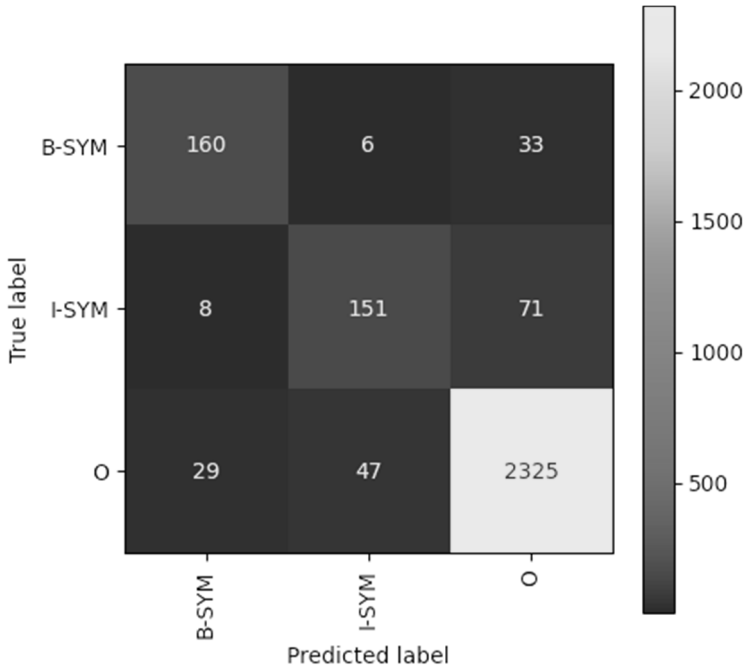

Рис. 2. Матрица ошибок с учетом класса «O»

Fig. 2. The error matrix taking into account class “O”

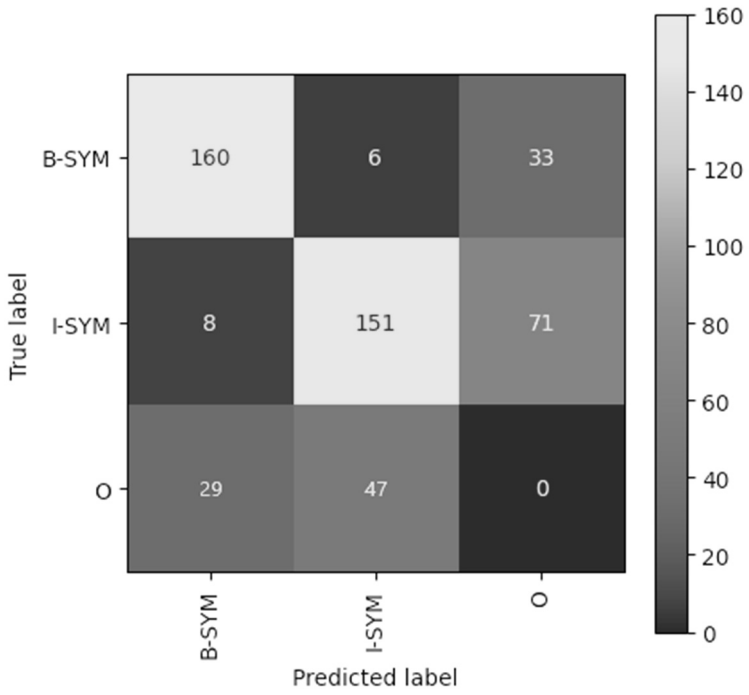

Рис. 3. Сфокусированная матрица ошибок

Fig. 3. The focused error matrix

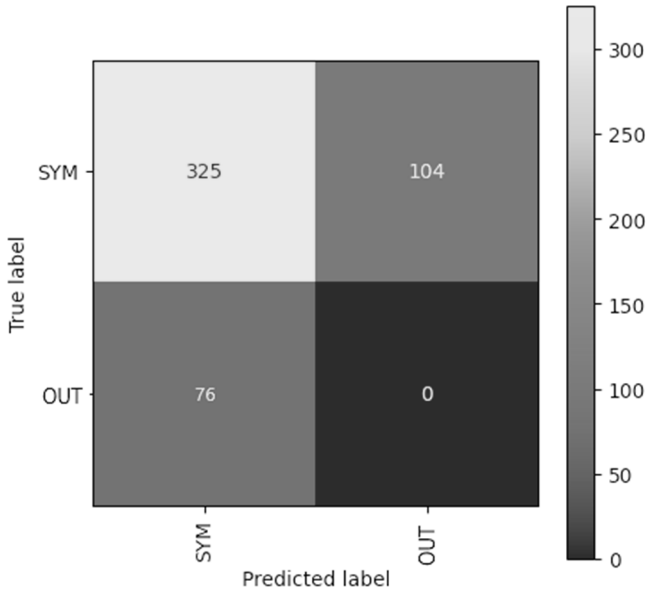

Рис. 4. Агрегированная матрица ошибок

Fig. 4. The aggregated error matrix

Таким образом, приведенные метрики качества и их визуализация в виде матриц ошибок подтверждают эффективность обученной модели NER в задаче распознавания симптомов пациентов.

2.4. НОРМАЛИЗАЦИЯ СИМПТОМОВ

Несмотря на достаточное качество распознавания сущностей описанной выше моделью, она предсказывает только факт наличия симптома, но не определяет его конкретный класс. Поэтому распознанные моделью симптомы требуют дальнейшей нормализации для отнесения их к конкретным медицинским категориям. Для решения этой задачи был разработан и реализован дополнительный модуль нормализации симптомов. Для приведения симптомов к стандартизированному виду используется метод семантической нормализации. Эталонные классы симптомов представлены в виде набора синонимичных формулировок, что обеспечивает адаптивность к разнообразию формулировок симптомов пациентами. Алгоритм нормализации использует предобученную модель LaBSE-en-ru для перевода эталонных формулировок и текста предсказанной сущности в векторные представления. Далее вектор предсказанного текста поочередно сравнивается с каждым вектором из набора эталонных примеров для каждого возможного класса симптомов, используя меру косинусного сходства *similarity* (SMT), вычисляемую по формуле [19]

$$SMT(A,\ B) = \frac{AB}{\| A \| \| B \|} = \frac{\sum_{i=1}^{N} a_i b_i}{\sqrt{\sum_{i=1}^{N} a_i^{\,2}}\ \sqrt{\sum_{i=1}^{N} b_i^{\,2}}},$$

где A и B – векторы текста сущности и эталонного примера соответственно.

Косинусное сходство позволяет оценить близость понятий, текст преобразуется в векторное пространство, такое что близкие по смыслу фразы имеют схожие направления векторов, даже если они выражены разными словами. Также возможно сравнение векторов с использованием евклидова расстояния, однако мера косинусного сходства не зависит от длины самих векторов. Инвариантность к длине текста важна при сравнении коротких названий симптомов с их развернутыми описаниями.

После вычисления косинусных расстояний вычисляется максимальная близость для каждого класса. Если близость превышает заданное пороговое значение, то симптом относится к соответствующему классу, в противном случае помечается как неизвестный. Более формализованное описание данного алгоритма включает в себя заданное входное множество классов симптомов:

$$C = \{c_1,\ c_2,...,\ c_n\}\ .$$

Для каждого класса симптомов задается множество синонимических наиболее распространенных названий этого симптома и разговорных формулировок его проявления:

$$E_c = \{e_1,\ e_2,...,\ e_n\}\ .$$

Векторное представление текста t в эмбеддинг v_t с помощью модели M задается формулой

$$v_t = M(t)\ .$$

Тогда для входного текста распознанной сущности t класс симптома вычисляется следующей формулой:

$$c'(t) = \arg\max_{c \in C} \left(\max_{e \in E_c} SMT(M(t),\ M(e)) \right),\ \text{если}\ \max \geq \theta,$$

иначе *unknown*.

Здесь c' – предсказанный класс симптома; t – входной текст симптома, распознанного моделью NER; C – множество классов симптомов; E_c – множество примеров для конкретного класса симптомов; M – модель векторизации; θ – порог уверенности.

Из формулы следует, что если близость для всех классов ниже некоторого порогового значения θ, то симптом помечается как *unknown*. Такая фильтрация обеспечивает возврат результата только при высокой степени схожести, что снижает количество ложных срабатываний и помогает отбросить симптомы, не имеющие отношения к эндокринологии. Подобные симптомы не передаются дальше в модель классификации. Таким образом, входной текст классифицируется по ближайшему классу на основе анализа векторных представлений, что позволяет учитывать разнообразие формулировок пациентов. Благодаря этому нормализатор преобразует разрозненные описания жалоб пациентов в стандартизированные классы, что позволяет перейти от распознавания факта упоминания симптома к его содержательной интерпретации, необходимой для последующего анализа и предсказания предварительного диагноза.

3. КЛАССИФИКАЦИЯ ЭНДОКРИНОЛОГИЧЕСКОГО ДИАГНОЗА

Для определения предварительного диагноза пациента по его жалобам был реализован модуль классификации по списку выявленных симптомов. В рамках настоящей работы были выбраны два алгоритма классификации: логистическая регрессия и наивный байесовский классификатор. Такой выбор обусловлен эффективностью этих алгоритмов в задачах с ограниченным объемом данных и интерпретируемостью их результатов. Логистическая регрессия способна моделировать вероятности принадлежности к классам через сигмоидальную функцию:

$$(y = 1 \vee x) = \frac{1}{1 + e^{-(wTx+b)}},$$

где w – веса признаков, T – операция транспонирования, b – смещение, x – вектор симптомов.

Обучение модели логистической регрессии [20] сводится к нахождению параметров w и b, минимизирующих функцию потерь L:

$$L(w,\ b) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(P(y_i = 1 \vee x_i)\right) + (1 - y_i)\log\left(1 - P(y_i = 1 \vee x_i)\right)\right],$$

где N – количество примеров в выборке, y_i – истинная метка класса для i-го примера.

Подбор параметров w и b выполняется методом градиентного спуска на протяжении заданного числа итераций:

$$w^{t+1} = w^t - \eta \nabla_w L,$$

$$b^{t+1} = b^t - \eta \frac{\partial L}{\partial b},$$

где η – скорость обучения, t – номер итерации, η – скорость обучения, $\nabla_w L$ – градиент функции потерь L по весам w, $\frac{\partial L}{\partial b}$ – частная производная функции по смещению b.

Наивный байесовский классификатор [21] основывается на теореме Байеса с предположением о независимости признаков:

$$P(y \vee x) = \frac{P(y)\prod_{i=1}^{n} P(x_i \vee y)}{P(x)}.$$

Априорная вероятность $P(y)$ вычисляется как доля примеров класса y в обучающей выборке:

$$P(y) = \frac{NCL_y}{N},$$

где NCL_y – количество примеров класса y.

Условные вероятности $P(x_i \vee y)$ для категориальных признаков вычисляются как

$$P(x_i \vee y) = \frac{NCL_{y(x_i)}}{NCL_y},$$

где $NCL_{y(xi)}$ – количество примеров класса y с признаками x_i.

Исходные данные для обучения алгоритмов были преобразованы в файл формата CSV, содержащего бинарные признаки наличия симптомов и текстовые названия возможных диагнозов. Признаки в файле закодированы таким образом, что значение «1» обозначат наличие симптома, а значение «–1» – отсутствие симптома. Текстовые названия диагнозов являются целевой переменной, которую должны предсказывать выбранные алгоритмы. Однако в силу возможности их работы только с числовыми переменными необходимо применение метода LabelEncoder, который преобразует текстовые метки в числовые и хранит в себе список соответствия. Оба алгоритма были обучены на одном наборе данных, разделенных на признаки и целевую переменную. Гиперпараметры моделей, такие как регуляризация для логистической регрессии или априорные вероятности для байесовского классификатора, использовались со значениями по умолчанию, так как объем данных не требовал сложной настройки. Обучение проводилось на полном наборе данных без выделения отдельной тестовой выборки из-за его ограниченного размера.

Модели после обучения были сохранены в виде объекта, готового к использованию в реальном времени. При поступлении списка симптомов, выявленных у пользователя, они преобразуются в вектор признаков, согласованный с порядком колонок в тренировочных данных. При этом значение «1» соответствует установленному у пациента наличию симптома, значение «–1» – отсутствию симптома, а пропущенные значения заполняются значением «0», показывающим, что информация о симптоме не предоставлена. Это позволяет модели обрабатывать случаи, когда информация о симптоме еще не была получена от пользователя. Диалог с пользователем реализован в виде консольного приложения. В результате модуль возвращает вероятности принадлежности набора симптомов к каждому возможному диагнозу, что позволяет ранжировать результаты и определить наиболее вероятный вариант. Настройки системы позволяют отображать только рекомендуемые для сдачи анализы, а предполагаемый диагноз демонстрировать врачу. В случае недостаточной уверенности модуля в наиболее вероятном варианте необходима дальнейшая его проработка и получение от пользователя дополнительной информации о наличии или отсутствии симптомов, информация о которых еще не известна.

ЗАКЛЮЧЕНИЕ

Представленная диалоговая система предназначена для поддержки принятия врачебных решений и упрощения процесса приема пациентов в условиях недостатка квалифицированных специалистов. В текущей версии система охватывает классификацию наиболее распространенных эндокринологических

заболеваний. Расширение этого списка требует сбора дополнительных данных. Большинство популярных рассмотренных диалоговых систем являются зарубежными разработками и обладают широким перечнем функций: от предоставления справочной информации по общим медицинским вопросам и рекомендаций по выбору специалиста до осуществления записи к врачу. Особенностью предложенной системы является привязка к конкретной отрасли медицины. В отличие от многих чат-ботов, ориентированных преимущественно на удобство пациентов, предложенная разработка позволяет снизить нагрузку на медицинских специалистов за счет сбора анамнеза и обработки полученных результатов. Основным недостатком системы является обучение на наборах данных малого объема, которые содержат ограниченное количество примеров, специфичных для эндокринологии, что может снижать точность редких или сложных случаев. Также система может испытывать трудности с пониманием неоднозначных формулировок жалоб.

Модель обработки естественного языка ориентирована на обработку только русскоязычных текстов, что может ограничивать ее применение в международных масштабах. Однако именно обработка только русскоязычных текстов делает систему уникальной среди популярных чат-ботов, осуществляющих диалог только на английском языке. Таким образом, на данный момент система не заменяет профессиональную медицинскую диагностику, а лишь оптимизирует процесс сбора первичной информации, предоставляя предварительные диагнозы, которые требуют обязательной проверки врачом.

Несмотря на текущие ограничения, система обладает значительным потенциалом для развития и внедрения в практическое здравоохранение. Дальнейшие исследования и доработки позволят повысить ее точность, универсальность и полезность для пациентов и врачей. Основными направлениями дальнейших исследований являются работы по расширению обучающих наборов данных, а также применение системы для других областей медицины.

Благодарности. Авторы выражают благодарность профессору кафедры вычислительной техники, доктору медицинских наук Мелдо Анне Александровне за консультирование в проведении исследования и разработки.

СПИСОК ЛИТЕРАТУРЫ

1. *Терещенко И.В.* Пропедевтика эндокринных заболеваний (методы исследования больных с заболеваниями эндокринной системы): монография. – Чебоксары: Среда, 2022. – 104 с. – ISBN 978-5-907411-89-0. – DOI: 10.31483/a-10348.

2. *Тимофеев Е.В., Парфенова Н.Н., Реева С.В.* Обследование эндокринной и кроветворной систем в клинике внутренних болезней: учебное пособие. – СПб.: Изд-во С.-Петерб. гос. педиатр. мед. ун-та, 2022. – 32 с.

3. *Медведева Е.И., Александрова О.А., Крошилин С.В.* Телемедицина в современных условиях: отношение социума и вектор развития // Экономические и социальные перемены: факты, тенденции, прогноз. – 2022. – Т. 15, № 3. – С. 200–222. – DOI: 10.15838/esc.2022.3.81.11.

4. *Баранов А.А, Вишнева Е.А., Намазова-Баранова Л.С.* Телемедицина – перспективы и трудности перед новым этапом развития // Педиатрическая фармакология. – 2013. – Т. 10, № 3. – С. 6–11. – DOI: 10.15690/pf.v10i3.691.

5. *Хайруллин А.М., Зарипова Р.С.* Использование чат-ботов и онлайн-консультации как будущее медицины // Информационные технологии в строительных, социальных и экономических системах. – 2020. – № 3 (21). – С. 74–76.

6. *Аксёнова Е.И., Медведева Е.И., Крошилин С.В.* Чат-боты – современная реальность консультирования в медицине // Здравоохранение Российской Федерации. – 2023. – Т. 67, № 5. – С. 403–410. – DOI: 10.47470/0044-197X-2023-67-5-403-410.

7. Babylon. – URL: https:// www.babylonhealth.com (accessed: 04.02.2025).

8. Ada Health. – URL: https://ada.com (accessed: 04.02.2025).

9. ByoHealth. – URL: https://www.buoyhealth.com (accessed: 04.02.2025).

10. Florence. – URL: https://florencehc.com (accessed: 04.02.2025).

11. Symptomate. – URL: https://symptomate.com (accessed: 04.02.2025).

12. Sensely. – URL: https://sensely.com (accessed 04.02.2025).

13. *McMahon G.T.* The risks and challenges of artificial intelligence in endocrinology // The Journal of Clinical Endocrinology & Metabolism. – 2024. – Vol. 109 (6). – P. e1468-e1471.

14. Roles, users, benefits, and limitations of chatbots in health care: rapid review / M. Laymouna, Y. Ma, D. Lessard, T. Schuster, K. Engler, B. Lebouche // Journal of Medical Internet Research. – 2024. – Vol. 26. – P. e56930.

15. Feature extraction method from electronic health records in Russia / D. Gavrilov, A. Gusev, I. Korsakov, R. Novitskiy, L. Serova // Proceedings of the 26th Conference of FRUCT Association. – 2020. – P. 497–500. – ISBN 978-952-69244-2-7.

16. The Russian drug reaction corpus and neural models for drug reactions and effectiveness detection in user reviews / E. Tutubalina, I. Alimova, Z. Miftahutdinov, A. Sakhovskiy, V. Malykh, S. Nikolenko // Bioinformatics. – 2020. – Vol. 36 (2). – P. 243–250. – DOI: 10.1093/bioinformatics/btaa675.

17. Exploring DeepSeek: A survey on advances, applications, challenges and future directions / Z. Deng, W. Ma, Q.-L. Han, W. Zhou, X. Zhu, S. Wen, Y. Xiang // IEEE/CAA Journal of Automatica Sinica. – 2025. – Vol. 12 (5). – P. 872–893.

18. *Chopra S., Hadsell R., LeCun Y.* Learning a similarity metric discriminatively, with application to face verification // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). – 2005. – Vol. 1 – P. 539–546.

19. *Han J., Morag C.* The influence of the sigmoid function parameters on the speed of backpropagation learning // From natural to artificial neural computation / ed. by J. Mira, F. Sandoval. – Springer, 1995. – P. 195–201. – (Lecture Notes in Computer Science; vol. 930).

20. *Maron M.E.* Automatic indexing: An experimental inquiry // Journal of the ACM. – 1961. – Vol. 8 (3). – P. 404–417.

Мамикьян Маргарита Сергеевна, студентка кафедры вычислительной техники Санкт-Петербургского государственного электротехнического университета «ЛЭТИ». Основное направление научных исследований – машинное обучение, системный анализ, информационные системы. E-mail: mmamikyan2001@gmail.com

Бекенева Яна Андреевна, кандидат технических наук, доцент кафедры вычислительной техники Санкт-Петербургского государственного электротехнического университета «ЛЭТИ». Основное направление научных исследований – методы обработки данных, машинное обучение, информационные системы. Имеет более 50 печатных работ. E-mail: yabekeneva@etu.ru

Mamikyan Margarita S., student at the Department of Computer Science and Engineering, Saint Petersburg Electrotechnical University “LETI”. Her research interests are currently focused on machine learning, system analysis, and information systems. E-mail: mmamikyan2001@gmail.com

Bekeneva Yana A., PhD (Eng.), associate professor at the Department of Computer Science and Engineering, Saint Petersburg Electrotechnical University “LETI”. Her research interests are currently focused on data processing, machine learning, and information systems. She has more than 50 publications. E-mail: yabekeneva@etu.ru

DOI: 10.17212/2782-2001-2026-3-75-96

Processing text data in medical dialogue systems using the example of a virtual assistant for an endocrinologist [*]

M.S. MAMIKYAN[a], *YA.A. BEKENEVA*[b]

Saint Petersburg Electrotechnical University "LETI", 5 Professora Popova Street, St-Petersburg, 197022, Russian Federation

[a] *mmamikyan2001@gmail.com* [b] *yabekeneva@etu.*

Abstract

The proportion of patients with endocrine pathologies is constantly increasing in all economically developed countries of the world. In Russia, over the past 6 years, the prevalence of endocrine diseases has increased by 26%, and primary morbidity rates by 36%. One of the main problems that aggravate the epidemiological situation is the shortage of local therapists and the uneven distribution of medical specialists between rural and urban regions. Residents of sparsely populated areas are forced to travel to neighboring cities to receive the necessary medical care, which entails additional time and financial costs. It is also worth noting that determining the necessary tests for delivery can take a lot of time, and rushing during the appointment can lead to the delivery of redundant tests. To optimize this process, it is important to determine in advance what tests the patient needs, which requires a preliminary diagnosis. A disease recognition system based on patient complaints can help with this task, which can automate the preliminary diagnosis and avoid unnecessary studies. The presented software solution is based on a combination of named entity recongition for extracting symptoms from free-form patient complaints and machine learning methods for determining a diagnosis. The RuDRec dataset was used for training, along with synthetic data. An algorithm for dialogue with the user to clarify symptoms and generate a list of necessary tests is described. Such a system will help reduce cases of self-medication, when patients, without precise recommendations, follow the advice of friends. Thus, the formation of directions for tests based on high-quality primary consultation of patients, especially in sparsely populated areas, is an urgent task. The practical significance of the work lies in optimizing the load on medical institutions, reducing the risks of self-medication and improving diagnostics in remote regions.

Keywords: virtual assistant of endocrinologist, medical data processing methods, large language models, chatbot, dialog interface, health complaint, diagnostics, endocrinological disease, telemedicine, named entities, intelligent machine learning system

REFERENCES

1. Tereshchenko I.V. *Propedevtika endokrinnykh zabolevanii (metody issledovaniya bol'nykh s zabolevaniyami endokrinnoi sistemy)* [Propaedeutics of endocrinisms (Methods of examination of patients with diseases of the endocrine system)] Cheboksary, Sreda Publ., 2022. 104 p. ISBN 978-5-907411-89-0. DOI: 10.31483/a-10348.

2. Timofeev E.V., Parfenova N.N., Reeva S.V. *Obsledovanie endokrinnoi i krovetvornoi sistem v klinike vnutrennikh boleznei* [Examination of the endocrine and hematopoietic systems in the clinic of internal diseases]. St. Petersburg, St. Petersburg State Pediatric Medical University Publ., 2022. 32 p.

3. Medvedeva E.I., Aleksandrova O.A., Kroshilin S.V. Telemeditsina v sovremennykh usloviyakh: otnoshenie sotsiuma i vektor razvitiya [Telemedicine in modern conditions: the attitude of society and the vector of development]. *Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz = Economic and Social Changes: Facts, Trends, Forecast*, 2022, vol. 15, no. 3, pp. 200–222. DOI: 10.15838/esc.2022.3.81.11.

[*] *Received 06 June 2025.*

4. Baranov A.A, Vishneva E.A., Namazova-Baranova L.S. Telemeditsina – perspektivy i trudnosti pered novym etapom razvitiya [Telemedicine – prospects and difficulties before a new stage of development]. *Pediatricheskaya farmakologiya = Pediatric Pharmacology*, 2013, vol. 10, no. 3, pp. 6–11. DOI: 10.15690/pf.v10i3.691.

5. Khairullin A.M., Zaripova R.S. Ispol'zovanie chat-botov i onlain-konsul'tatsii kak budushchee meditsiny [Using chat bots and online consultation as the future of medicine]. *Informatsionnye tekhnologii v stroitel'nykh, sotsial'nykh i ekonomicheskikh sistemakh*, 2020, no. 3 (21), pp. 74–76. (In Russian).

6. Aksenova E.I., Medvedeva E.I., Kroshilin S.V. Chat-boty – sovremennaya real'nost' konsul'tirovaniya v meditsine [Chatbots is the modern reality of consulting in medicine]. *Zdravookhranenie Rossiiskoi Federatsii = Health Care of the Russian Federation*, 2023, vol. 67, no. 5, pp. 403–410. DOI: 10.47470/0044-197X-2023-67-5-403-410.

7. *Babylon*. Available at: https:// www.babylonhealth.com (accessed 04.02.2025).

8. *Ada Health*. Available at: https://ada.com (accessed 04.02.2025).

9. *ByoHealth*. Available at: https://www.buoyhealth.com (accessed 04.02.2025).

10. *Florence.* Available at://florencehc.com (accessed 04.02.2025).

11. *Symptomate*. Available at: https://symptomate.com (accessed 04.02.2025).

12. *Sensely*. Available at: https://sensely.com (accessed 04.02.2025).

13. McMahon G.T. The risks and challenges of artificial intelligence in endocrinology. *The Journal of Clinical Endocrinology & Metabolism*, 2024, vol. 109 (6), p. e1468-e1471.

14. Laymouna M., Ma Y., Lessard D., Schuster T., Engler K., Lebouche B. Roles, users, benefits, and limitations of chatbots in health care: rapid review. *Journal of Medical Internet Research*, 2024, vol. 26, p. e56930.

15. Gavrilov D., Gusev A., Korsakov I., Novitskiy R., Serova L. Feature extraction method from electronic health records in Russia. *Proceedings of the 26th Conference of FRUCT Association*, 2020, pp. 497–500. ISBN 978-952-69244-2-7.

16. Tutubalina E., Alimova I., Miftahutdinov Z., Sakhovskiy A., Malykh V., Nikolenko S. The Russian drug reaction corpus and neural models for drug reactions and effectiveness detection in user reviews. *Bioinformatics*, 2020, vol. 36 (2), pp. 243–250. DOI: 10.1093/bioinformatics/btaa675.

17. Deng Z., Ma W., Han Q.-L., Zhou W., Zhu X., Wen S., Xiang Y. Exploring DeepSeek: A survey on advances, applications, challenges and future directions. *IEEE/CAA Journal of Automatica Sinica*, 2025, vol. 12 (5), pp. 872–893.

18. Chopra S., Hadsell R., LeCun Y. Learning a similarity metric discriminatively, with application to face verification. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 1, pp. 539–546.

19. Han J., Morag C. The influence of the sigmoid function parameters on the speed of backpropagation learning. *From Natural to Artificial Neural Computation*. Ed. by J. Mira, F. Sandoval. *Lecture Notes in Computer Science*, vol. 930. Springer, 1995, pp. 195–201.

20. Maron M.E. Automatic indexing: An experimental inquiry. *Journal of the ACM*, 1961, vol. 8 (3), pp. 404–417.

Для цитирования:

Мамикьян М.С., Бекенева Я.А. Обработка текстовых данных в медицинских диалоговых системах на примере виртуального ассистента эндокринолога // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 75–96. – DOI: 10.17212/2782-2001-2026-3-75-96.

For citation:

Mamikyan M.S., Bekeneva Ya.A. Obrabotka tekstovykh dannykh v meditsinskikh dialogovykh sistemakh na primere virtual'nogo assistenta endokrinologa [Processing text data in medical dialogue systems using the example of a virtual assistant for an endocrinologist]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 75–96. DOI: 10.17212/2782-2001-2026-3-75-96.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 97–114

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 97–114

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 519.876

DOI: 10.17212/2782-2001-2026-3-97-114

Применение алгоритма муравьиной колонии в оптимизации управления мультипроектными разработками*

Ю.А. МЕЗЕНЦЕВ[a], А.А. БОРИСОВА[b]

РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет

[a] *mesyan@yandex.ru* [b] *alina_10112004@mail.ru*

В статье рассматривается применение муравьиного алгоритма (Ant Colony Optimization, ACO) для решения задачи оптимизации управления мультипроектными разработками с ограниченными ресурсами. Мультипроектная среда характеризуется конкуренцией между задачами за ресурсы, технологическими и ресурсными зависимостями, что делает задачу комбинаторной и NP-трудной.

В работе описан алгоритм муравьиной колонии, включающий формирование допустимых расписаний одновременно несколькими искусственными муравьями, вероятностный выбор операций с учетом уровня феромона и эвристической информации, а также обновление феромонных следов в процессе итераций. Проведены эксперименты на задачах различной размерности, показавшие, что при корректной настройке параметров алгоритм стабильно находит оптимальные или близкие к оптимальным решения. Для малых задач решение полностью совпадает с результатами точной оптимизации в IBM ILOG CPLEX, для средних и больших задач метод обеспечивает эффективное нахождение качественных расписаний за существенно меньшее время по сравнению с точными методами.

Результаты исследования демонстрируют полиномиальную вычислительную трудоемкость алгоритма ACO и его высокую эффективность при планировании мультипроектных разработок, что делает его перспективным инструментом для применения в практических задачах управления ресурсами и сроками проектов.

Ключевые слова: мультипроектное планирование, комбинаторная оптимизация, алгоритм муравьиной колонии, NP-трудная задача, управление ресурсами, смешанные графы, календарное планирование

* *Статья получена 20 марта 2026 г.*

ВВЕДЕНИЕ

Сегодня разработки программных и инженерных систем всё чаще планируются с помощью проектного подхода. Суть такого планирования состоит в том, что множество одновременно выполняющихся проектов состоит из каких-либо задач, которые должны выполнить ограниченное число исполнителей. При этом исполнитель является ресурсом, без которого задача не может быть начата. Такая мультипроектная среда характеризуется ограниченностью ресурсов, конкуренцией между проектами и необходимостью эффективного планирования сроков исполнения проектов [1, 2].

Ресурс в задачах планирования и управления проектами понимается как любой ограниченный объект, средство или субъект, необходимый для выполнения задачи. Использование ресурса связано с ограничениями по его количеству, доступности или времени применения. Задача может быть начата только при наличии необходимых ресурсов, поэтому в мультипроектной среде несколько задач могут одновременно конкурировать за одни и те же ресурсы. Это приводит к необходимости их распределения во времени и между различными проектами.

Можно выделить разделяемые и неразделимые ресурсы [3]. Разделяемые ресурсы могут использоваться несколькими задачами одновременно, при этом их общий объем распределяется между задачами. Каждая задача получает лишь часть ресурса, а суммарное использование не должно превышать его доступный объем. К таким ресурсам относятся, например, вычислительные мощности, финансовые средства или пропускная способность каналов связи. Разделяемые ресурсы могут быть как складируемыми, так и нескладируемыми. Складируемые ресурсы обладают свойством накопления, т. е. неиспользованный в данный момент ресурс может быть сохранен и использован позже. Примером таких ресурсов могут служить материалы, комплектующие или финансовые средства. Нескладируемые ресурсы, напротив, не могут накапливаться: если они не используются в определенный момент времени, возможность их использования теряется. К таким ресурсам относится, например, рабочее время специалистов или вычислительное время оборудования.

Неразделимые ресурсы характеризуются тем, что в один момент времени они могут быть использованы только одной задачей. Если такой ресурс занят выполнением некоторой задачи, остальные задачи вынуждены ожидать его освобождения. В качестве примеров неразделимых ресурсов можно рассматривать конкретных специалистов, производственное оборудование или уникальные технические установки. Наличие неразделимых ресурсов в системе приводит к возникновению конфликтов доступа между задачами и существенно усложняет процесс планирования [4].

Исполнитель в нашей задаче является неразделимым ресурсом. Это означает, что только после окончания текущей операции исполнитель начнет следующую. Таким образом, получается, что помимо технологической последовательности есть также ресурсная.

Основной сложностью мультипроектного управления является распределение во времени ресурсов между задачами различных проектов с целью минимизации общего времени выполнения и повышения эффективности использования ресурсов. Эта задача относится к классу комбинаторных оптимизационных задач и является NP-трудной (Nondeterministic Polynomial-Time Hard – недетерминированно полиномиально трудная), т. е. задачей, для которой

не существует известного алгоритма с полиномиальной зависимостью времени решения от размерности задачи [5].

В последние годы для решения подобных задач активно применяются методы метаэвристической оптимизации. Одним из наиболее перспективных является муравьиный алгоритм (Ant Colony Optimization), основанный на моделировании поведения колоний муравьев при поиске кратчайших путей.

Муравьиный алгоритм напрямую относится к направлению роевого интеллекта и является одним из наиболее известных его представителей. Роевой интеллект представляет собой направление искусственного интеллекта, основанное на моделировании коллективного поведения децентрализованных систем, состоящих из большого числа относительно простых агентов. Такие системы не имеют централизованного управления, однако за счет взаимодействия между агентами способны демонстрировать сложное и организованное поведение [6].

Такой принцип коллективного поиска решения лежит в основе муравьиного алгоритма. В алгоритме множество искусственных агентов (муравьев) параллельно строят возможные решения задачи, используя информацию о накопленном феромоне и локальные эвристики. Взаимодействие агентов осуществляется через общую структуру данных, в которой хранится уровень феромона на элементах решения. Благодаря этому возникает механизм положительной обратной связи: решения, приводящие к лучшим значениям целевой функции, усиливаются и становятся более вероятными для выбора в последующих итерациях.

Целью настоящей работы является усовершенствование формальной постановки задачи мультипроектного планирования с учетом ресурсных ограничений и разработка эффективного алгоритма ее решения в парадигме роевого интеллекта.

1. ОБЗОР СУЩЕСТВУЮЩИХ ПОДХОДОВ

В статье [7] рассматривается подход к решению задачи календарного планирования с ограниченными ресурсами (Resource-Constrained Project Scheduling Problem – RCPSP), в которой необходимо определить моменты начала выполнения работ таким образом, чтобы были соблюдены технологические зависимости между работами и ограничения на использование ресурсов, а общий срок завершения проекта был минимальным. Для решения этой задачи авторы предлагают использовать генетический алгоритм, адаптированный с учетом особенностей ресурсных ограничений.

Предложенный подход основан на представлении возможного расписания работ в виде хромосомы, которая кодирует порядок выполнения операций проекта. Такой порядок не является непосредственно готовым расписанием, а используется как приоритетный список, на основе которого строится допустимое расписание с учетом всех технологических и ресурсных ограничений. Для преобразования хромосомы в расписание применяется процедура построения графика, которая последовательно назначает моменты начала работ, учитывая их длительность, зависимости предшествования и доступность ресурсов. В результате из каждого генетического решения формируется корректное расписание проекта, для которого вычисляется значение целевой функции – длина расписания или время завершения всех работ.

Основным недостатком такого подхода является его вычислительная сложность, обусловленная необходимостью построения допустимого расписания

для каждой хромосомы в популяции на каждой итерации. Процедура декодирования, учитывающая ресурсные ограничения и технологические зависимости, требует значительных вычислительных ресурсов, что критично при увеличении размерности задачи. Кроме того, эффективность генетического алгоритма сильно зависит от выбора параметров эволюции (размер популяции, вероятность мутации и скрещивания), а также от способа кодирования решений, что требует предварительной настройки и может приводить к преждевременной сходимости к субоптимальным решениям.

В статье [8] представлен подход к решению NP-трудной задачи календарного планирования типа job shop, основанный на адаптивной модификации метода роя частиц. Суть подхода заключается в создании двухуровневой иерархической системы оптимизации, где один рой частиц решает исходную задачу, а другой рой динамически настраивает его параметры для повышения эффективности поиска.

Для применения метода роя частиц, который изначально разработан для непрерывных пространств, к дискретной задаче составления расписания авторы используют механизм декодирования. Координаты каждой частицы нижнего уровня интерпретируются как приоритеты операций, на основе которых строится допустимое расписание и вычисляется его качество.

Верхний уровень системы реализует адаптивную настройку параметров основного алгоритма. Рой частиц верхнего уровня подбирает такие значения коэффициентов, как инерция и ускорение, которые максимизируют эффективность поиска роя нижнего уровня. Оценкой для частиц верхнего уровня служит качество решений, найденных основным алгоритмом с соответствующим набором параметров.

Недостатком предложенной двухуровневой иерархической системы является ее высокая ресурсоемкость, поскольку она требует параллельной работы двух роев частиц, что удваивает вычислительные затраты по сравнению с классическим методом. Эффективность подхода также критически зависит от корректности механизма декодирования непрерывных координат частиц в дискретные приоритеты операций, так как неудачное отображение может привести к потере информации и снижению качества решений. Кроме того, несмотря на адаптивную настройку, алгоритм сохраняет чувствительность к начальным значениям параметров верхнего уровня, которые задаются в начале работы.

В статье [9] рассматривается метод, основанный на итеративном улучшении исходного допустимого расписания. На каждом шаге алгоритм генерирует небольшое возмущение текущего решения, например, меняя порядок задач или переназначая ресурсы, и оценивает полученный вариант. Если новое расписание оказывается лучше с точки зрения времени выполнения, оно принимается. В случае ухудшения решение также может быть принято, но с вероятностью, зависящей от значения ухудшения и текущего значения управляющего параметра, называемого температурой. На начальных этапах температура высокая, что позволяет алгоритму активно исследовать пространство решений, не застревая в локальных оптимумах. Постепенно температура снижается, и алгоритм переходит в режим локальной оптимизации, отдавая предпочтение только улучшающим изменениям. Такая динамика обеспечивает баланс между глобальным поиском и локальной сходимостью, и в результате после большого числа итераций удается получить расписание, близкое к оптимальному и превосходящее по качеству простые эвристики.

К достоинствам метода можно отнести способность преодолевать локальные минимумы за счет вероятностного принятия ухудшающих решений на

ранних стадиях, что выгодно отличает его от жадных алгоритмов. Кроме того, алгоритм хорошо масштабируется и допускает гибкую настройку вычислительных ресурсов через ограничение числа итераций и выбор закона изменения температуры.

Вместе с тем подход не гарантирует достижения глобального оптимума и чувствителен к выбору параметров. Особенно важна корректная настройка механизма генерации соседних решений: слишком малые изменения замедляют сходимость, а чрезмерно крупные приводят к неэффективному перебору и росту числа неудачных попыток. Это требует адаптации алгоритма под конкретную задачу, что может усложнить его практическое применение.

2. ПОСТАНОВКА ЗАДАЧИ

Имеется P проектов, содержащих в сумме I задач (операций), также имеется ограниченное число исполнителей E. Каждая задача относится к одному проекту и выполняется одним исполнителем. Для описания операций можно воспользоваться тройной индексацией (p, e, q), где p – номер проекта, e – номер исполнителя, q – кратность использования, т. е. операция (p, e, q) выполняется e-м исполнителем в p-м проекте в q-й раз.

Однако для упрощения записи перейдем от тройной индексации к одинарной: $(p,\ e,\ q) \Leftrightarrow (i)$. Для всех задач операции пронумерованы $i = \overline{1, I}$. Известно t_i^n – время выполнения в нормальном режиме, также известна технологическая последовательность выполнения задач в проектах. Задачи, которые выполняются одним исполнителем, не могут происходить одновременно.

Необходимо найти решение задачи с критерием минимального времени окончания всех проектов.

В таком случае формальная постановка задачи [10] может быть представлена следующим образом. Целевая функция имеет вид

$$T_k \to \min \tag{1}$$

при следующих ограничениях:

$$T_j + t_j^n \le T_k \quad \text{при} \quad j = \overline{1, I}, \tag{2}$$

$$T_i + t_i^n \le T_j, \text{ если операция } i \text{ предшествует } j,\ i, j = \overline{1, I}, \tag{3}$$

$$T_j \ge 0 \quad \text{при} \quad j = \overline{1, I}, \tag{4}$$

$$\begin{cases} T_j - T_i - Bw_{i,j} \le -t_i^n, \\ T_i - T_j + Bw_{i,j} \le B - t_j^n, \end{cases} \tag{5}$$

если операции i и j не могут выполняться одновременно, $w_{i,j}$ – переменная порядка выполнения пары операций (i, j):

$$w_{i,j} = \begin{cases} 1, \text{ если операция } j \text{ выполняется после } i, \\ 0 \text{ в противном случае,} \end{cases} \tag{6}$$

где T_j – время начала операции j; T_k – время завершения последнего из проектов; $B > 0$ – достаточно большое положительное число, используемое для взаимоисключающего выполнения операций.

Ограничение (3) для всех технологических зависимостей гарантирует, что если операция j не может начаться до завершения операции i согласно технологическому процессу проекта, то момент начала T_j будет не меньше, чем момент начала T_i, увеличенный на время выполнения t_i^n операции i. Ограничение (4) задает естественное условие неотрицательности моментов начала всех операций. Ограничение (5) связывает временные переменные с переменными порядка: если операция i предшествует операции j на одном ресурсе $(w_{i,j}=1)$, то ее завершение должно наступать раньше начала j; если же $w_{i,j}=0$, то большая константа B делает это неравенство автоматически выполненным, не накладывая ограничений.

Эту задачу можно показать посредством смешанного графа для небольшого примера мультипроектной разработки (рис. 1).

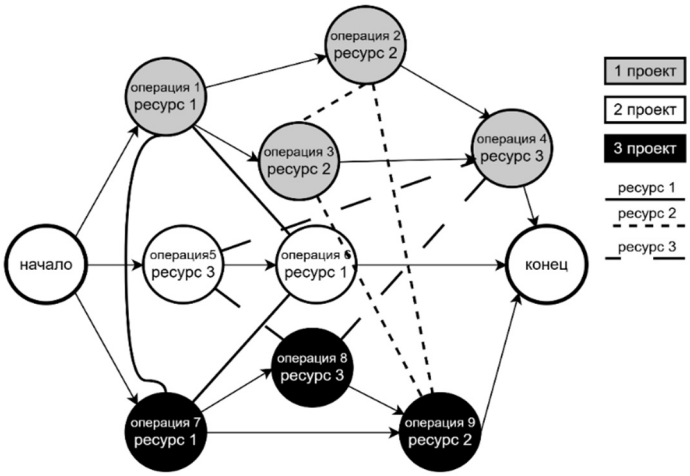

Рис. 1. Задача мультипроектных разработок в виде сетевого графа

Fig. 1. A multi-project development task in the form of a network graph

На сетевом графе вершины отображают операции, а технологические последовательности по проектам представлены дугами (связи со стрелками). Операции с одним ресурсом (исполнителем) соединены неориентированными ребрами, так как определенного порядка между ними в начале анализа нет. На рисунке операции с разными ресурсами соединены разными ребрами: ресурс 1 – жирной линией, ресурс 2 – мелким пунктиром, ресурс 3 – крупным пунктиром.

Используемая модель относится к классу смешанных, или дизъюнктивных, графов. Смешанный граф представляет собой граф, в котором одновременно присутствуют ориентированные дуги и неориентированные ребра. Ориентированные дуги задают жестко фиксированные технологические зависимости между операциями, определяющие допустимую последовательность их выполнения. Такие дуги отражают порядок выполнения работ внутри каждого проекта и формируют технологическую структуру задач. Неориентированные ребра, в свою очередь, отражают ресурсные ограничения и соединяют операции, которые используют один и тот же ресурс. Между такими операциями

на этапе построения модели порядок выполнения заранее не задан, поэтому соответствующие связи являются дизъюнктивными и должны быть ориентированы в процессе построения расписания.

В задачах теории расписаний смешанные графы позволяют одновременно учитывать технологические зависимости между операциями и ограничения по ресурсам. Формирование допустимого расписания в таком случае сводится к выбору направления для дизъюнктивных ребер, т. е. к определению порядка выполнения операций, конкурирующих за один и тот же ресурс. После ориентации этих ребер граф становится ориентированным ациклическим графом, по которому можно определить моменты начала выполнения операций и длительность всего расписания. Таким образом, использование смешанного графа позволяет наглядно представить структуру мультипроектной задачи и формализовать процесс поиска оптимального или близкого к оптимальному календарного плана выполнения работ.

3. АЛГОРИТМ МУРАВЬИНОЙ КОЛОНИИ

В природных условиях муравьи способны находить кратчайшие пути между муравейником и источником пищи, несмотря на ограниченные возможности отдельной особи. Обмен информацией между муравьями осуществляется косвенно – посредством выделения специального химического вещества, называемого феромоном. Передвигаясь по маршруту, муравьи оставляют феромон, который служит сигналом для других особей. Чем больше муравьев проходит по определенному пути, тем выше концентрация феромона и тем более привлекательным становится данный маршрут [11].

Феромон со временем испаряется, поэтому менее эффективные маршруты постепенно теряют привлекательность. В результате совместного действия процессов накопления и испарения феромона колония постепенно концентрирует движение на наиболее коротких и выгодных путях [12].

Популяция состоит из l искусственных муравьев, каждый из которых строит последовательность операций на каждом ресурсе заданное количество итераций N. На рис. 2 представлен пример решения конфликта на ресурсе 1. Ребра становятся дугами.

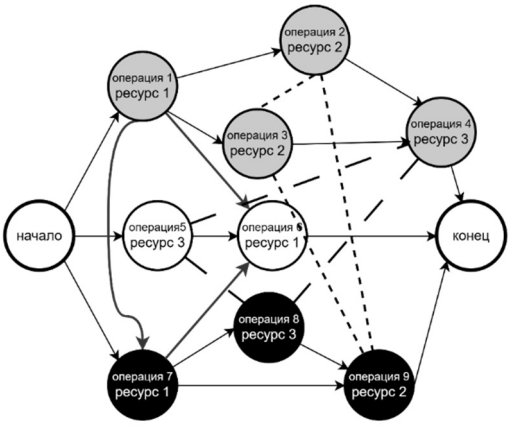

Рис. 2. Решение задачи на ресурсе 1

Fig. 2. Solution to the problem on resource 1

Можем заметить, что дуга «операция 1 – операция 6» является фиктивной, операция 6 всё равно будет осуществляться после операции 7, так как та идет после операции 1. В таком случае можно сделать вывод, что муравью достаточно построить прямую последовательность операций для каждого вида ресурса [1, 6, 7], подобно задаче коммивояжера.

Следовательно, у каждого муравья для каждого ресурса будет список с еще не посещенными вершинами. Выбор следующей вершины для посещения осуществляется по вероятностному правилу перехода:

$$p_{i,j} = \frac{\tau_{i,j}^{\alpha} d_{i,j}^{\beta}}{\sum\limits_{j \in V} \tau_{i,j}^{\alpha} d_{i,j}^{\beta}},$$

здесь $d_{i,j}$ – эвристическое расстояние между операциями i и j:

$$d_{i,j} = \frac{1}{T_j + 1},$$

где $p_{i,j}$ – вероятность перехода из вершины i в вершину j; V – вершины, доступные для посещения из вершины i (еще не посещенные операции, использующие тот же ресурс); $\tau_{i,j}$ – уровень феромона на ребре i, j; α – параметр, определяющий влияние феромона; β – параметр, определяющий влияние эвристической информации.

В рамках алгоритма ACO каждый искусственный муравей формирует допустимое расписание выполнения операций, последовательно выбирая операции из множества доступных к выполнению. Доступной считается операция, для которой выполнены все технологические предшествования и которая может быть назначена на соответствующий ресурс без нарушения ресурсных ограничений. При построении решения муравей выбирает следующую операцию i из множества допустимых операций A с вероятностью, зависящей от значения феромона $\tau_{i,j}$, связанного с переходом между операциями, и эвристической информации $d_{i,j}$, отражающей предпочтительность выбора операции, в нашем случае на основе времени начала операции. Можно эвристическую информацию выбирать на основе продолжительности операции, однако в случае, когда мы пересчитываем время после построения последовательности операций для каждого ресурса, предпочтительнее опираться на время начала, чтобы не было больших простоев. Вероятность выбора определяется стандартным соотношением алгоритма муравьиной колонии, в котором параметры α и β задают относительное влияние феромонной информации и эвристической оценки.

После того как муравей сформировал полную последовательность выполнения операций, на ее основе строится расписание с учетом технологических зависимостей и ограничений по ресурсам. Для каждой операции определяется момент начала выполнения как максимальное значение между моментом завершения всех ее технологических и ресурсных предшественников, затем расчет времени окончания всех проектов (самого позднего проекта) T_k. Если T_k меньше лучшего решения (или лучшего решения еще нет), то данная последовательность становится лучшим решением.

По завершении итерации алгоритма происходит обновление феромонов на дугах графа. Решения, приводящие к меньшей длительности общего расписания, усиливают соответствующие феромонные следы, что повышает вероятность выбора аналогичных последовательностей операций в последующих итерациях. Таким образом, алгоритм постепенно концентрирует поиск в области наиболее перспективных вариантов упорядочивания операций, что позволяет находить эффективные расписания выполнения мультипроектных разработок с учетом технологических и ресурсных ограничений. Сначала происходит процесс испарения феромонов:

$$\tau_{i,j} = \tau_{i,j}(1-\rho),$$

где ρ – коэффициент испарения феромона (от нуля до единицы включительно).

Происходит локальное добавление феромона на ребрах, входящих в найденные решения:

$$\tau_{i,j} = \tau_{i,j} + \Delta\tau_{i,j},$$

$$\Delta\tau_{i,j} = \begin{cases} \dfrac{Q}{T_k}, \text{ если } w_{i,j} = 1, \\ 0 \text{ в противном случае,} \end{cases}$$

где Q – константа, определяющая количество добавляемого феромона.

Значение феромона должно удовлетворять условию

$$\tau_{\min} \le \tau_{i,j} \le \tau_{\max},$$

где $\tau_{\min}$ – минимальный уровень феромона; $\tau_{\max}$ – максимальный уровень феромона.

Временную сложность можно рассчитать по формуле

$$TC = Nl\left(\sum_{e=1}^{E} n^2 + \varphi\right), \tag{7}$$

где TC – временная сложность (Time Complexity), N – количество итераций; l – количество муравьев; n – количество операций, использующих ресурс e; φ – время расчета временных характеристик.

Следовательно, временная сложность определяется совокупностью числа итераций и количества задействованных муравьев, а также затратами на обработку ресурсов в ходе вычислений. Полученное выражение позволяет выявить факторы, влияющие на производительность алгоритма и проанализировать характер роста вычислительной сложности.

4. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ПРИМЕНЕНИЯ МУРАВЬИНОГО АЛГОРИТМА В ОПТИМИЗАЦИИ МУЛЬТИПРОЕКТНОГО ПЛАНИРОВАНИЯ

Для начала изучим влияние параметров на работу алгоритма. Для этого была взята задача с 36 операциями, состоящая из шести проектов и шести ресурсов с оптимальным решением (временем окончания), равным 132. Коли-

чество муравьев на всех запусках программы будет фиксированным (l = 36), коэффициент испарения ρ = 0,01, а коэффициенты α и β будем менять, чтобы понять, при каких их значениях за сколько итераций достигается оптимум. Результаты представлены в табл. 1.

Таблица 1

Table 1

Количество итераций, необходимое для достижения оптимума

The number of iterations required to reach the optimum

β \ α	α = 0	α = 1	α = 2	α = 3	α = 4	α = 5	α = 6	α = 7	α = 8
β = 0	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
β = 1	>1000	98	163	102	82	72	>1000	>1000	>1000
β = 2	10	1	1	21	16	11	2	11	6
β = 3	3	5	2	6	2	2	9	1	4
β = 4	1	1	1	1	1	1	1	1	1
β = 5	1	1	1	1	1	1	1	1	1
β = 6	1	1	1	1	1	1	1	1	1
β = 7	1	1	1	1	1	1	1	1	1
β = 8	1	1	1	1	1	1	1	1	1

Проведенный эксперимент позволил оценить влияние параметров алгоритма муравьиной колонии на скорость нахождения оптимального решения. Полученные результаты показывают, что при $\beta = 0$ алгоритм не способен найти оптимальное решение даже за 1000 итераций независимо от значения параметра α. Это объясняется тем, что при $\beta = 0$ муравьи не учитывают эвристическую информацию в выборе следующего перехода и ориентируются только на уровень феромона, который на начальных этапах практически одинаков для всех путей. В результате поиск становится близким к случайному и эффективность алгоритма резко снижается.

При небольших значениях параметра β (например, $\beta = 1$) алгоритм начинает учитывать эвристическую информацию, однако скорость сходимости остается относительно низкой. В этом случае оптимальное решение достигается за десятки или сотни итераций, а при некоторых значениях параметра α оптимум не достигается даже за 1000 итераций. Это указывает на то, что влияние эвристической составляющей в данном случае недостаточно для эффективного направленного поиска.

С увеличением значения параметра β наблюдается значительное улучшение результатов. Уже при $\beta = 2$ оптимальное решение достигается за небольшое количество итераций, обычно от одной до нескольких десятков. Это свидетельствует о том, что усиление влияния эвристической информации позволяет алгоритму быстрее концентрировать поиск в перспективных областях пространства решений.

При $\beta \geq 4$ алгоритм демонстрирует наилучшие результаты: оптимальное решение находится уже на первой итерации практически при всех значениях

параметра α. Это говорит о том, что при достаточно высоком влиянии эвристической информации поиск становится более направленным и алгоритм быстро находит оптимальный маршрут.

Таким образом, результаты эксперимента показывают, что параметр β оказывает существенно большее влияние на скорость сходимости алгоритма, чем параметр α. Для рассматриваемой задачи наилучшие результаты достигаются при значениях β не менее 4, при которых алгоритм стабильно и быстро находит оптимальное решение независимо от выбранного значения параметра α. Полученные результаты подтверждают необходимость корректной настройки параметров алгоритма для обеспечения его высокой эффективности при решении задач календарного планирования.

Далее алгоритм тестировался на трех задачах по три раза (так как алгоритм содержит случайную компоненту) и сравнивался с результатами, полученными в IBM ILOG CPLEX Optimization Studio (программный комплекс для решения масштабных задач оптимизации) (табл. 2).

Название задач представлено в виде *I* – *P* – *E* (количество исполнителей), где *I* – количество операций, *P* – количество проектов, *E* – количество исполнителей. Точнее говоря, в задаче 25–5–5 всего 25 операций, 5 проектов и 5 ресурсов; в задаче 100–10–10 операций 100, проектов 10, ресурсов 10; задача 121–11–11 включает 121 операцию, 11 проектов и 11 ресурсов.

Проведенные эксперименты позволили оценить эффективность применения муравьиного алгоритма для решения задачи мультипроектного планирования и сравнить полученные результаты с решениями, найденными с использованием системы оптимизации IBM ILOG CPLEX Optimization Studio.

Таблица 2

Table 2

Результаты тестирования алгоритма

Algorithm testing results

Задача	CPLEX		Муравьиный алгоритм		
	T_k	Время решения	T_k	Отклонение от оптимума	Время решения
25–5–5	97	00:00:04	97	0 %	00:00:00
			97	0 %	00:00:00
			97	0 %	00:00:00
100–10–10	226	00:00:06	226	0 %	00:00:17
			234	3,5 %	00:00:10
			229	1,3 %	00:00:10
121–11–11	322	> 05:00:00	322	0 %	00:00:07
			322	0 %	00:00:02
			322	0 %	00:00:02

В качестве иллюстрации получаемых решений рассмотрим пример расписания, построенного алгоритмом для задачи небольшой размерности. Для задачи 25–5–5 алгоритм сформировал следующую последовательность выполнения операций (рис. 3).

```
=== РЕЗУЛЬТАТЫ ===
Время окончания всех проектов (Tk) = 97
Операция № 1: время начала (Ti) = 0, продолжительность = 12
Операция № 2: время начала (Ti) = 30, продолжительность = 9
Операция № 3: время начала (Ti) = 39, продолжительность = 8
Операция № 4: время начала (Ti) = 49, продолжительность = 22
Операция № 5: время начала (Ti) = 73, продолжительность = 7
Операция № 6: время начала (Ti) = 19, продолжительность = 11
Операция № 7: время начала (Ti) = 30, продолжительность = 12
Операция № 8: время начала (Ti) = 42, продолжительность = 21
Операция № 9: время начала (Ti) = 63, продолжительность = 24
Операция № 10: время начала (Ti) = 90, продолжительность = 5
Операция № 11: время начала (Ti) = 0, продолжительность = 19
Операция № 12: время начала (Ti) = 19, продолжительность = 23
Операция № 13: время начала (Ti) = 47, продолжительность = 13
Операция № 14: время начала (Ti) = 60, продолжительность = 5
Операция № 15: время начала (Ti) = 80, продолжительность = 10
Операция № 16: время начала (Ti) = 0, продолжительность = 15
Операция № 17: время начала (Ti) = 19, продолжительность = 8
Операция № 18: время начала (Ti) = 42, продолжительность = 7
Операция № 19: время начала (Ti) = 63, продолжительность = 17
Операция № 20: время начала (Ti) = 80, продолжительность = 8
Операция № 21: время начала (Ti) = 0, продолжительность = 19
Операция № 22: время начала (Ti) = 39, продолжительность = 22
Операция № 23: время начала (Ti) = 65, продолжительность = 8
Операция № 24: время начала (Ti) = 73, продолжительность = 17
Операция № 25: время начала (Ti) = 90, продолжительность = 7
Ресурс № 0; последовательность операций = 1, 8, 19, 15, 25
Ресурс № 4; последовательность операций = 11, 6, 2, 22, 20
Ресурс № 2; последовательность операций = 21, 17, 3, 13, 9
Ресурс № 1; последовательность операций = 12, 18, 4, 24, 10
Ресурс № 3; последовательность операций = 16, 7, 14, 23, 5
Время выполнения программы: 0,421 сек.
```

Рис. 3. Решение задачи 25 – 5 – 5

Fig. 3. Solution to problem 25 – 5 – 5

Решение этой же задачи в CPLEX:

// solution (optimal) with objective 97
// Quality Incumbent solution:
// MILP objective 9,7000000000e+01
// MILP solution norm |x| (Total, Max) 1,26900e+03 9,70000e+01

x = [0 30 39 51 80 19 30 42 66 90 0 19 47 75 80 0 31 42 63 80 0 39 61 73 90 97];
w = [0 0 0 0 0 0 0 0 1 0 0 1 1 0 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 0 1 1 0 0 1 0 0 0 0 1];

В обоих случаях значение целевой функции совпадает. Время завершения всех проектов составляет $T_k = 97$ (временных единиц). Это означает, что разработанный муравьиный алгоритм смог найти решение, совпадающее с оптимальным решением, полученным методом точной оптимизации в CPLEX. Следовательно, по критерию длительности расписания результаты алгоритма полностью соответствуют оптимальному значению.

При сравнении координат переменных, т. е. времени начала отдельных операций, можно заметить, что для части операций значения полностью совпадают. Например, операции 1, 2, 3, 6, 7, 8, 10, 11, 12, 13, 16, 18, 22, 24 и 25 имеют одинаковое время начала в обоих решениях. Для некоторых операций наблюдаются небольшие различия. Так, операция 4 в муравьином алгоритме начинается в момент времени 49, тогда как в CPLEX – в момент 51, операция 5 – 73 против 80, операция 9 – 63 против 66, операция 14 – 60 против 75,

операция 17 – 19 против 31, операция 19 – 49 против 63, операция 20 – 66 против 80, операция 23 – 65 против 61.

Такие различия объясняются тем, что задача допускает существование нескольких различных расписаний с одинаковым значением целевой функции. При этом суммарная длительность выполнения всех проектов остается неизменной. Муравьиный алгоритм формирует одно из допустимых расписаний, удовлетворяющих всем ограничениям по ресурсам и технологическим зависимостям, тогда как CPLEX находит одно из оптимальных решений задачи линейного целочисленного программирования.

Анализ результатов показывает, что для задачи небольшого размера (25–5–5) муравьиный алгоритм во всех запусках смог найти оптимальное решение, совпадающее с результатом CPLEX. При этом время работы алгоритма оказалось крайне малым и практически мгновенным, что свидетельствует о высокой эффективности метода для задач небольшой размерности.

Для более наглядной оценки посчитаем количество ограничений, непрерывных и булевых переменных. Так как операций 25, то непрерывных переменных, обозначающих время начала операций и окончания операций, в два раза больше:

$$S = 2I,$$

где S – количество непрерывных переменных.

Количество булевых переменных (6), которые необходимо найти, можно рассчитать по формуле

$$W = \sum_{e=1}^{E} C_n^2 = \sum_{e=1}^{E} \frac{n(n-1)}{2},$$

где W – количество переменных $w_{i,j}$ порядка выполнения для всех операций i, j, которые не могут выполняться одновременно; n – количество операций, использующих ресурс e.

Количество ограничений определяется по формуле

$$C = C_{\text{техн}} + C_{\text{неотр}} + C_{\text{рес}} + C_{\text{цел}}, \tag{8}$$

$$C_{\text{техн}} = I - P,$$

$$C_{\text{неотр}} = I,$$

$$C_{\text{рес}} = 2W,$$

$$C_{\text{цел}} = I,$$

где C – общее количество ограничений для задачи; $C_{\text{техн}}$ – количество технологических ограничений (3); $C_{\text{неотр}}$ – количество ограничений неотрицательности для моментов начала операций (5); $C_{\text{рес}}$ – количество ограничений, связанных с ресурсными ограничениями (6); $C_{\text{цел}}$ – количество ограничений, связанных с целевой функцией (2).

Подставив эти выражения в формулы (8), получим общую формулу количества ограничений в упрощенном виде:

$$C = 3I - \mathrm{P} + 2W.$$

Следовательно, для задачи 25–5–5 $S = 50,\ W = 10,\ C = 90.$

Для задачи большого размера (100–10–10) результаты оказались более вариативными. В одном из запусков муравьиный алгоритм также достиг оптимального решения, совпадающего с результатом CPLEX. В других запусках были получены решения с небольшим отклонением от оптимума – 3,5 % и 1,3 %. При этом время работы алгоритма составило от 10 до 17 секунд, что превышает время решения CPLEX, однако отклонение от оптимального значения остаётся относительно небольшим. Это подтверждает, что муравьиный алгоритм способен находить решения хорошего качества даже для более сложных задач, хотя из-за стохастической природы метода результаты отдельных запусков могут отличаться. Для данной задачи $S = 200,\ W = 450,\ C = 1190,$ что существенно больше, чем в задаче 25–5–5 и усложняет поиск оптимального решения.

Наиболее показательные результаты были получены для задачи еще большего размера (121–11–11). В данном случае точный метод, реализованный в CPLEX, не смог найти решение в течение более чем пяти часов вычислений, поэтому мы воспользовались CP, реализованный в модуле IBM ILOG CP Optimizer [13]. CP (Constraint Programming – программирование с ограничениями) представляет собой метод комбинаторной оптимизации, основанный на задании системы ограничений, применении специализированных процедур распространения ограничений и поиска по дереву решений. Такой подход широко применяется для решения задач календарного планирования и дискретной оптимизации. В результате использование CP позволило получить решение за 0,74 секунды с отклонением 0 % (оптимум). Муравьиный алгоритм во всех трех запусках также смог получить оптимальное значение целевой функции за несколько секунд. Это демонстрирует одно из ключевых преимуществ метаэвристических методов: способность эффективно находить хорошие решения для задач большой размерности, для которых точные методы требуют значительных вычислительных ресурсов или становятся практически неприменимыми. Для данной задачи $S = 242,\ W = 679,\ C = 1710.$ При таком количестве переменных и ограничений точные методы становятся практически неприменимы из-за комбинаторной сложности [14].

Таким образом, результаты экспериментов показывают, что муравьиный алгоритм является эффективным инструментом для решения задач мультипроектного планирования. Для задач малой размерности он способен находить оптимальные решения за минимальное время, а для более сложных задач обеспечивает получение решений, близких к оптимальным, при существенно меньших вычислительных затратах по сравнению с точными методами оптимизации.

В рамках проведенных вычислительных экспериментов была оценена вычислительная трудоемкость алгоритма муравьиной колонии. В задаче 121–11–11 использовалось $N = 10$ итераций и $l = 11$ муравьев при $I = 121$ операций. На каждой итерации каждый муравей последовательно формирует допустимое решение задачи, выбирая операции из множества доступных. На каждом шаге

выполняется вычисление временных характеристик расписания и проверка ограничений по ресурсам, что определяет трудоемкость одного шага алгоритма. Расчет временных характеристик φ после построения последовательности требует определения моментов начала всех операций с учетом технологических и ресурсных предшествований, что выполняется за $O(I^2)$ операций, так как для каждой операции необходимо проверить всех ее предшественников [15]. Так, трудоемкость одного шага для задачи 121–11–11 при равномерно распределенных операциях по ресурсам равна

$$T = \sum_{e=1}^{11} 11^2 + 121^2 = 1131 + 14\,641 = 15\,772 .$$

С учетом того, что в эксперименте использовалось $N = 10$ итераций и $l = 11$ муравьев, общая трудоемкость алгоритма для задачи размерности $I = 121$ составит

$$T_{\text{общ}} = Nl\left(\sum_{e=1}^{11} 11^2 + 121^2\right) = 10 \cdot 11 \cdot 15\,772 = 1\,734\,920 \text{ операций.}$$

Для оценки зависимости трудоемкости от размерности задачи проанализируем структуру полученной формулы (7). Слагаемое $\sum_{e=1}^{E} n^2$ характеризует затраты на упорядочивание операций в рамках отдельных ресурсов и в худшем случае не превышает I^2. Слагаемое φ, равное I^2, соответствует расчету временных характеристик расписания. Таким образом, оба слагаемых имеют порядок $O(I^2)$, и, следовательно, общая вычислительная сложность алгоритма составляет

$$TC = O(I^2).$$

Это означает, что алгоритм муравьиной колонии обладает полиномиальной временной сложностью: при увеличении числа операций в k раз время работы алгоритма увеличивается не более чем в k^2 раз.

Можно сделать вывод, что алгоритм муравьиной колонии обладает полиномиальной вычислительной трудоемкостью. Это означает, что время его работы увеличивается пропорционально росту основных параметров задачи, таких как число операций, количество итераций и число муравьев, участвующих в построении решений. Благодаря этому алгоритм демонстрирует достаточно высокое быстродействие и может эффективно применяться для любых реальных задач мультипроектного планирования с большим числом операций и ресурсных ограничений.

ЗАКЛЮЧЕНИЕ

В ходе работы была исследована возможность применения алгоритма муравьиной колонии для оптимизации распределения ресурсов в мультипроектной среде. Выполнена математическая постановка задачи с использованием

смешанных графов, учитывающих технологические зависимости и ресурсные ограничения. Разработанный алгоритм адаптирован для построения допустимых расписаний путем вероятностного выбора последовательности операций. Экспериментально подтверждена высокая эффективность подхода: на малых задачах достигается точное совпадение с оптимальными решениями CPLEX, на средних и больших – близкие к оптимуму результаты за приемлемое время, тогда как точные методы (и программные средства на их основе) становятся неприменимы. Полученные результаты позволяют рекомендовать алгоритм и разработанную его программную реализацию для практического использования в системах мультипроектного управления. Разработанный подход может быть применен в различных областях, где требуется планирование множества взаимосвязанных задач при ограниченных ресурсах: управление разработкой программного обеспечения, планирование производственных процессов, управление инженерными и строительными проектами, а также задачи распределения вычислительных ресурсов в информационных системах. Использование муравьиного алгоритма особенно целесообразно при большой размерности задачи, когда применение точных методов становится вычислительно затруднительным.

СПИСОК ЛИТЕРАТУРЫ

1. *Лазарев А.А., Гафаров Е.Р.* Теория расписаний. Задачи и алгоритмы. – М.: МГУ, 2011.

2. *Manne A.S.* On the job-shop scheduling problem // Operations Research. – 1960. – Vol. 8 (2). – P. 219–223.

3. *Олейникова С.А.* Математическая модель и оптимизационная задача составления расписания для мультипроектной системы с временными и ресурсными ограничениями и критерием равномерной загрузки // Вестник ВГТУ. – 2013. – № 6-3. – С. 58–61.

4. *Матвеев А.И.* Алгоритм оптимизации планирования ресурсов (на примере метода отжига) // Перспективные информационные технологии (ПИТ 2018). – Самара, 2018. – С. 1046–1050.

5. *Xiong H.* A survey of job shop scheduling problem: The types and models // Computers & Operations Research. – 2022. – Vol. 142. – P. 105731.

6. *Курейчик В.М., Кажаров А.А.* Использование роевого интеллекта в решении NP-трудных задач // Известия ЮФУ. Технические науки. – 2011. – № 7. – С. 30–36.

7. *Гончаров Е.Н., Леонов В.В.* Генетический алгоритм для задачи календарного планирования с ограниченными ресурсами // Автоматика и телемеханика. – 2017. – № 6. – С. 173–189.

8. *Блок И.Н., Секаев В.Г.* Разработка, исследование и адаптация метода роя частиц в задачах дискретной оптимизации теории расписаний // Глобальный научный потенциал. – 2015. – № 5. – С. 70–75.

9. *Chakraborty S., Bhowmik S.* Job shop scheduling using simulated annealing // First International Conference on Computation and Communication Advancement. – 2013. – Vol. 1 (1). – P. 69–73.

10. *Mezentsev Y.A., Chubko N.Y.* On one bicriterion discrete optimization problem and a hybrid ant colony algorithm for its approximate solution // Lecture Notes in Computer Science. – 2021. – Vol. 12689. – P. 289–300.

11. *Саймон Д.* Алгоритмы эволюционной оптимизации. – М.: ДМК Пресс, 2020. – 1002 с.

12. *Dorigo M., Stützle T.* Ant colony optimization: overview and recent advances // Handbook of Metaheuristics. – Springer, 2018. – P. 311–351.

13. IBM. Constraint programming in CP Optimizer. – URL: https://www.ibm.com/products/ilog-cplex-optimization-studio/cplex-cp-optimizer (accessed: 16.03.2026).

14. *Клеванский Н.Н., Красников А.А.* Формирование календарных графиков мультипроектного планирования // Образовательные ресурсы и технологии. – 2015. – № 3. – С. 11–33.

15. *Son P.V.H., Khoi L.N.Q., Loc L.X.* Multi-project scheduling optimization with artificial intelligence: a novel metaheuristic framework // Cluster Computing. – 2026. – Vol. 29. – P. 63.

Мезенцев Юрий Анатольевич, доктор технических наук, профессор кафедры автоматизированных систем управления Новосибирского государственного технического университета. Основные направления научных исследований – автоматика, вычислительная техника, компьютерные науки, математика, кибернетика. Имеет более ста печатных работ. E-mail: mesyan@yandex.ru

Борисова Алина Алексеевна, бакалавр кафедры автоматизированных систем управления Новосибирского государственного технического университета. Основное направление научных исследований – компьютерные науки. Имеет две печатные работы. E-mail: alina_10112004@mail.ru

Mezentsev Yurii A., PhD (Eng.), professor at the department of automated control systems, Novosibirsk State Technical University. The main areas of his scientific research are automation, computer engineering, computer science, mathematics, and cybernetics. He has more than a hundred printed works. E-mail: mesyan@yandex.ru

Borisova Alina A., bachelor at the department of automated control systems, Novosibirsk State Technical University. The main area of her scientific research is computer science. She has 2 printed works. E-mail: alina_10112004@mail.ru

DOI: 10.17212/2782-2001-2026-3-97-114

Application of the ant colony algorithm in optimizing the management of multi-project development[*]

Y. A. MEZENTSEV[a], *A. A. BORISOVA*[b]

Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation

[a] *mesyan@yandex.ru* [b] *alina_10112004@mail.ru*

Abstract

The article discusses the application of the Ant Colony Optimization (ACO) algorithm to solve the problem of optimizing the management of multi-project developments with limited resources. A multi-project environment is characterized by competition between tasks for resources, technological and resource dependencies, which makes the task combinatorial and NP-hard.

The paper describes an algorithm for an ant colony which includes the formation of acceptable schedules in parallel with several artificial ants, a probabilistic choice of operations taking into account the pheromone level and heuristic information, as well as updating pheromone traces during iterations. Experiments have been conducted on problems of various dimensions, which have shown that with the correct parameter setting, the algorithm consistently finds optimal or close to optimal solutions. For small tasks, the solution completely coincides with the results of precise optimization in IBM ILOG CPLEX, for medium and large tasks, the method provides efficient search for high-quality schedules in significantly less time compared to precise methods.

The research results demonstrate the polynomial computational complexity of the ACO algorithm and its high efficiency in planning multi-project developments, which makes it a promising tool for use in practical tasks of managing project resources and deadlines.

Keywords: multi-project planning, combinatorial optimization, ant colony algorithm, NP-hard problem, resource management, mixed graphs, calendar planning

REFERENCES

1. Lazarev A.A., Gafarov E.R. *Teoriya raspisanii. Zadachi i algoritmy*. Moscow, Moscow State University Publ., 2011.

2. Mann A.S. On the task of planning working hours in the workshop. *Operation Research*, 1960, vol. 8 (2), pp. 219–223.

[*] *Received 20 March 2026.*

3. Oleynikova S.A. Matematicheskaya model' i optimizatsionnaya zadacha sostavleniya raspisaniya dlya mul'tiproektnoi sistemy s vremennymi i resursnymi ogranicheniyami i kriteriem ravnomernoi zagruzki [Mathematical model and optimization problem of scheduling for multi-project system with time and resource constraints and criteria of uniform downloads]. *Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta = Bulletin of Voronezh State Technical University*, 2013, no. 6-3, pp. 58–61.

4. Matveev A.I. [Resource planning optimization algorithm]. *Perspektivnye informatsionnye tekhnologii (PIT 2018)* [Advanced information technologies and scientific computing (PIT 2018)]. Samara, 2018, pp. 1046–1050. (In Rssian).

5. Xiong H. An overview of the problem of workplace planning in workshops: types and models. *Computers & Operations Research*, 2022, vol. 142, p. 105731.

6. Kureichik V.M., Kazharov A.A. Ispol'zovanie roevogo intellekta v reshenii NP-trudnykh zadach [The use of new intelligence in the analysis of NP-labor resources]. *Izvestiya Yuzhnogo federal'nogo universiteta. Tekhnicheskie nauki = Izvestiya Southem Federal University. Engineering sciences*, 2011, no. 7, pp. 30–36.

7. Goncharov E.N., Leonov V.V. Geneticheskii algoritm dlya zadachi kalendarnogo planirovaniya s ogranichennymi resursami [Genetic algorithm for the resource-constrained project scheduling problem]. *Avtomatika i telemekhanika = Automation and Remote Control*, 2017, no. 6, pp. 173–189. (In Russian).

8. Blok I.N., Sekaev V.G. Razrabotka, issledovanie i adaptatsiya metoda roya chastits v zadachakh diskretnoi optimizatsii teorii raspisanii [Development, research and adaptation of the method of the particles' swarm in discrete optimization problems of scheduling theory]. *Global'nyi nauchnyi potentsial = Global Scientific Potential*, 2015, no. 5, pp. 70–75.

9. Chakraborty S., Boumik S. Workflow planning using simulated annealing. *The First International Conference on the Development of Computer Technology and Communications*, 2013, vol. 1 (1), pp. 69–73.

10. Mezentsev Y.A., Chubko N.Y. On a two-criterion discrete optimization problem and a hybrid ant colony algorithm for its approximate solution. *Lecture Notes in Computer Science*, 2021, vol. 12689, pp. 289–300.

11. Simon D. *Algoritmy evolyutsionnoi optimizatsii* [Algorithms of evolutionary optimization]. Moscow, DMK Press, 2020. 1002 p. (In Russian).

12. Dorigo M., Stützle T. Optimization of the ant colony: an overview and recent achievements. *Handbook of Metaheuristics.* Springer, 2018, pp. 311–351.

13. IBM. *Constraint programming in CP Optimizer.* Available at: https://www.ibm.com/products/ilog-cplex-optimization-studio/cplex-cp-optimizer (accessed 16.03.2026).

14. Klevansky N.N., Krasnikov A.A. Formirovanie kalendarnykh grafikov mul'tiproektnogo planirovaniya [Multi-project scheduling problems techniques]. *Obrazovatel'nye resursy i tekhnologii = Educational Resources and Technologies*, 2015, no. 3, pp. 11–33.

15. Son P.V.H., Khoi L.N.Q., Loc L.X. Multi-project scheduling optimization with artificial intelligence: a novel metaheuristic framework. *Cluster Computing*, 2026, vol. 29, p. 63.

Для цитирования:

Мезенцев Ю.А., Борисова А.А. Применение алгоритма муравьиной колонии в оптимизации управления мультипроектными разработками // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 97–114. – DOI: 10.17212/2782-2001-2026-3-97-114.

For citation:

Mezentsev Y.A., Borisova A.A. Primenenie algoritma murav'inoi kolonii v optimizatsii upravleniya mul'tiproektnymi razrabotkami [Application of the ant colony algorithm in optimizing the management of multi-project development]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 97–114. DOI: 10.17212/2782-2001-2026-3-97-114.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 115–130

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 115–130

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 303.732

DOI: 10.17212/2782-2001-2026-3-115-130

Анализ влияния гиперпараметров ансамблевых моделей на процесс их обучения в задаче прогнозирования электропотребления предприятия газовой промышленности*

А.И. СТЕПАНОВА[a], А.И. ХАЛЬЯСМАА[b], П.В. МАТРЕНИН[c]

РФ, 620062, г. Екатеринбург, ул. Мира, 19, Уральский федеральный университет имени первого Президента России Б.Н. Ельцина

[a] *a.i.stepanova@urfu.ru* [b]*a.i.khaliasmaa@urfu.ru* [c]*p.v.matrenin@urfu.ru*

В настоящей статье представлен анализ влияния гиперпараметров ансамблевых моделей на точность, стабильность и вычислительную эффективность в задаче краткосрочного прогнозирования потребления электрической энергии. Объектом исследования являлась реальная компрессорная станция, находящаяся на севере Центральной России. Эксперименты проводились для моделей, построенных на основании быстрого бустинга, экстремального градиентного бустинга, адаптивного бустинга и на основании случайного леса. В качестве гиперпараметров были выбраны эффективная глубина деревьев и количество базовых моделей. Общее количество комбинаций эффективной глубины деревьев и количества базовых деревьев составило 72 для каждой модели. Вклад работы заключается в проведении эксперимента по полному перебору комбинацией гиперпараметров на реальных данных предприятия газовой промышленности, оценки склонности моделей к переобучению и недообучению сравнением средней относительной ошибки по модулю на обучающей и тестовой выборках. Также были выявлены оптимальные конфигурации гиперпараметров для каждой модели с учетом компромисса между точностью, скоростью и стабильностью. По результатам экспериментов лучшей стабильностью обладают модели, построенные на основе случайного леса и адаптивного бустинга, но из-за высоких показателей времени обучения и низких показателей точности эти модели не могут быть применены для краткосрочного прогнозирования потребления предприятием газовой промышленности. Лучшими характеристиками точности и предсказуемости ошибки обладают модели, построенные на основе экстремального градиентного бустинга; хорошими показателями точности и самым низким временем обучения обладают модели, построенные на основе быстрого бустинга.

* *Статья получена 11 февраля 2026 г.*

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (тема № FEUZ-2025–0005).

Ключевые слова: ансамблевые методы, дерево решений, эффективная глубина дерева, количество базовых моделей, прогнозирование электропотребления, предприятие газовой промышленности, склонность к недообучению, склонность к переобучению, точность обучения, скорость обучения, стабильность обучения

ВВЕДЕНИЕ

Повышение энергетической эффективности и управление ресурсами являются ключевыми приоритетами предприятий газовой промышленности. Одним из перспективных мероприятий для анализа показателей энергетической эффективности является создание и внедрение системы краткосрочного прогнозирования потребления электрической энергии (1–3 суток). Внедрение такой системы позволяет:

- выровнять график потребления электрической энергии [1, 2];
- оптимизировать режимы работы объектов потребления, генерации и накопления электрической энергии [3, 4];
- минимизировать затраты на использование потребляемой мощности за счет регулирования тарифов на оптовом рынке электрической энергии и мощности путем планирования нагрузки [1, 5].

Внедрение усложнено из-за необходимости учета сложной структуры энергетического комплекса предприятий и сильной зависимости потребления электрической энергии от технологических процессов, климатических условий и режимов эксплуатации оборудования [5].

Для краткосрочного прогнозирования электропотребления применяются различные методы, которые можно разделить на детерминированные (статистические) методы и методы машинного обучения. К первым относятся сезонные модели [6] и методы на основе авторегрессии [7–9].

Методы машинного обучения способны учитывать большое количество факторов, в том числе метеорологических [10, 11] и производственных [12], а также зависимости между ними [13]. Преимущества применения методов машинного обучения по сравнению с детерминированными показаны в работах [14, 15]. Как правило, наилучшие результаты достигаются с помощью нейросетевых моделей [16–18], включая рекуррентные [19] и глубокие нейронные сети [17, 20, 21], ансамбли деревьев решений [22].

Однако, несмотря на широкое применение этих методов, точность и обобщающая способность методов зависят от выбора гиперпараметров. Большинство существующих исследований используют эвристический метод выбора параметров, в то время как работы [23–25] показывают, что даже небольшие изменения в гиперпараметрах могут приводить к колебаниям точности и повышению времени обучения в условиях ограниченности данных и высокого уровня шума [26].

В настоящее время отсутствуют комплексные исследования, посвященные эмпирическому анализу чувствительности ансамблевых моделей к гиперпараметрам для предприятий газовой промышленности.

Настоящая работа является продолжением исследования прогнозирования потребления электрической энергии предприятия газовой промышленности [27]. Целью исследования является анализ влияния ключевых гиперпараметров на точность, стабильность и вычислительную эффективность

моделей, построенных на адаптивном бустинге, экстремальном градиентном бустинге, быстром бустинге, случайном лесе (бустинг от англ. *boosting* – метод формирования ансамблей моделей машинного обучения, при котором несколько относительно простых моделей обучаются последовательно так, чтобы каждая следующая модель уменьшала ошибки предыдущих).

Вклад работы заключается в следующем.

1. Оценена склонность каждой модели к переобучению и недообучению сравнением метрик на обучающей и тестовой выборках.

2. Выявлены оптимальные конфигурации гиперпараметров для каждой модели с учетом компромисса между точностью, скоростью и стабильностью.

3. Сформулированы практические рекомендации по выбору модели и ее настройке для предприятий газовой промышленности.

В разделе 1 представлено описание используемых материалов и методов. В разделе 2 показаны результаты исследования влияния гиперпараметров на результаты обучения моделей. В заключении представлено описание будущих направлений исследования.

1. ИСПОЛЬЗУЕМЫЙ МЕТОД КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ

В работе используются архивные данные электропотребления компрессорной станции магистрального газопровода и метеорологические параметры, применяемые в качестве входных признаков моделей машинного обучения для краткосрочного прогнозирования потребления электрической энергии.

Объектом исследования в статье является компрессорная станция магистрального газопровода, которую можно определить как комплекс оборудования для повышения давления газа и его охлаждения при транспортировке. Рассматриваемая в статье станция расположена на севере Центральной России и состоит из трех компрессорных цехов.

Исходный набор данных включал трехлетние значения (791 день):

- почасового потребления электроэнергии компрессорной станцией для всего предприятия, полученные с автоматизированной системы коммерческого учета;
- суточного потребления электроэнергии аппаратами воздушного охлаждения (АВО) трех установок охлаждения газа;
- суточного потребления электроэнергии газоперекачивающими агрегатами (ГПА) трех компрессорных цехов.

Также в экспериментах рассматривались метеорологические данные, которые были взяты из архива сайта rp5.ru. Вместо прогнозных данных использовался архив фактических метеоданных из-за отсутствия открытого доступа к архивам прогнозных значений. Такое упрощение может быть оправдано тем, что метеорологические параметры являются дополнительными факторами, а точность прогноза погоды сопоставима с фактическими параметрами погоды на краткосрочном горизонте.

Описание предобработки данных представлено в статье [27]. Выбросы в наборе данных отсутствовали.

После корреляционного анализа, проведенного в статье [27], прогнозируемое потребление электрической энергии описывается как

$$y_i^* = f(y_{i-48}, y_{i-60}, y_{i-72}, y_{i-84}, y_{i-96}, X),$$

где y_i^* – прогнозируемое потребление электрической энергии в час i; f – прогнозная модель; X – другие значения (метеорологические данные, данные производственного календаря, суточное потребление ГПА, АВО).

Следует отметить, что наличие мультиколлинеарности входных признаков потенциально может влиять на устойчивость моделей. Однако для ансамблевых методов на основе деревьев решений данный эффект частично нивелируется за счет случайного выбора признаков и структуры разбиений.

Весь набор данных был разделен на обучающие и тестовые наборы данных в соотношении 90:10. Для повышения достоверности полученных результатов дополнительно использовалась процедура кросс-валидации на обучающей выборке (с использованием $k = 5$ подвыборок), в рамках которой модель обучалась на $k - 1$ подвыборках и тестировалась на оставшейся выборке. Итоговые значения метрик определялись как усредненные по всем разбиениям.

Для решения задачи краткосрочного прогнозирования потребления электрической энергии в работе выбраны ансамблевые методы машинного обучения: случайный лес (СЛ), адаптивный бустинг (АБ) [28], экстремальный градиентный бустинг (ГБ) [29], быстрый бустинг (ББ) [30].

Для обеспечения сопоставимости и контроля сложности моделей в задаче оценки влияния параметров моделей на точность и скорость ее работы в настоящей работе были выбраны эффективная глубина дерева и количество базовых деревьев.

Эффективная глубина дерева определяет максимальную сложность базового решающего дерева. Для адаптивного бустинга, экстремального градиентного бустинга и случайных деревьев использовался гиперпараметр max_depth. Так как в быстром бустинге рост дерева контролируется числом листьев, для него использовался параметр num_leaves.

Количество базовых моделей определяет число деревьев в ансамбле и выражается n_estimators. Этот параметр является общим для всех рассматриваемых моделей.

Выбор параметров обусловлен влиянием на компромисс смещения и дисперсии (bias-variance tradeoff), а также влиянием на вычислительную сложность модели.

Установлены диапазоны значений гиперпараметров:

- max_depth (num_leaves): 3, 5, 10, 20, 30, 50, 70, 100;
- n_estimators: от 10, 30, 50, 70, 90, 100, 150, 200, 300.

При значении max_depth (num_leaves) до 10 анализируются слабые модели. При значении глубины больше 50 анализируются сложные модели, что позволяет выявить режимы недообучения и переобучения.

Минимальные значения n_estimators позволяют покрыть минимальные конфигурации для оценки нижней границы точности и насыщенные ансамбли для оценки предела улучшения модели.

Общее количество комбинаций – 72 для каждой модели.

Выбор дискретных значений гиперпараметров с относительно крупным шагом обусловлен необходимостью ограничения вычислительной сложности полного перебора при сохранении репрезентативности исследуемого пространства параметров.

Для сравнения результатов и оценки склонности каждой модели к переобучению и недообучению были рассчитаны время обучения, средняя абсолютная ошибка (MAE), средняя относительная ошибка по модулю (MAPE), коэффициент детерминации (R2).

Предобработка данных, построение и тестирование моделей машинного обучения было выполнено на языке программирования Python 3 с применением библиотек с открытым исходным кодом: Scikit-Learn (методы адаптивного бустинга, случайного леса) [28], XGBoost (метод экстремального градиентного бустинга) [29], LightGBM (метод быстрого бустинга) [30].

Эксперименты проводились на вычислительной системе со следующими характеристиками: центральный процессор Intel Core i5 G7, 4 ядра 2,4 ГГц, 16 ГБ оперативной памяти.

2. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

На рис. 1–4 представлены поверхности, которые показывают зависимость времени обучения модели от эффективной глубины и количества базовых моделей для экстремального градиентного бустинга, адаптивного бустинга, быстрого бустинга и случайного леса.

Поверхности строятся для наглядного анализа вычислительной эффективности, формы поверхности.

Наибольшей скоростью обучения обладают модели, построенные на основе быстрого бустинга. Это обосновано следующими особенностями метода:

- лист-ориентированный рост дерева;
- использование гистограммной бинаризации признаков;
- фильтрация малозначимых признаков.

Неровности формы поверхности быстрого бустинга обоснованы оптимизацией, зависящей от структуры данных, и динамическим выбором признаков.

Волнистость поверхности быстрого бустинга выражена слабо, что говорит о высокой эффективности оптимизации.

Время обучения моделей, построенных на основе адаптивного бустинга, линейно зависит от количества базовых моделей на всей поверхности. Зависимость от эффективной глубины дерева имеет линейный характер до max_depth = 20, после этого значения время обучения не изменяются при возрастании глубины дерева. Адаптивный бустинг обучается медленнее быстрого бустинга, так как деревья обучаются последовательно, пересчитывая веса объектов на каждом шаге без использования оптимизации.

Форма поверхности адаптивного бустинга отражает детерминированную сложность (сложность модели полностью определяется параметрами).

Время обучения моделей, построенных на основе случайного леса, линейно зависит от количества базовых моделей.

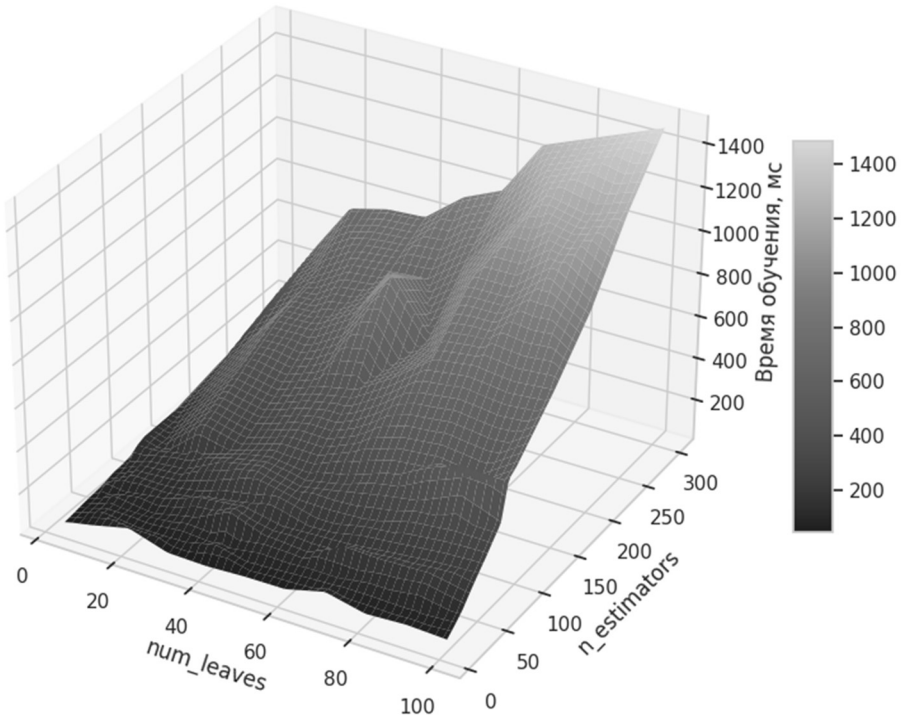

Рис. 1. Зависимость времени обучения от гиперпараметров для ББ

Fig. 1. Training time as a function of hyperparameters for LightGBM

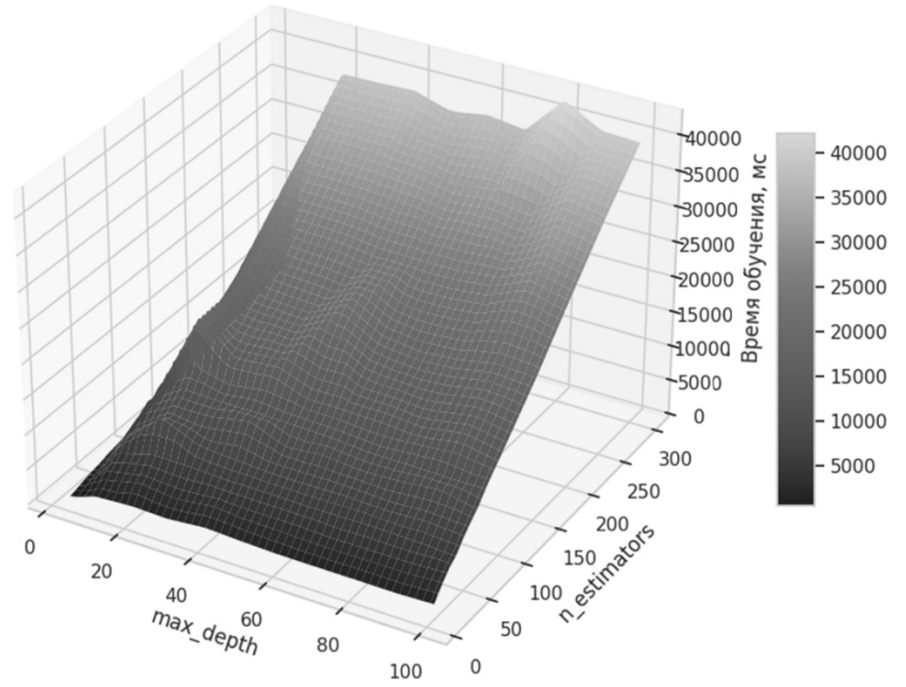

Рис. 2. Зависимость времени обучения от гиперпараметров для АБ

Fig. 2. Training time as a function of hyperparameters for AdaBoost

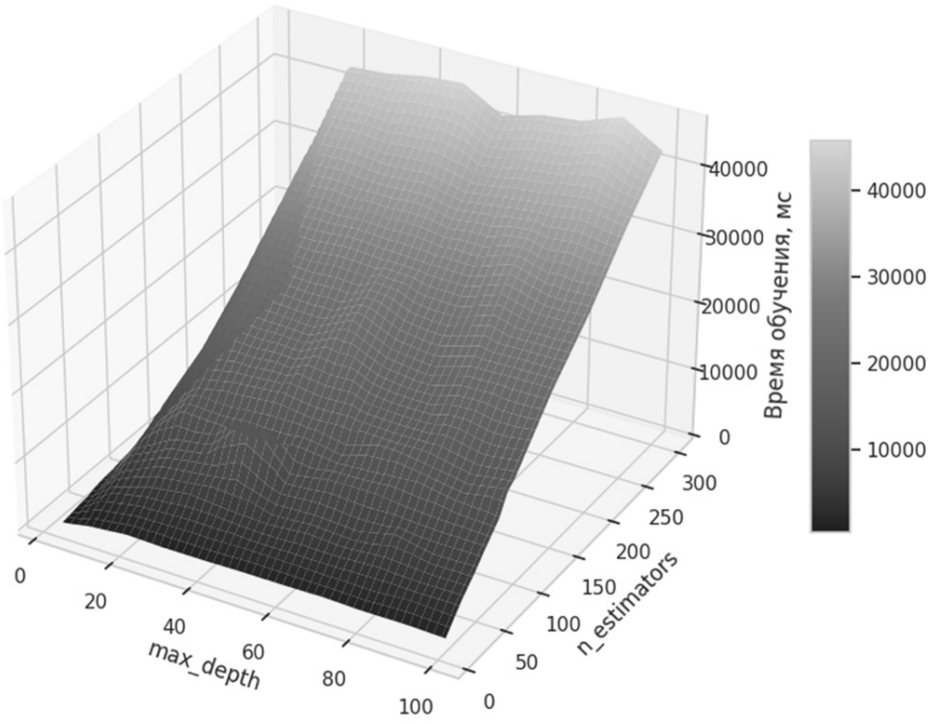

Рис. 3. Зависимость времени обучения от гиперпараметров для СЛ

Fig. 3. Training time as a function of hyperparameters for Random Forest

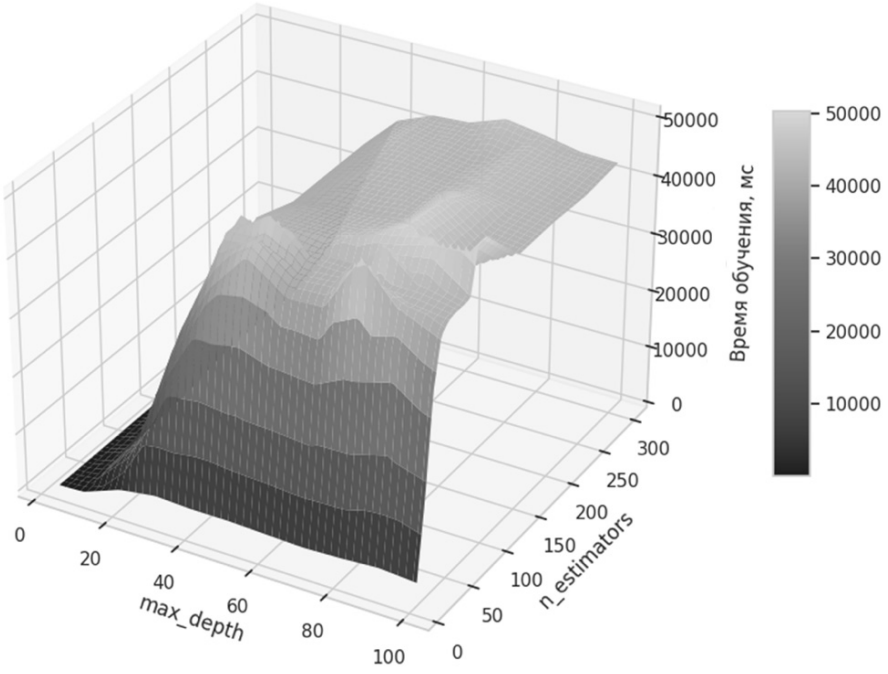

Рис. 4. Зависимость времени обучения от гиперпараметров для ГБ

Fig. 4. Training time as a function of hyperparameters for XGBoost

Зависимость от эффективной глубины дерева имеет линейный характер до max_depth = 20, после этого значения времени обучения не изменяются. Большее время обучения по сравнению с быстрым бустингом обосновано тем, что каждое дерево строится на полной выборке без оптимизации.

Гладкая поверхность случайного леса показывает стабильность работы модели при изменении параметров.

Время обучения моделей, построенных на основе экстремального градиентного бустинга, линейно зависит от эффективной глубины дерева до max_depth = 40, после этого не изменяется. При этом значения времени обучения для эффективной глубины больше max_depth = 40 совпадают. Время обучения модели экстремального градиентного бустинга определяется сочетанием глубины деревьев и количества базовых моделей. При значениях параметра max_depth < 30 увеличение количества базовых моделей практически не влияет на время обучения. При большей глубине деревьев (max_depth ≥ 30) и количестве базовых моделей до 100 время обучения возрастает линейно с увеличением их числа. При дальнейшем увеличении количества базовых моделей существенного изменения времени обучения не наблюдается. Время обучения экстремального градиентного бустинга имеет наибольшие значения, так как он использует точный поиск разбиений и вычисляет вторые производные функции потерь. Это обеспечивает более высокую точность, но сказывается на вычислительной сложности.

Неровность поверхности обоснована оптимизацией, зависящей от структуры данных, и кэшированием градиентов.

Сильная волнистость поверхности экстремального градиентного бустинга отображает высокую чувствительность к структуре данных при max_depth >30, что приводит к уменьшению предсказуемости времени обучения. Высокая чувствительность к гиперпараметрам может увеличить операционные риски при автоматическом развертывании.

Гладкость поверхностей адаптивного бустинга и случайного леса обеспечивает предсказуемость, но ценой низкой скорости. На рис. 5 представлен график зависимости MAPE от количества базовых моделей для анализа склонности моделей к переобучению. Сплошные линии соответствуют значению MAPE на тестовой выборке, пунктирные – на обучающей. Для повышения читаемости рисунка представлены зависимости только для четырех значений эффективной глубины дерева.

При анализе графика зависимости MAPE от количества базовых моделей для адаптивного бустинга можно заключить, что для всех значений max_depth наблюдается переобучение. С ростом количества базовых моделей ошибка тестовой выборки растет, что указывает на неустойчивость модели при увеличении сложности ансамбля.

График зависимости MAPE от количества базовых моделей для экстремального градиентного бустинга показывает наилучшую устойчивость. Так как разница MAPE между тестовой и обучающей выборками минимальная, можно заключить, что у экстремального градиентного бустинга низкая склонность к переобучению. Также видно, что MAPE незначительно зависит от эффективной глубины выборки.

Для графика зависимости MAPE от количества базовых моделей быстрого бустинга при значениях количества базовых моделей до 50 наблюдается резкое

падение MAPE, что говорит о быстрой адаптации модели. При количестве базовых моделей больше 100 MAPE выходит на плато, дальнейшее увеличение числа деревьев не улучшает точность. Зависимость от max_depth выражена слабо.

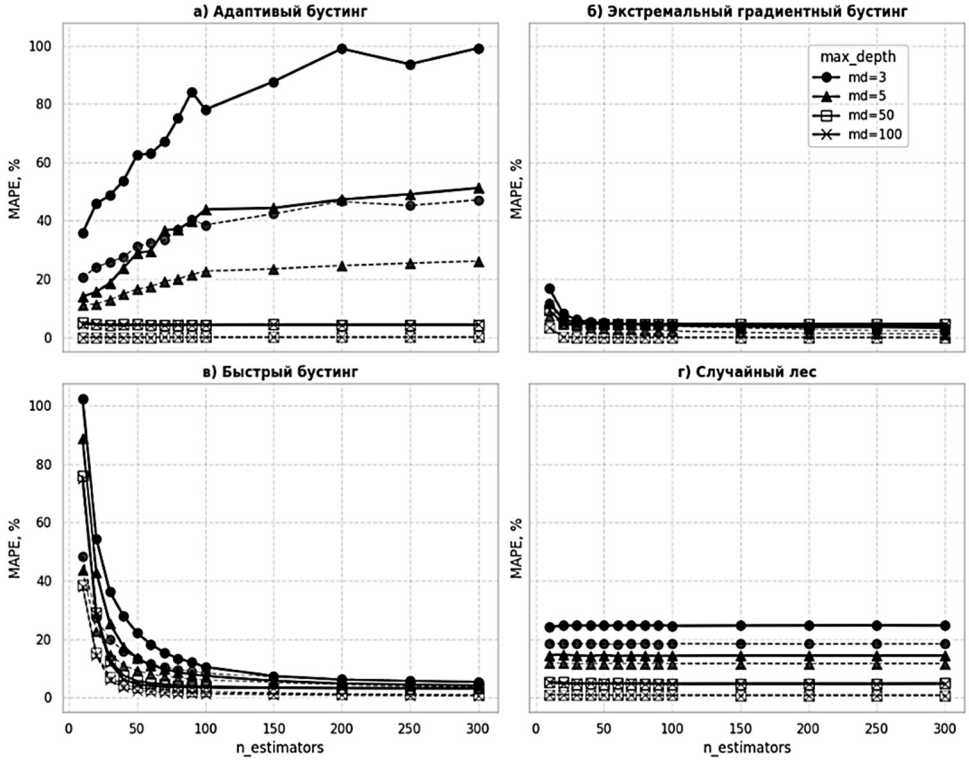

Рис. 5. Зависимость MAPE от количества базовых моделей для адаптивного бустинга (*а*), экстремального градиентного бустинга (*б*), быстрого бустинга (*в*), случайного леса (*г*)

Fig. 5. MAPE as a function of the total number of trees for AdaBoost (*a*), XGBoost (*b*), LightGBM (*c*), Random Forest (*d*)

Для графика зависимости MAPE от количества базовых моделей случайного леса MAPE находится в диапазоне 5…25 % в зависимости от эффективной глубины дерева. При эффективной глубине дерева 3 модель недообучена. Параметры точности моделей, построенных с использованием моделей случайного леса, стабилизируются при max_depth = 20. При значениях эффективной глубины дерева ошибка снижается и стабилизируется. Это значение является точкой оптимального компромисса между точностью и скоростью.

Сравнение лучших результатов обучения по точности и скорости для всех моделей представлено в таблице. Выбор оптимальной модели осуществлялся на основе многокритериального подхода, учитывающего одновременно точность, устойчивость (разница между обучающей и тестовой ошибкой) и вычислительную эффективность (время обучения). В рамках исследования применялся принцип Парето-оптимальности, при котором предпочтение отдавалось моделям, обеспечивающим наилучший компромисс между указанными критериями.

Модель, построенная на основе экстремального градиентного бустинга, показывает лучший результат: MAPE на тестовой выборке составляет 3,24 %, R2 составляет 0,98. Несмотря на то что использование такой модели нецелесообразно при частом повторном обучении из-за высокого и непредсказуемого времени обучения, так как в задаче краткосрочного прогнозирования потребления электрической энергии предприятия газовой промышленности нет необходимости часто обучать модель на новых данных, модель на основе экстремального градиентного бустинга может быть выбрана для применения на объекте.

Сравнение лучших результатов обучения

Comparison of the best training results

Модель	max_ depth	n_ estima- tors	MAE обуч, кВт	MAE тест, кВт	MAPE тест, %	R2	Время обучения, мс
ББ	5	300	23,11	13,40	4,24	0,96	324
ГБ	3	300	12,73	9,20	3,24	0,98	359
АД	10	100	13,79	21,92	7,75	0,91	9909
СЛ	30	100	5,73	16,02	4,78	0,93	15665

Также хорошими показателями обладает модель, построенная с использованием быстрого бустинга: MAPE на тестовой составляет 4,24 %, R2 составляет 0,96. При внедрении модели и выборе между моделями на основе экстремального градиентного бустинга и быстрого бустинга необходимо учитывать, что модели на основе быстрого бустинга демонстрируют высокую стабильность при правильных параметрах max_depth, что делает их оптимальными для промышленного внедрения при частом переобучении модели.

Случайный лес является наиболее предсказуемым по производительности, однако уступает в точности моделям на основе быстрого бустинга и экстремального градиентного бустинга. Лучшие показатели: MAPE на тестовой выборке составляет 4,78 % при времени обучения 15 665 мс.

Адаптивный бустинг склонен к переобучению. Для лучшего показателя время обучения модели меньше, чем у случайного леса, но склонность к переобучению и более выраженная по сравнению со случайным лесом нестабильность делает бустинг непрактичным для применения для рассматриваемой задачи.

С практической точки зрения снижение ошибки прогнозирования на уровне 1–2 процентных пункта MAPE является значимым для энергетических объектов, поскольку позволяет более точно планировать график потребления, снижать штрафы за отклонение от заявленных режимов и оптимизировать

загрузку оборудования. Для компрессорных станций это выражается в снижении эксплуатационных затрат и повышении энергоэффективности за счет более точного управления режимами работы.

ЗАКЛЮЧЕНИЕ

В настоящей статье для задачи краткосрочного прогнозирования потребления электрической энергии на основе данных потребления электрической энергии компрессорной станцией был проведен эксперимент по полному перебору комбинаций гиперпараметров для выбора модели с целью промышленного внедрения.

На основе анализа форм поверхностей времени обучения наиболее стабильными являются модели случайного леса и адаптивного бустинга, однако они демонстрируют худшие показатели сходимости MAPE по сравнению с методами экстремального градиентного бустинга и быстрого бустинга. Модели, построенные на основе быстрого бустинга, имеют высокие способности к адаптации и высокую скорость обучения. Предлагаемой к внедрению моделью является модель, построенная на экстремальном градиентном бустинге, так как она обеспечивает минимальную ошибку (3,24 %). Однако при необходимости частого переобучения или изменения постановки задачи для прогнозирования потребления в режиме реального времени необходимо рассмотреть применение быстрого бустинга.

При интерпретации полученных результатов необходимо учитывать ряд ограничений: качество прогнозирования существенно зависит от полноты и достоверности исходных данных, включая корректность измерений и отсутствие пропусков. Результаты получены на данных одного промышленного объекта.

В ходе дальнейших исследований планируется провести дообучение моделей, построенных на основе экстремального градиентного бустинга и быстрого бустинга, данными других объектов газовой промышленности с последующим анализом поведения моделей.

СПИСОК ЛИТЕРАТУРЫ

1. *Lee E., Kim J., Jang D.* Load profile segmentation for effective residential demand response program: Method and evidence from Korean pilot study // Energies. – 2020. – Vol. 13 (6). – P. 1348. – DOI: 10.3390/en13061348.

2. A short-term forecasting approach for regional electricity power consumption by considering its co-movement with economic indices / K. Li, Z. Yang, D. Li, Y.Y. Xing, W. Nai // IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC). – IEEE, 2020. – P. 551–555. – DOI: 10.1109/ITOEC49072.2020.9141928.

3. Renewable on-site power generation for manufacturing companies – Technologies, modeling, and dimensioning / J. Schulz, D. Leinmüller, A. Misik, M.F. Zaeh // Sustainability. – 2021. – Vol. 13 (7). – P. 3898. – DOI: 10.3390/su13073898.

4. Optimal maintenance policy for equipment submitted to multi-period leasing as a circular business model / A.B. Mabrouk, A. Chelbi, M.S. Aguir, S. Dellagi // Sustainability. – 2024. – Vol. 16 (12). – P. 5238. – DOI: 10.3390/su16125238.

5. *Stepanova A.I., Matrenin P.V.* New load forecasting ensemble model based on LightGBM for gas industry enterprises // Proceedings of the 2024 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT). – IEEE, 2024. – P. 1–4. – DOI: 10.1109/USBEREIT61901.2024.10583981.

6. Power load forecasting research based on neural network and Holt-winters method / J. Fan, X. Liu, Z. Li, X. Wang, S. Cao, J. Lei // IOP Conference Series: Earth and Environmental Science. – 2021. – Vol. 692. – P. 022120. – DOI: 10.1088/1755-1315/692/2/022120.

7. Short-term forecast of electricity load for LLC “Omsk Energy Retail Company” using neural network / V. Potapov, R. Khamitov, V. Makarov, A. Gritsay, I. Chervenchuk, D. Tyunkov // Proceeding of 2018 Dynamics of Systems, Mechanisms and Machines (Dynamics). – IEEE, 2018. – P. 1–5. – DOI: 10.1109/Dynamics.2018.8601430.

8. *Chodakowska E., Nazarko J., Nazarko L.* ARIMA models in electrical load forecasting and their robustness to noise // Energies. – 2021. – Vol. 14 (23). – P. 7952. – DOI: 10.3390/en14237952.

9. Blind Kalman filtering for short-term load forecasting / S. Sharma, A. Majumdar, V. Elvira, E. Chouzenoux // IEEE Transactions on Power Systems. – 2020. – Vol. 35. – P. 4916–4919. – DOI: 10.1109/TPWRS.2020.3018623.

10. *Caro E., Juan J., Nouhitehrani S.* Optimal selection of weather stations for electric load forecasting // IEEE Access. – 2023. – Vol. 11. – P. 42981–42990. – DOI: 10.1109/ACCESS.2023.3270933.

11. Regression model-based short-term load forecasting for university campus load / M. Madhukumar, A. Sebastian, X. Liang, M. Jamil, M.N.S.K. Shabbir // IEEE Access. – 2022. – Vol. 10. – P. 8891–8905. – DOI: 10.1109/ACCESS.2022.3144206.

12. *Sergeev N., Matrenin P.* Improving accuracy of machine learning based short-term load forecasting models with correlation analysis and feature engineering // Proceeding of 2023 IEEE 24th International Conference of Young Professionals in Electron Devices and Materials (EDM). – IEEE, 2023. – P. 1000–1004. – DOI: 10.1109/EDM58354.2023.10225058.

13. Load forecasting techniques and their applications in smart grids / H. Habbak, M. Mahmoud, K. Metwally, M.M. Fouda, M.I. Ibrahem // Energies. – 2023. – Vol. 16 (3). – P. 1480. – DOI: 10.3390/en16031480.

14. *Ryu S., Noh J., Kim H*. Deep neural network based demand side short term load forecasting // Energies. – 2017. – Vol. 10 (1). – P. 3. – DOI: 10.3390/en10010003.

15. *Szczepaniuk H., Szczepaniuk E.K*. Applications of artificial intelligence algorithms in the energy sector // Energies. – 2023. – Vol. 16 (1). – P. 347. – DOI: 10.3390/en16010347.

16. A short-term load forecasting method based on GRU-CNN hybrid neural network model / L. Wu, C. Kong, X. Hao, W. Chen // Mathematical Problems in Engineering. – 2020. – P. 1428104. – DOI: 10.1155/2020/1428104.

17. Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting / Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, Z. Zhu // IEEE Access. – 2019. – Vol. 7. – P. 88058–88071. – DOI: 10.1109/ACCESS.2019.2926137.

18. Short-term residential load forecasting based on LSTM recurrent neural network / W. Kong, Z.Y. Dong, Y. Jia, D.J. Hill, Y. Xu, Y. Zhang // IEEE Transactions on Smart Grid. – 2019. – Vol. 10. – P. 841–851. – DOI: 10.1109/TSG.2017.2753802.

19. *Cai M., Pipattanasomporn M., Rahman S*. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques // Applied Energy. – 2019. – Vol. 236 (3). – P. 1078–1088. – DOI: 10.1016/j.apenergy.2018.12.042.

20. A deep learning method for short-term residential load forecasting in smart grid / Y. Hong, Y. Zhou, Q. Li, W. Xu, X. Zheng // IEEE Access. – 2020. – Vol. 8. – P. 55785–55797. – DOI: 10.1109/ACCESS.2020.2981817.

21. Load forecasting techniques for power system: Research challenges and survey / N. Ahmad, Y. Ghadi, M. Adnan, M. Ali // IEEE Access. – 2022. – Vol. 10. – P. 71054–71090. – DOI: 10.1109/ACCESS.2022.3187839.

22. Forecasting planned electricity consumption for the united power system using machine learning / R.V. Klyuev, A.D. Morgoeva, O.A. Gavrina, I.I. Bosikov, I.D. Morgoev // Journal of Mining Institute. – 2023. – Vol. 261. – P. 392–402.

23. *Mohammadagha M*. Hyperparameter optimization strategies for tree-based machine learning models prediction: A comparative study of AdaBoost, decision trees, and random forest // SSRN Electronic Journal. – 2025. – DOI: 10.2139/ssrn.5226457.

24. *Dasgupta S., Sen J*. A comparative study of hyperparameter tuning methods // Data Science in Theory and Practice. – 2024. – DOI: 10.48550/arXiv.2408.16425.

25. A unified framework for tuning hyperparameters in clustering problems / X. Fan, Y. Yue, P. Sarkar, Y.X.R. Wang // arXiv. – 2019. – DOI: 10.48550/arXiv.1910.08018.

26. *Hidayaturrohman Q.A., Hanada E.* A comparative analysis of hyper-parameter optimization methods for predicting heart failure outcomes // Applied Science. – 2025. – Vol. 15 (6). – P. 3393. – DOI: 10.3390/app15063393.

27. Application of SHAP and multi-agent approach for short-term forecast of power consumption of gas industry enterprises / A.I. Stepanova, A.I. Khalyasmaa, P.V. Matrenin, S.A. Eroshenko // Algorithms. – 2024. – Vol. 17 (10). – P. 447. – DOI: 10.3390/a17100447.

28. Scikit-Learn libraries. – URL: https://scikit-learn.org/stable/supervised_learning.html#supervised-learning (accessed: 17.01.2026).

29. XGBoost. Introduction to Boosted Trees. – URL: https://xgboost.readthedocs.io/en/stable/tutorials/model.html (accessed: 17.01.2026).

30. LightGBM. – URL: https://lightgbm.readthedocs.io/en/stable/ (accessed: 17.01.2026).

Степанова Алина Игоревна, аспирант, младший научный сотрудник научной лаборатории цифровых двойников в электроэнергетике Уральского энергетического института, Уральский федеральный университет имени первого Президента России Б.Н. Ельцина. Основное направление научных исследований – методы искусственного интеллекта в электроэнергетике, объяснимый искусственный интеллект, анализ временных рядов. Имеет более 10 печатных работ. E-mail: a.i.stepanova@urfu.ru

Хальясмаа Александра Ильмаровна, кандидат технических наук, доцент, заведующий научной лабораторией цифровых двойников в электроэнергетике Уральского энергетического института, Уральский федеральный университет имени первого Президента России Б.Н. Ельцина. Основное направление научных исследований – техническая диагностика электрооборудования, методы искусственного интеллекта в электроэнергетике, робототехника. Имеет более 250 печатных работ. E-mail: a.i.khaliasmaa@urfu.ru

Матренин Павел Викторович, кандидат технических наук, доцент, ведущий научный сотрудник научной лаборатории цифровых двойников в электроэнергетике Уральского энергетического института, Уральский федеральный университет имени первого Президента России Б.Н. Ельцина. Основное направление научных исследований – методы искусственного интеллекта в электроэнергетике, объяснимый искусственный интеллект, анализ временных рядов. Имеет более 250 печатных работ. E-mail: p.v.matrenin@urfu.ru

Stepanova Alina I., PhD student, research assistant at the scientific laboratory of digital twins in the power industry, Ural Power Engineering Institute, Ural Federal University named after the first President of Russia, B.N. Yeltsin. The main field of her scientific research is methods of artificial intelligence in power industry, eXplainable artificial intelligence, and time series analysis. She is the author of more than 10 scientific publications. E-mail: a.i.stepanova@urfu.ru

Khalyasmaa Alexandra I., Ph.D, Head of scientific laboratory of digital twins in the power industry, Ural Power Engineering Institute, Ural Federal University named after the first President of Russia, B.N. Yeltsin. The main field of her scientific research is technical diagnostics of electrical equipment, methods of artificial intelligence in power industry, and robotics. She is the author of more than 250 scientific publications. E-mail: a.i.khaliasmaa@urfu.ru

Matrenin Pavel V., Ph.D, associate professor, leading researcher at the scientific laboratory of digital twins in the power industry, Ural Power Engineering Institute, Ural Federal University named after the first President of Russia, B.N. Yeltsin. The main field of his scientific research is methods of artificial intelligence in power industry, eXplainable artificial intelligence, and time series analysis. He is the author of more than 250 scientific publications. E-mail: p.v.matrenin@urfu.ru

DOI: 10.17212/2782-2001-2026-3-115-130

*Analysis of the impact of hyperparameters of ensemble models on their training process in the task of predicting electricity consumption for a gas industry enterprise**

A.I. STEPANOVA [a], A.I. KHALYASMAA [b], P.V. MATRENIN [c]

Ural Federal University, named after the first President of Russia, B.N. Yeltsin, 19, Mira Street, Ekaterinburg, 620062, Russian Federation

[a] *a.i.stepanova@urfu.ru* [b] *a.i.khaliasmaa@urfu.ru* [c] *p.v.matrenin@urfu.ru*

Abstract

This paper presents an analysis of the impact of hyperparameters on the accuracy, stability, and computational efficiency of ensemble models in short-term electrical energy consumption forecasting. The study was conducted on a real compressor station located in northern Central Russia. Experiments were performed using models based on LightGBM, XGBoost, AdaBoost, and Random Forest. The effective tree depth and the number of base learners were selected as hyperparameters, yielding 72 combinations of these parameters for each model. The contribution of this work lies in conducting a full hyperparameter grid search on real data from a gas industry enterprise assessing model tendencies toward overfitting and underfitting by comparing a mean absolute percentage error on training and test sets, and identifying optimal hyperparameter configurations for each model that balance accuracy, speed, and stability. Experimental results show that Random Forest and AdaBoost models demonstrate the highest stability; however, due to long training times and lower accuracy, these models are unsuitable for short-term forecasting in the gas industry applications. XGBoost-based models achieve the best accuracy and predictable error behavior, while LightGBM-based models provide competitive accuracy with the shortest training time.

Keywords: ensemble methods, decision tree, effective tree depth, number of base learners, electricity consumption forecasting, gas industry enterprise, underfitting tendency, overfitting tendency, prediction accuracy, training speed, training stability

REFERENCES

1. Lee E., Kim J., Jang D. load profile segmentation for effective residential demand response program: Method and evidence from Korean pilot study. *Energies*, 2020, vol. 13 (6), p. 1348. DOI: 10.3390/en13061348.

2. Li K., Yang Z., Li D., Xing Y.Y., Nai W. A short-term forecasting approach for regional electricity power consumption by considering its co-movement with economic indices. *IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*. IEEE, 2020, pp. 551–555. DOI: 10.1109/ITOEC49072.2020.9141928.

3. Schulz J., Leinmüller D., Misik A., Zaeh M.F. Renewable on-site power generation for manufacturing companies – Technologies, modeling, and dimensioning. *Sustainability*, 2021, vol. 13 (7), p. 3898. DOI: 10.3390/su13073898.

4. Mabrouk A.B., Chelbi A., Aguir M.S., Dellagi S. Optimal maintenance policy for equipment submitted to multi-period leasing as a circular business model. *Sustainability*, 2024, vol. 16 (12), p. 5238. DOI: 10.3390/su16125238.

5. Stepanova A.I., Matrenin P.V. New load forecasting ensemble model based on LightGBM for gas industry enterprises. *Proceedings of the 2024 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*. IEEE, 2024, pp. 1–4. DOI: 10.1109/USBEREIT61901.2024.10583981.

* *Received 11 February 2026.*

The work was carried out within the framework of a state assignment with financial support from the Ministry of Science and Higher Education of the Russian Federation (topic No. FEUZ-2025–0005).

6. Fan J., Liu X., Li Z., Wang X., Cao S., Lei J. Power load forecasting research based on neural network and Holt-winters method. *IOP Conference Series: Earth and Environmental Science*, 2021, vol. 692, p. 022120. DOI: 10.1088/1755-1315/692/2/022120.

7. Potapov V., Khamitov R., Makarov V., Gritsay A., Chervenchuk I., Tyunkov D. Short-term forecast of electricity load for LLC "Omsk Energy Retail Company" using neural network. *Proceeding of 2018 Dynamics of Systems, Mechanisms and Machines (Dynamics)*. IEEE, 2018, pp. 1–5. DOI: 10.1109/Dynamics.2018.8601430.

8. Chodakowska E., Nazarko J., Nazarko L. ARIMA models in electrical load forecasting and their robustness to noise. *Energies*, 2021, vol. 14 (23), p. 7952. DOI: 10.3390/en14237952.

9. Sharma S., Majumdar A., Elvira V., Chouzenoux E. Blind Kalman filtering for short-term load forecasting. *IEEE Transactions on Power Systems*, 2020, vol. 35, pp. 4916–4919. DOI: 10.1109/TPWRS.2020.3018623.

10. Caro E., Juan J., Nouhitehrani S. Optimal selection of weather stations for electric load forecasting. *IEEE Access*, 2023, no. 11, pp. 42981–42990. DOI: 10.1109/ACCESS.2023.3270933.

11. Madhukumar M., Sebastian A., Liang X., Jamil M., Shabbir M.N.S.K. Regression model-based short-term load forecasting for university campus load. *IEEE Access*, 2022, no. 10, pp. 8891–8905. DOI: 10.1109/ACCESS.2022.3144206.

12. Sergeev N., Matrenin P. Improving accuracy of machine learning based short-term load forecasting models with correlation analysis and feature engineering. *Proceeding of 2023 IEEE 24th International Conference of Young Professionals in Electron Devices and Materials (EDM)*, 2023, pp. 1000–1004. DOI: 10.1109/EDM58354.2023.10225058.

13. Habbak H., Mahmoud M., Metwally K., Fouda M.M., Ibrahem M.I. Load forecasting techniques and their applications in smart grids. *Energies*, 2023, vol. 16 (3), p. 1480. DOI: 10.3390/en16031480.

14. Ryu S., Noh J., Kim H. Deep neural network based demand side short term load forecasting. *Energies*, 2017, vol. 10 (1), p. 3. DOI: 10.3390/en10010003.

15. Szczepaniuk H., Szczepaniuk E.K. Applications of artificial intelligence algorithms in the energy sector. *Energies*, 2023, vol. 16 (1), p. 347. DOI: 10.3390/en16010347.

16. Wu L., Kong C., Hao X., Chen W. A short-term load forecasting method based on gru-cnn hybrid neural network model. *Mathematical Problems in Engineering*, 2020, p. 1428104. DOI: 10.1155/2020/1428104.

17. Deng Z., Wang B., Xu Y., Xu T., Liu C., Zhu Z. Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting. *IEEE Access*, 2019, vol. 7, pp. 88058–88071. DOI: 10.1109/ACCESS.2019.2926137.

18. Kong W., Dong Z., Jia Y., Hill D.H., Xu Y., Zhang Y. Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, 2019, vol. 10, pp. 841–851. DOI: 10.1109/TSG.2017.2753802.

19. Cai M., Pipattanasomporn M., Rahman S. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. *Applied Energy*, 2019, vol. 236 (3), pp. 1078–1088. DOI: 10.1016/j.apenergy.2018.12.042.

20. Hong Y., Zhou Y., Li Q., Xu Q., Zhou Y., Li Q. A deep learning method for short-term residential load forecasting in smart grid. *IEEE Access*, 2020, vol. 8, pp. 55785–55797. DOI: 10.1109/ACCESS.2020.2981817.

21. Ahmad N., Ghadi Y., Adnan M., Ali M. Load forecasting techniques for power system: Research challenges and survey. *IEEE Access*, 2022, vol. 10, pp. 71054–71090. DOI: 10.1109/ACCESS.2022.3187839.

22. Klyuev R.V., Morgoeva A.D., Gavrina O.A., Bosikov I.I., Morgoev I.D. Forecasting planned electricity consumption for the united power system using machine learning. *Journal of Mining Institute*, 2023, vol. 261, pp. 392–402.

23. Mohammadagha M. Hyperparameter optimization strategies for tree-based machine learning models prediction: A comparative study of AdaBoost, decision trees, and random forest. *SSRN Electronic Journal*, 2025. DOI: 10.2139/ssrn.5226457.

24. Dasgupta S., Sen J. A comparative study of hyperparameter tuning methods. *Data Science in Theory and Practice*, 2024. DOI: 10.48550/arXiv.2408.16425.

25. Xinjie F., Yuguang Y., Sarkar P., Wang Y.X.R. A unified framework for tuning hyperparameters in clustering problems. *arXiv*, 2019. DOI: 10.48550/arXiv.1910.08018.

26. Hidayaturrohman Q.A., Hanada E. A comparative analysis of hyper-parameter optimization methods for predicting heart failure outcomes. *Applied Science*, 2025, vol. 15 (6), p. 3393. DOI: 10.3390/app15063393.

27. Stepanova A.I., Khalyasmaa A.I., Matrenin P.V., Eroshenko S.A. Application of SHAP and multi-agent approach for short-term forecast of power consumption of gas industry enterprises. *Algorithms*, 2024, vol. 17 (10), p. 447. DOI: 10.3390/a17100447.

28. *Scikit-Learn libraries*. Available at: https://scikit-learn.org/stable/supervised_learning.html#supervised-learning (accessed 17.01.2026).

29. XGBoost. *Introduction to Boosted Trees*. Available at: https://xgboost.readthedocs.io/en/stable/tutorials/model.html (accessed 17.01.2026).

30. LightGBM. Available at: https://lightgbm.readthedocs.io/en/stable/ (accessed 17.01.2026).

Для цитирования:

Степанова А.И., Хальясмаа А.И., Матренин П.В. Анализ влияния гиперпараметров ансамблевых моделей на процесс их обучения в задаче прогнозирования электропотребления предприятия газовой промышленности // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 115–130. – DOI: 10.17212/2782-2001-2026-3-115-130.

For citation:

Stepanova A.I., Khalyasmaa A.I., Matrenin P.V. Analiz vliyaniya giperparametrov ansamblevykh modelei na protsess ikh obucheniya v zadache prognozirovaniya elektropotrebleniya predpriyatiya gazovoi promyshlennosti [Analysis of the impact of hyperparameters of ensemble models on their training process in the task of predicting electricity consumption for a gas industry enterprise]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 115–130. DOI: 10.17212/2782-2001-2026-3-115-130.

ISSN 2782-2001
Системы анализа
и обработки данных
том 103, № 3, 2026, с. 131–150

http://journals.nstu.ru/vestnik
Analysis and data
processing systems
Vol. 103, No. 3, 2026, pp. 131–150

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 65.011.56

DOI: 10.17212/2782-2001-2026-3-131-150

Достоверность измерений размеров канатной проволоки с использованием цифровых технологий и нейронных сетей*

В.С. ХОРЕВ[a], С.Г. ЦАПКО[b], Е.В. КУЗЬМИНСКАЯ[c], П.А. КОМАНДИНОВ[d]

634050, г. Томск, пр. Ленина, 30, Национальный исследовательский Томский политехнический университет

[a] *horevvs@tpu.ru* [b] *tsapko@tpu.ru* [c] *bedareva@tpu.ru* [d] *pak37@tpu.ru*

Актуальность исследования обусловлена необходимостью повышения достоверности и производительности контроля геометрических параметров канатной проволоки, поскольку традиционные контактные методы не позволяют проводить измерения в реальном времени на движущихся объектах, а визуальный контроль субъективен и не обеспечивает количественной оценки. Переход к бесконтактным оптическим системам в сочетании с интеллектуальной обработкой данных открывает возможность автоматизированного неразрушающего контроля с требуемой точностью и повторяемостью результатов.

Целью работы является разработка и метрологическое обоснование метода измерения диаметра и выявления поверхностных дефектов канатной проволоки на основе машинного зрения и нейросетевых алгоритмов. Для достижения поставленной цели решались задачи калибровки оптической системы с применением рабочих эталонов, создания алгоритмов сегментации и анализа изображений, а также классификации дефектов с использованием сверточной нейронной сети.

Предложен и реализован алгоритм обработки изображений, включающий адаптивную бинаризацию по методу Оцу, преобразование расстояний для определения диаметра, контурный анализ для локализации дефектов и их последующую классификацию нейросетевой моделью. Экспериментальная апробация на проволоке диаметром 1,80 мм с наработкой от 500 до 3500 ч подтвердила работоспособность подхода: погрешность измерений диаметра составила ±0,022 мм (менее 1,5 %), что обеспечивает дискретность результатов до 0,001 мм. Разработанный программный комплекс на языке Python позволяет осуществлять автоматизированный контроль в режиме, близком к реальному времени, и может быть рекомендован для внедрения в системы неразрушающего контроля качества.

Ключевые слова: канатная проволока, компьютерное зрение, бинарное изображение, видеофиксация, визуализация, фильтрация, зажатие, задир, перегиб, нейронная сеть, дефект, эталон, Python

* *Статья получена 07апреля 2026 г.*

ВВЕДЕНИЕ

Обеспечение достоверности измерений геометрических параметров канатной проволоки представляет собой комплексную метрологическую задачу, решение которой в современных условиях базируется на интеграции прецизионных оптических систем и методов интеллектуальной обработки данных на основе нейронных сетей. Традиционные контактные методы, включая использование рычажных скоб и цифровых микрометров, обеспечивают высокую точность в лабораторных условиях, однако их применение ограничено низкой производительностью и невозможностью контроля движущихся объектов в реальном времени. Поэтому доминирующее положение занимают бесконтактные оптические методы, среди которых промышленное распространение получили лазерные сканирующие микрометры и системы теневой проекции. Например, измерительная система TLAser 103 производства компании LaserLink Inc., основанная на сканирующем лазерном луче, демонстрирует возможность выполнения 1200 измерений в секунду с повторяемостью ±0,025 мкм и позиционной погрешностью ±0,25 мкм, что позволяет эффективно контролировать проволоку диаметром от 15 мкм, движущуюся с высокой скоростью, даже в неблагоприятных условиях промышленного производства [1, 4].

Лазерные дифракционные измерители представляют собой альтернативный подход, особенно эффективный для очень тонкой проволоки. Как показано в работе Ходиера, опубликованной в журнале Optics & Laser Technology, анализ комбинированной интерференционной и дифракционной картины, возникающей при облучении вертикальной металлической проволоки лазерным лучом, позволяет измерять диаметр в диапазоне 170…450 мкм с расширенной неопределенностью от ±1,45 мкм для тонкой проволоки до ±0,57 мкм для более толстой. Исследование подчеркивает, что метод является высокочувствительным, но требует контролируемых условий окружающей среды и отсутствия вибраций, а рассеяние света может влиять на точность измерений для проволок малого диаметра, что потребовало применения поляроидных пленок для подавления рассеянного излучения [2]. Развитие этого направления нашло отражение в европейском проекте ABSODIAM, целью которого стало создание прототипа абсолютной измерительной системы для цилиндрических объектов диаметром от 0,005 до 0,2 мм на основе анализа дифракционной картины с усовершенствованным аппаратным обеспечением и уточненной дифракционной формулой для обработки сигнала, а также разработка эталонных материалов для данного диапазона диаметров [8]. Коммерческие реализации дифракционного метода, такие как лазерный микрометр TRUBACH Microline L360, обеспечивают измерение диаметра сверхтонких проволок от 0,005 до 0,3 мм с точностью ±(0,03 + 0,03 % × D) мкм, при этом отсутствие движущихся частей гарантирует стабильность измерений в течение длительного срока эксплуатации [7]. Лазерные дифракционные системы также применяются для онлайн-измерения диаметра нитей в диапазоне 5…160 мкм с абсолютной точностью около 1 % от диаметра и чувствительностью к изменениям диаметра около 0,2 % [5].

Дальнейшее повышение достоверности измерений связывают с внедрением методов компьютерного зрения и нейросетевой обработки сигналов. Исследователи Национального института стандартов и технологий США (NIST)

разработали лазерный микрометр четвертого поколения, использующий усовершенствованный лазерный интерферометр для измерения диаметров объектов шириной менее 50 мм с неопределенностью всего 2 нм. Прибор, изготовленный практически полностью из инвара для обеспечения термической стабильности, позволяет варьировать силу контакта и экстраполировать недеформированный диаметр, что ставит его на уровень лучших мировых аналогов при меньшей стоимости и простоте эксплуатации [8]. Принципиально иной подход к обеспечению достоверности предлагают методы трехмерной реконструкции на основе глубокого обучения. В работе Харьюхахто с соавторами, представленной на международной конференции Jicable и доступной в архиве препринтов arXiv, описан разработанный метод топографического сканирования для мониторинга процесса изоляции силовых кабелей. Современные измерительные технологии в сочетании с высокопроизводительными встроенными вычислениями позволяют строить полные детализированные 3D-карты поверхности изолированной жилы, а для обнаружения поверхностных дефектов применяются методы глубокого обучения, при этом сверточные нейронные сети хорошо подходят для анализа данных измерений поверхности в реальном времени [3].

Критическим аспектом применения нейронных сетей в измерительных задачах является оценка достоверности их выходных данных. Исследование Мишани и Синтова, опубликованное в журнале Engineering Applications of Artificial Intelligence, посвящено изучению возможности оценки формы провода на основе измерений силы и крутящего момента без использования визуального восприятия. Авторы предлагают обучать супервизируемый автоэнкодер со сверточными слоями, который реконструирует пространственную форму провода, одновременно принуждая латентное пространство соответствовать пространству измерений силы и крутящего момента. Декодер модели, обученной на нитиноловых и стандартных электрических проводах, позволяет точно оценивать форму по данным силомоментных датчиков, а энкодер решает обратную задачу отображения желаемой формы в требуемое состояние силы и крутящего момента, что демонстрирует способность нейросетей учитывать неопределенности, такие как нелинейность датчиков, начальная кривизна и неоднородность провода [6]. Таким образом, современное состояние исследований в области обеспечения достоверности измерений размеров проволоки характеризуется переходом от автономных контактных средств к гибридным цифровым системам, интегрирующим высокоскоростные оптические сенсоры с алгоритмами глубокого обучения и методами количественной оценки неопределенности, что позволяет достигать субмикронной точности в широком диапазоне диаметров и условий эксплуатации.

Если продолжить обсуждение этой темы, то стоит особо отметить: точность и надежность результатов напрямую зависит от того, насколько грамотно организовано метрологическое обеспечение измерений. Сегодняшние методы уже умеют работать с погрешностями, которые связаны с калибровкой системы и нестабильностью условий эксперимента. Помимо этого, специалисты могут выделять и те составляющие ошибок, что возникают на этапах обработки снимков и распознавания дефектов.

Чтобы система работала надежнее, авторы настоящей статьи предлагают использовать рабочие эталоны с заранее известными размерами. Такой подход позволил свести к минимуму инструментальные погрешности и заметно повысить эффективность измерительного тракта в системе машинного

зрения. В ходе исследований было разработано несколько ключевых решений. Например, подход, при котором эталонные образцы калибруют измерительную систему прямо во время ее работы. Также был создан алгоритм, который корректирует оптические искажения и автоматически выделяет границы проволоки на цифровых изображениях. Инженеры ТПУ сконструировали устройство, которое совмещает эталон и исследуемый образец в одном оптическом поле. И наконец, участники исследовательской работы, результатам которой посвящена статья, написали программу для статистической обработки данных и нейросетевого анализа дефектов поверхности. Причем разработка ПО была выполнена с поправкой на метрологические характеристики самого тракта.

В итоге внедрение этих предложенных подходов и наработок позволяет резко снизить неопределенность измерений, а контроль геометрических параметров проволоки становится куда более достоверным [3].

1. КОНЦЕПЦИЯ ИЗМЕРЕНИЯ РАЗМЕРОВ

В процессе контроля канатной проволоки могут быть обнаружены различные типы поверхностных дефектов, примеры которых приведены на рис. 1 и рис. 2, где показаны имитации прижатия коренной задвижкой и локальной эллипсности на проволоке диаметром 1,8 мм.

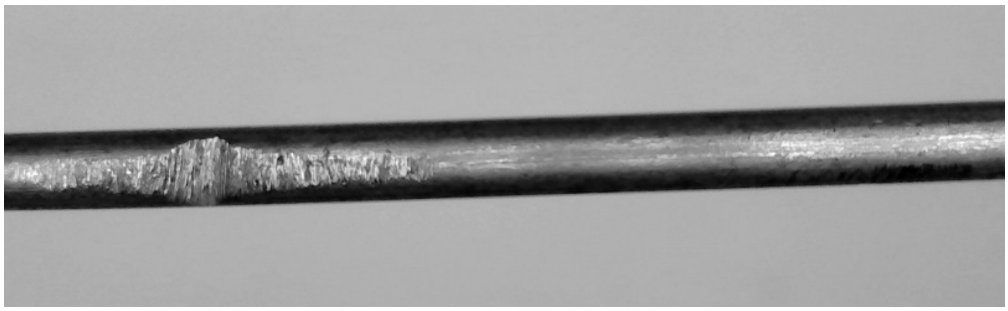

Рис. 1. Пример дефекта «вмятина» на канатной проволоке

Fig. 1. An example of a “dent” defect on the rope wire

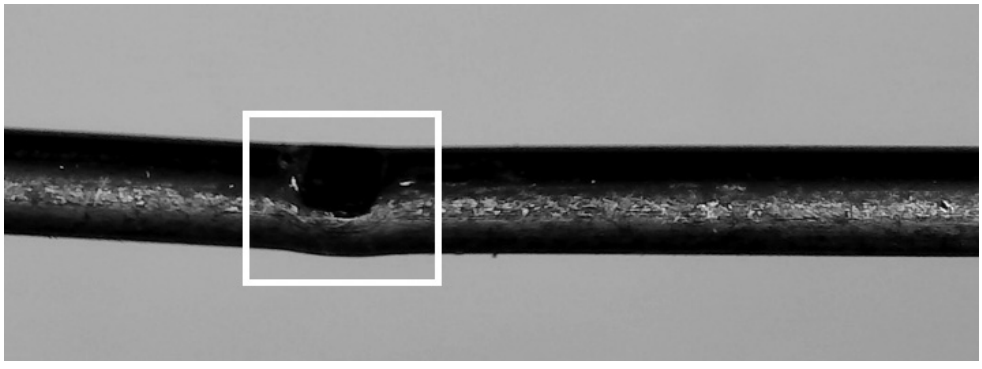

Рис. 2. Пример дефекта «поджатие» на канатной проволоке

Fig. 2. An example of a “compression” defect on the rope wire

Для количественной оценки геометрических параметров и выявления подобных неоднородностей был разработан специализированный алгоритм обработки изображений, позволяющий выполнять измерение диаметра объекта контроля с последующей детекцией дефектов. Перед проведением измерений необходимо подготовить исходное изображение. Для этого применяется метод бинаризации, преобразующий цветное изображение в черно-белое, где фон отображается черным цветом, а проволока – белым, что обеспечивает более четкое выделение границ. Бинарное изображение анализируется с использованием пороговой сегментации, при которой извлекается контур наибольшего объекта, соответствующего проволоке, после чего формируется маска и подготавливаются данные для дальнейшей обработки. Для сегментации проволоки от фона используется адаптивная пороговая бинаризация по методу Оцу, автоматически определяющая оптимальный порог путем максимизации межклассовой дисперсии интенсивностей пикселей, что обеспечивает устойчивость к вариациям освещенности и в результате дает бинарное изображение с проволокой, представленной белыми пикселями, а с фоном – черными.

Перед выполнением количественных измерений производится калибровка пиксельной системы для определения коэффициента пересчета пикселей в миллиметры. Пользователю предъявляется бинарное изображение, на котором необходимо последовательно указать две точки, соответствующие известному физическому расстоянию на эталонном участке проволоки, причем координаты точек фиксируются с учетом масштабирования окна отображения. Расстояние между выбранными точками в пикселях вычисляется по формуле Евклида

$$d_{px} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\ , \tag{1}$$

где x_1 и y_1 – координаты эталона (начальная точка); x_2 и y_2 – координаты эталона (конечная точка).

После ввода реального расстояния в миллиметрах коэффициент разрешения определяется как отношение пиксельного расстояния к физическому согласно выражению (2), причем полученное значение используется для последующих линейных и площадных измерений.

$$k = \frac{d_{px}}{d_{mm}}, \tag{2}$$

где d_{px} – диаметр в пикселях; d_{mm} – диаметр в миллиметрах (исходные размеры эталона).

Для определения диаметра применяется метод преобразования расстояний, который включает следующие этапы: на бинарном изображении находятся внешние контуры, среди которых выбирается контур с наибольшей площадью, соответствующий контуру проволоки; создается маска проволоки, заполненная значением 255; для каждой точки внутри объекта вычисляется евклидово расстояние до ближайшей границы, после чего находится глобальный максимум преобразования расстояний, соответствующий радиусу наибольшей вписанной окружности, центр которой принимается за геометрический центр сечения проволоки. Диаметр в пикселях определяется как удвоенное значение этого радиуса, а диаметр в миллиметрах получается делением пиксельного

диаметра на калибровочный коэффициент. Результат измерения диаметра для проволоки диаметром 1,8 мм показан на рис. 3.

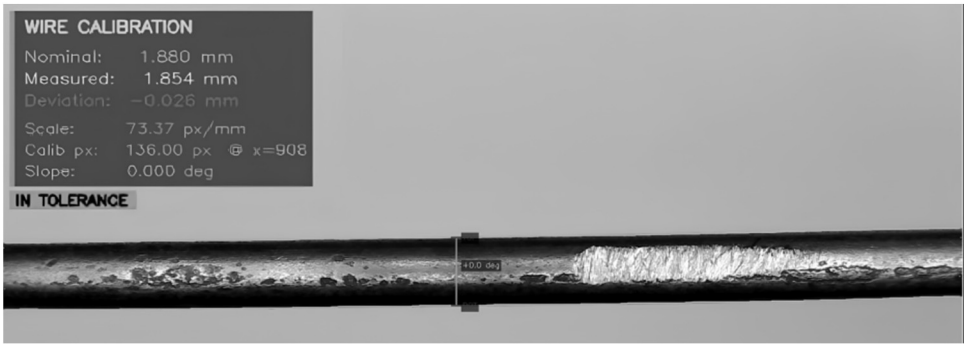

Рис. 3. Пример измерения диаметра на канатной проволоке диаметром 1,8 мм

Fig. 3. An *e*xample of diameter measurement on the rope wire with a diameter of 1,8 mm

Детекция дефектов основана на анализе темных областей внутри маски проволоки, которые соответствуют локальным нарушениям сплошной закраски поверхности. Алгоритм включает инвертирование бинарного изображения с последующим наложением маски основного контура проволоки на инвертированное изображение, что позволяет выделить только внутренние темные области. Затем выполняется поиск контуров потенциальных дефектов с фильтрацией мелких контуров по площади (порог не менее пяти пикселей) для исключения шума, причем для каждого значимого контура вычисляется площадь, а результаты визуализируются путем нанесения ограничивающих рамок и меток на исходное бинарное изображение в цвете, как показано на рис. 4, где представлены примеры детекции дефектов различных размеров.

Рис. 4. Пример детекции дефекта на канатной проволоке диаметром 1,8 мм:
1 – дефект размером 1,11 × 0,8 мм; *2* – дефект размером 0,36 × 0,46 мм; *3* – дефект размером 1,23 × 1,7 мм

Fig. 4. An example of defect detection on the rope wire with a diameter of 1,8 mm:
1 is a defect measuring 1,11 × 0,8 mm; *2* is a defect measuring 0,36 × 0,46 mm; *3* is a defect measuring 1,23 × 1,7 mm

Повреждения определяются как темные области внутри маски проволоки, для каждого дефекта определяются контуры, площадь, размеры ограничивающего прямоугольника и реальные значения ширины и высоты в миллиметрах, после чего на изображение наносится визуальная разметка в виде рамок и масок. Таким образом, использование описанного алгоритма позволяет измерять геометрические параметры и обнаруживать наличие дефектов в канатной проволоке.

2. АЛГОРИТМ КОНТРОЛЯ КАНАТНОЙ ПРОВОЛОКИ

Изложенная концепция измерения размеров закладывает теоретический фундамент для количественной оценки геометрических параметров канатной проволоки. В ее рамках определены ключевые методы обработки изображений: адаптивная бинаризация по Оцу, калибровка пространственного разрешения, преобразование расстояний для определения диаметра и алгоритмы поиска темных областей внутри маски для выявления поверхностных дефектов. Эти методы образуют основу, на которой строится дальнейшая практическая реализация системы контроля.

Предложенная концепции была реализована в программно-аппаратном комплексе на основе разработанного детализированного алгоритма. Его особенностью является то, что он последовательно реализует все этапы обработки данных. Сначала обеспечивается подготовка оборудования и калибровка устройства по эталонному образцу. В последующем выполняется захват видео движущейся проволоки, покадровая обработка с применением описанных математических процедур. На завершающем этапе выполняется агрегация результатов. По мнению авторов статьи, алгоритм выступает связующим звеном между теорией и инженерной реализацией, обеспечивая автоматизированное выполнение измерений в соответствии с установленными метрологическими требованиями. Сам алгоритм состоит из нескольких этапов, рассмотренных далее.

Этап 1. На этом этапе, как было указано ранее, выполняется подготовка и калибровка системы.

1. Производится выбор эталонного образца. Для этого используется самая распространенная проволока с известным диаметром (например, 1,80 мм по ГОСТ 7372–79) и геометрическими параметрами, аттестованными производителем проволоки.

2. В процессе настройки оптической системы устанавливаются камеры высокого разрешения, а также обеспечивается равномерное освещение и фокусировка.

3. Калибровка разрешения выполняется в пикселях на миллиметр. Для этого получаем бинарное изображение эталона по методу Оцу. Далее осуществляется выбор двух точек на изображении, расстояние между которыми известно в миллиметрах. При этом данные предоставляются оператором системы или с помощью автоматического алгоритма выбора. На основе выбранных параметров вычисляется евклидово расстояние в пикселях по формуле (1) и коэффициент разрешения по формуле (2).

4. Для оценки погрешности системы выполняется серия измерений эталона (не менее 20 раз) в одинаковых условиях. В последующем вычисляется средняя ошибка, СКО и доверительный интервал. Авторы статьи на основе проведенных апробаций получили погрешность ±0,022 мм, что значительно меньше допустимой, составляющей 1,5 %. Конечно, при необходимости выполняется повторная корректировка персональных настроек, таких как освещение и пороги бинаризации.

Этап 2. Этот этап алгоритма является унифицированным и заключается в основном в сборе данных на основе видеофиксации. В качестве основных действий авторы выделяют:

– протяжку исследуемой проволоки с наработкой 500…3500 ч перед камерами;

– видеозахват с высокой частотой кадров;

– декомпозицию видео на отдельные кадры (например, каждый кадр или с заданным шагом).

Этап 3. На этом этапе проводится измерение диаметра и поиск дефектов, что подразумевает последовательную обработку каждого кадра. Для каждого кадра выполняется последовательность операций.

3.1. Операция предварительной обработки изображения подразумевает три действия: преобразование в оттенки серого; адаптивная бинаризация методом Оцу, при котором фон окрашивается в черный цвет, а проволока – в белый; извлечение основного контура, определенного по максимальной площади, который является маской проволоки.

3.2. При необходимости выполняется операция калибровки текущего кадра, если условия съемки меняются и можно применить калибровку по реперным точкам. Однако в настоящей работе авторы предполагают неизменность коэффициента k, полученного на этапе 1.

3.3. Операция измерения диаметра подразумевает следующие действия:

– преобразование расстояний к маске проволоки;

– определение глобального максимума $r_{\max}$, который определяется как радиус максимальной вписанной окружности;

– расчет диаметра в пикселях: $D_{px} = 2r_{\max} D_{px} = 2r_{\max}$;

– пересчет результата в миллиметрах: $D_{mm} = D_{px} / k$;

– сохранение результата как текущий диаметр сечения.

3.4. Операция выявления поверхностных дефектов состоит из последовательности следующих действий:

– инвертирование бинарного изображения, при котором дефекты преобразуются в темные области на светлой проволоке и становятся светлыми на темном фоне;

– маскирование проволоки используется для того, чтобы рассматривать только внутренние области;

– поиск контуров потенциальных дефектов;

– фильтрация шумов по площади (например, порог ≥5 пикселей);

– вычисление для каждого найденного дефекта:

  o площади (в пикселях в квадрате и миллиметрах в квадрате после умножения на $k2k2$);

- габаритов ограничивающего прямоугольника (ширина, высота в миллиметрах);
- координат центра;

– визуализация дефектов рамками на изображении.

3.5. Операция нейросетевой классификация дефектов основана на использовании нейросетевой модели, разработанной авторами статьи, которая предварительно обучена. При этом изображения дефектных областей в виде вырезанных фрагментов подаются на вход сверточной нейронной сети. Далее сеть определяет дефект и относит его к одному из классов: «поджатие», «задир», «перегиб», «эллипсность», «прочие». При этом также возможно использование нейросети для уточнения границ дефекта, например сегментация на основе U-Net, если качество бинаризации недостаточное.

3.6. Операция записи результатов обеспечивает сохранение в базу данных или файл таких характеристик дефекта, как номер кадра, измеренный диаметр, список дефектов с характеристиками и классами.

Этап 4. На этапе агрегации и анализа по всей длине проволоки строится профиль диаметра, который определяется графиком зависимости диаметра от координаты. Также вычисляются статистические параметры: средний диаметр, минимальный / максимальный, отклонения от номинала. При этом отчет о количестве дефектов каждого типа формируется относительно единиц длины. Если выявляется превышение заданных допусков, система сигнализирует о браке.

Этап 5. На последнем этапе выполняется периодическая верификация как повторное измерение эталонного образца для контроля дрейфа характеристик системы. Также при необходимости выполняется перекалибровка или дообучение нейросети на новых размеченных данных.

Представленный алгоритм охватывает все стадии автоматизированного контроля канатной проволоки: от первичной калибровки оптической системы до финальной агрегации результатов и периодической верификации. Он объединяет классические методы обработки изображений (бинаризация по Оцу, преобразование расстояний, контурный анализ) с современными нейросетевыми подходами к классификации дефектов, что позволяет не только количественно оценивать геометрические параметры, но и идентифицировать типы повреждений. Последовательное выполнение этапов обеспечивает получение достоверных данных о диаметре, выявление поверхностных дефектов и формирование статистических отчетов, необходимых для контроля качества.

Для практической реализации описанного алгоритма разработан программный комплекс на языке Python, структура которого полностью соответствует логике технологического процесса. Модульная архитектура включает компоненты калибровки, предварительной обработки, измерения диаметра, детекции и нейросетевой классификации дефектов, а также модули обработки видеопотока и анализа погрешностей. Использование библиотек OpenCV для операций компьютерного зрения и TensorFlow/Keras для нейросетевых моделей обеспечивает высокую производительность и гибкость настройки системы под конкретные условия эксплуатации.

3. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА КОНТРОЛЯ ГЕОМЕТРИЧЕСКИХ ПАРАМЕТРОВ КАНАТНОЙ ПРОВОЛОКИ

Разработанный алгоритм контроля канатной проволоки реализован в виде программного комплекса, построенного по модульному принципу на языке Python с использованием библиотек компьютерного зрения и машинного обучения. Выбор языка Python обусловлен наличием развитой экосистемы библиотек для обработки изображений и научных вычислений, а также возможностью быстрого прототипирования и отладки алгоритмов. Архитектура программного комплекса обеспечивает гибкость настройки, возможность замены отдельных компонентов и масштабирования системы при переходе к промышленной эксплуатации. Основными функциональными модулями являются модуль захвата и предварительной обработки изображений, модуль калибровки, модуль измерения геометрических параметров, модуль детекции дефектов и модуль нейросетевой классификации. Взаимодействие между модулями организовано через четко определенные интерфейсы, что позволяет независимо развивать каждый компонент и адаптировать систему под конкретные условия эксплуатации.

Модуль калибровки системы предназначен для определения коэффициента пересчета пикселей в миллиметры, что является необходимым условием для получения количественных результатов измерений. Процесс калибровки реализован в интерактивном режиме с использованием графического интерфейса на основе библиотеки OpenCV. Оператору предъявляется бинарное изображение эталонного участка проволоки с известным физическим расстоянием, на котором необходимо последовательно указать две точки. Координаты выбранных точек фиксируются с учетом масштабирования окна отображения, после чего вычисляется евклидово расстояние между ними в пикселях. Коэффициент разрешения определяется как отношение полученного пиксельного расстояния к известному физическому расстоянию в миллиметрах. Полученное значение сохраняется в файл конфигурации и используется во всех последующих измерениях. Для повышения точности калибровки предусмотрена возможность усреднения результатов по нескольким парам точек и проведения повторной калибровки при изменении условий съемки.

Модуль предварительной обработки изображений выполняет сегментацию проволоки на исходном изображении и формирование маски для последующих измерений. Основным методом сегментации является адаптивная пороговая бинаризация по методу Оцу, который автоматически определяет оптимальный порог на основе анализа гистограммы интенсивностей пикселей, обеспечивает устойчивость к вариациям освещенности и не требует ручной настройки пороговых значений. После бинаризации выполняется поиск всех внешних контуров на изображении, среди которых выбирается контур с максимальной площадью, соответствующий искомой проволоке. На основе выбранного контура формируется бинарная маска, где область проволоки заполнена значением 255, а фон представлен нулями. Для удаления мелких шумовых артефактов применяются морфологические операции открытия и закрытия с ядром небольшого размера, что позволяет сгладить границы маски и устранить единичные выбросы.

Модуль измерения диаметра реализует определение геометрического центра и диаметра проволоки на основе маски, полученной на предыдущем этапе. Ключевым методом является преобразование расстояний (distance transform), которое для каждой точки внутри маски вычисляет евклидово расстояние до ближайшей границы. Результатом преобразования является матрица, значения которой соответствуют расстоянию от каждой внутренней точки до границы объекта. Глобальный максимум этой матрицы соответствует радиусу наибольшей вписанной окружности, а его координаты определяют геометрический центр сечения проволоки. Диаметр в пикселях вычисляется как удвоенное значение этого радиуса, после чего с использованием калибровочного коэффициента выполняется пересчет в миллиметры. Для оценки стабильности измерений модуль поддерживает режим многократных измерений с небольшими пространственными смещениями области интереса, что позволяет оценить воспроизводимость результатов и выявить возможные аномалии.

Модуль детекции дефектов предназначен для выявления поверхностных повреждений проволоки, таких как поджатие, задиры, перегибы и локальная эллипсность. Алгоритм детекции основан на анализе темных областей внутри маски проволоки, которые соответствуют локальным нарушениям сплошности поверхности. Исходное бинарное изображение инвертируется, после чего на него накладывается маска проволоки, что позволяет выделить только внутренние области, потенциально содержащие дефекты. Далее выполняется поиск контуров на полученном изображении с последующей фильтрацией по минимальной площади для исключения шумовых артефактов. Для каждого значимого контура вычисляются геометрические характеристики: площадь в пикселях и миллиметрах, габариты ограничивающего прямоугольника, ширина и высота в миллиметрах, координаты центра. Полученные данные сохраняются для последующей классификации и визуализации. Для повышения надежности детекции предусмотрена возможность настройки пороговых значений и применения дополнительных морфологических операций для улучшения качества выделения границ дефектов.

Модуль нейросетевой классификации и детекции дефектов реализован на базе архитектуры YOLOv8. Данная модель выбрана благодаря высокой эффективности обработки видеопотока и способности выполнять одновременную локализацию и классификацию объектов в рамках одного прохода. Архитектура модели включает *backbone*-сеть (на базе C2f blocks) для извлечения иерархических признаков, *neck*-структуру для агрегации признаков разных масштабов и *head*-блоки, ответственные за предсказание границ дефектов (bounding boxes) и их вероятностную классификацию.

В отличие от классических классификаторов с функцией softmax, в YOLOv8 для классификации используется функция активации sigmoid, что позволяет независимо оценивать вероятность принадлежности объекта к каждому из целевых классов: «поджатие», «задир», «перегиб», «эллипсность» и др. Обучение модели проводилось методом transfer learning с использованием предобученных весов при разрешении входного изображения 640×640 пикселей. Для оптимизации процесса обучения применялся оптимизатор Adam и комплексная функция потерь, объединяющая компоненты регрессии ограничивающих рамок (CIoU loss) и классификации (Binary Cross-Entropy). Использование

методов геометрической и фотометрической аугментации позволило повысить устойчивость модели к вариативности условий освещенности и ракурса съемки.

Модуль обработки видеопотока обеспечивает возможность непрерывного контроля движущейся проволоки в реальном времени или при обработке предварительно записанных видеоматериалов; выполняет декомпозицию видео на отдельные кадры с заданной частотой (например, каждый кадр или с определенным шагом), после чего последовательно применяет к каждому кадру описанные выше алгоритмы предварительной обработки, измерения диаметра, детекции и классификации дефектов. Для обеспечения производительности при работе с высокоскоростными видеопотоками реализована поддержка многопоточной обработки с использованием библиотеки multiprocessing, что позволяет эффективно загружать многоядерные процессоры. Результаты обработки каждого кадра сохраняются в структурированном формате JSON, включая измеренный диаметр, координаты центра, количество обнаруженных дефектов и их характеристики. Визуализированные кадры с нанесенной разметкой сохраняются в отдельную директорию для последующего визуального контроля и анализа.

Модуль анализа погрешности и валидации реализует статистическую обработку результатов многократных измерений эталонных образцов в соответствии с требованиями ГОСТ Р 8.736–2011. Модуль вычисляет основные статистические характеристики: среднее арифметическое значение, среднее квадратическое отклонение, доверительные границы случайной погрешности, а также систематическую составляющую погрешности на основе сравнения с аттестованным значением эталона. Для выявления грубых погрешностей и выбросов применяется критерий Граббса. Результаты анализа представляются в виде отчета, содержащего оценку точности измерительной системы, что позволяет обоснованно судить о достоверности получаемых результатов и принимать решения о необходимости дополнительной калибровки или корректировки алгоритмов.

Для тестирования разработанного программного комплекса был подобран набор эталонных изображений и видеоматериалов, данные для которого предоставило ООО «Альтаир» (г. Томск). Полученные результаты тестирования подтвердили адекватность и корректность разработанного алгоритма, а также соответствие требуемым метрологическим характеристикам. При этом модульность, заложенная в программное обеспечение, позволяет в дальнейшем выполнить расширение функциональности. Это позволит в будущем обеспечить интеграцию дополнительных методов анализа, реализовать поддержку новых типов дефектов и провести адаптацию к различным условиям эксплуатации. В работе авторов были использованы оптимизированные библиотеки компьютерного зрения и методы распараллеливания вычислительной нагрузки. Это, в свою очередь, позволяет достичь производительности, достаточной для использования системы в режимах, приближенных к реальному времени. По мнению авторов, работа открывает широкие перспективы для внедрения разработанных методов и подходов для промышленных линий контроля качества канатной проволоки.

4. АПРОБАЦИЯ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

На основе серии экспериментальных исследований была доказана достоверность и адекватность предложенного алгоритма, а также подтверждена его пригодность для применения в практических задачах поиска дефектов канатной проволоки. Необходимость автоматизации контроля качества проволоки обусловлена ограничениями традиционных визуальных методов, которые характеризуются субъективностью и подвержены влиянию человеческого фактора, что особенно критично при больших объемах производства. Внедрение систем технического зрения позволяет не только повысить объективность контроля, но и обеспечить его непрерывность в реальном времени с высокой точностью измерений.

Методология апробации разработанного алгоритма основывалась на сравнительном анализе результатов измерений, полученных с помощью программного комплекса, и эталонных значений. В качестве объекта исследований использовалась проволока диаметром 1,80 мм, изготовленная по ГОСТ 7372–79, с наработкой от 500 до 3500 ч, на которой был выявлен ряд потенциальных дефектов, включая поджатие, задиры, перегиб и прочие поверхностные повреждения. Эксперимент проводился в статических условиях при фиксированном положении проволоки перед камерами высокого разрешения. Для обеспечения репрезентативности результатов было выполнено двадцать последовательных измерений диаметра эталонного образца в идентичных условиях, что соответствует требованиям ГОСТ Р 8.736–2011 к обработке результатов прямых многократных измерений. Полученные данные представлены в табл. 1, демонстрирующей незначительный разброс значений от 1,79 до 1,81 мм при номинальном диаметре 1,80 мм, метрики – в табл. 2.

На точность измерения геометрических размеров объектов с применением системы машинного зрения влияет не только пространственное расположение камеры относительного самого объекта, но также внутренние технические характеристики камеры, дискретизация яркости и шумы изображения [10].

При использовании веб-камеры ключевым источником систематической погрешности измерения диаметра такой проволоки является погрешность квантования, обусловленная дискретной (пиксельной) природой цифрового изображения.

При настройке измерительной системы с использованием образца с известными геометрическими размерами был определен масштабный коэффициент системы $k = 0{,}0328$ мм/пиксель.

Поскольку диаметр исследуемой проволоки достаточно малого размера и его изображение не занимает целое число пикселей, это приводит к погрешности квантования. При этом погрешность квантования можно рассчитать по формуле

$$\Delta d = \sqrt{\Delta x_1{}^2 + \Delta x_2{}^2} \cdot k = 0{,}7k,$$

где Δx_1, Δx_2 – погрешность определения левой и правой границ проволоки, вызванная нецелочисленным заполнением пикселя, каждая из которых принимается равной $\pm 0{,}5$ пикселя; k – масштабный коэффициент, мм/пиксель.

Таким образом, погрешность квантования равна 0,022 мм.

При использовании веб-камеры однозначно возникают случайные погрешности, которые могут быть вызваны недостаточным освещением и яркостью изображения.

Для определения суммарной погрешности измерения диаметра с учетом систематической и случайной составляющих погрешности воспользуемся ГОСТ Р 8.736.

Благодаря возможностям измерительной системы был получен массив данных измерения диаметра канатной проволоки в различных точках. Для расчета случайной составляющей погрешности учитывались только первые 20 измерений. При этом среднее арифметическое значение равно 1,80 мм, среднеквадратическое отклонение 0,0013, следовательно, случайная составляющая погрешности измерения диаметра проволоки составляет ± 0,002 мм. В соответствии с алгоритмом обработки прямых многократных измерений получена суммарная погрешность измерения диаметра с применением разработанной измерительной системы, и она составила ± 0,023 мм.

Таблица 1

Table 1

Группа измерений

Measurement Group

Номер измерения	1	2	3	4	5	6	7	8	9	10
Диаметр, мм	1,80	1,79	1,81	1,80	1,79	1,81	1,80	1,80	1,81	1,79
Номер измерения	11	12	13	14	15	16	17	18	19	20
Диаметр, мм	1,80	1,80	1,79	1,81	1,80	1,80	1,81	1,79	1,80	1,80

Обработка экспериментальных данных включала расчет суммарной и случайной погрешностей для каждого измерения, результаты которого представлены в табл. 2.

Таблица 2

Table 2

Обработанные экспериментальные данные

Processed experimental data

Группа измерений	Номер измерения	Суммарная погрешность	Случайная погрешность, мм
1	1,80	0,00	0,0000
2	1,79	–0,01	0,0001
3	1,81	0,01	0,0001
4	1,80	0,00	0,0000
5	1,79	–0,01	0,0001

Окончание табл. 2

End of the Tab. 2

Группа измерений	Номер измерения	Суммарная погрешность	Случайная погрешность, мм
6	1,81	0,01	0,0001
7	1,80	0,00	0,0000
8	1,80	0,00	0,0000
9	1,81	0,01	0,0001
10	1,79	-0,01	0,0001
11	1,80	0,00	0,0000
12	1,80	0,00	0,0000
13	1,79	–0,01	0,0001
14	1,81	0,01	0,0001
15	1,80	0,00	0,0000
16	1,80	0,00	0,0000
17	1,81	0,01	0,0001
18	1,79	–0,01	0,0001
19	1,80	0,00	0,0000
20	1,80	0,00	0,0000

Таблица 3

Table 3

Метрики качества

Quality metrics

Класс дефекта	Точность (Precision)	Полнота (Recall)	F1-Score
Поджатие	0,94	0,92	0,93
Задир	0,91	0,89	0,90
Перегиб	0,96	0,95	0,95
Эллипсность	0,88	0,91	0,89
Прочие	0,85	0,82	0,83
Среднее (mAP@0.5)	0,91	0,90	0,90

Анализ погрешностей показал, что отклонения измеренных значений от эталонного диаметра носят случайный характер и не превышают ±0,01 мм для

отдельных измерений. На основе полученной выборки была рассчитана итоговая погрешность измерительной системы, составившая ±0,022 мм, что соответствует относительной погрешности не более 1,5 % от номинального диаметра. Достигнутая точность позволяет проводить измерения диаметра и геометрических параметров дефектов с погрешностью менее 1 % и дискретностью до 0,001 мм.

Полученные результаты согласуются с современными тенденциями развития оптических измерительных систем. В качестве подтверждения был проведен анализ результатов проекта VisionPlus Швейцарской высшей школой прикладных наук. Результаты этого проекта продемонстрировали, что возможность измерения диаметра проволоки в диапазоне от 0,010 до 5,000 мм и с точностью 1 % именно для диаметров свыше 0,200 мм и около 1 мкм можно использовать для более тонких изделий [12]. Представленные исследования в области контроля волоконно-оптических катушек также можно интерпретировать как подтверждение эффективности использования методов машинного зрения. Однако надо отметить, что разработанный алгоритм использует метод бинаризации Оцу, а преобразования расстояний, как правило, обеспечивают точность, очень близкую к коммерческим системам, имеющим стоимость на порядок выше стоимости разработки авторов статьи. Было доказано, что при меньшей вычислительной сложности можно достичь сопоставимых результатов, что подтверждается успешным применением аналогичных подходов в задачах комбинации оптических элементов.

В заключение следует отметить, что экспериментальная апробация полностью подтвердила работоспособность предложенного алгоритма, показано его соответствие современным требованиям к точности измерений в системах технического зрения. Достигнутые показатели погрешности позволяют рекомендовать разработанный метод для практического применения в задачах неразрушающего контроля канатной проволоки.

ЗАКЛЮЧЕНИЕ

В рамках выполненного исследования разработан алгоритмический и программный базис системы неразрушающего контроля геометрических параметров канатной проволоки, интегрирующий методы компьютерного зрения и нейросетевой классификации поверхностных дефектов. Предложенный подход включает калибровку оптической системы с использованием эталонных образцов, адаптивную бинаризацию по методу Оцу для сегментации проволоки, преобразование расстояний при определении диаметра и контурный анализ для выявления локальных повреждений, классификация которых выполняется сверточной нейронной сетью. Программная реализация на языке Python с применением библиотек OpenCV и TensorFlow обеспечивает модульность, масштабируемость и возможность обработки видеопотока в режиме, близком к реальному времени.

Экспериментальная апробация на эталонных образцах проволоки диаметром 1,80 мм с наработкой от 500 до 3500 ч подтвердила достоверность получаемых результатов. После двадцати циклов измерений удалось выйти на среднеквадратическую погрешность ±0,022 мм. В относительном выражении это не превышает 1,5 %. При таких показателях диаметр проволоки и геомет-

рические параметры дефектов можно фиксировать с дискретностью до 0,001 мм.

Полученные метрологические характеристики вполне сопоставимы с теми, которые дают современные коммерческие оптические системы. А значит, разработанный метод вполне пригоден для автоматизированного контроля качества прямо в производственных условиях.

Дальнейшие планы авторы связывают с доработкой нейросетевых моделей как для сегментации, так и для классификации дефектов. Кроме того, предстоит адаптировать алгоритмы к работе с высокоскоростными видеопотоками непосредственно на технологической линии. В качестве отдельного направления дальнейших исследований можно выделить задачу встраивания системы в единую цифровую платформу управления качеством. Это позволит накапливать статистику и в перспективе прогнозировать остаточный ресурс канатной проволоки.

СПИСОК ЛИТЕРАТУРЫ

1. Новый лазерный микрометр от компании LaserLink // RusCable.ru. – 2016. – URL: https://www.ruscable.ru/news/2016/06/27/Novyj_lazernyj_mikrometr_ot_kompanii_LaserLink/ (дата обращения: 19.02.2026).

2. *Khodier S.A.* Measurement of wire diameter by optical diffraction // Optics & Laser Technology. – 2004. – Vol. 36 (1). – P. 63–67. – DOI: 10.1016/S0030-3992(03)00134-8.

3. Topography scanning as a part of process monitoring in power cable insulation process / J. Harjuhahto, J. Harjuhahto, M. Lahti, J. Hanhirova, B. Sonerud // arXiv.org. – 2026. – arXiv:2602.06519. – URL: https://arxiv.org/abs/2602.06519 (accessed: 19.02.2026).

4. США: LaserLink Inc. выпустила сканирующий лазерный микрометр // Ukrmach. – 2016. – URL: https://ukrmach.dp.ua/2016/07/04/ssha-laserlink-inc-vypustila-skaniruyushhij-lazernyj-mikrometr.html (дата обращения: 19.02.2026).

5. *West P.* On-line gauge for wire and fibre diameter measurement // Measurement and Control. – 1974. – Vol. 7 (2). – P. 45–46. – DOI: 10.1177/002029407400700201.

6. *Mishani I., Sintov A.* Learning configurations of wires for real-time shape estimation and manipulation planning // Engineering Applications of Artificial Intelligence. – 2023. – Vol. 121. – DOI: 10.1016/j.engappai.2023.105967.

7. Лазерный микрометр TRUBACH Microline L360 // Катай машиностроение: сайт. – URL: https://czcathay.ru/equipment/lazernyj-mikrometr-trubach-microline-l360 (дата обращения: 19.02.2026).

8. Absolute diameter measurement of thin wire, thread and fibre // CORDIS. – URL: https://cordis.europa.eu/project/id/SMT4952042 (accessed: 19.02.2026).

9. As Good as it Gets: NIST Develops Its Fourth Generation Wire Micrometer that Rivals Best in the World // National Institute of Standards and Technology. – 2021. – URL: https://www.nist.gov/news-events/news/2021/03/good-it-gets-nist-develops-its-fourth-generation-wire-micrometer-rivals (accessed: 19.02.2026).

10. Комплексный подход к неразрушающему контролю исследовательской проволоки с использованием цифровых технологий и нейронных систем / С.Г. Цапко, В.С. Хорев, И.В. Цапко, А.А. Друки // Перспективы науки. – 2025. – № 11 (194). – С. 89–96. – URL https://moofrnk.com/assets/files/journals/science-prospects/194/science-prospect-11(194)-main.pdf (дата обращения: 31.03.2026).

11. New system for measuring wire diameter with linear cameras // Scuola universitaria professionale della Svizzera Italiana (SUPSI): website. – 2024. – URL: https://www.supsi.ch/en/nuovo-sistema-per-misurare-il-diametro-dei-fili-con-telecamere-lineari (accessed: 31.03.2026).

12. VisionPlus: Next generation inspection system for wires and filaments // ARIS: Accessible Research Information System of SUPSI. – URL: https://aris.supsi.ch/entities/funding/239b9917-28b0-4028-af2f-479ea00d8cc9 (accessed: 31.03.2026).

Хорев Владимир Сергеевич, исполняющий обязанности руководителя отделения информационных технологий инженерной школы информационных технологий и робототехники Томского политехнического университета, доцент, кандидат технических наук. Основное направление научных исследований – неразрушающие методы контроля. Имеет более 10 печатных работ. E-mail: horevvs@tpu.ru

Цапко Сергей Геннадьевич, директор научно-образовательного центра топ-инженерии в сфере информационных технологий инженерной школы информационных технологий и робототехники Томского политехнического университета, доцент, кандидат технических наук. Основное направление научных исследований – разработка огнетушащих средств специального назначения, разработка нейросетвых алгоритмов идентификации личности, предсказания аварийных ситуаций в сложных технических системах, разработка и синтез высокопрочных изделий на основе новых композитных материалов, автоматизация промышленных производств, использование нейросетвых алгоритмов в задачах прогнозирования динамики изменения курса акций, имитационное моделирование сложных технических систем. Имеет более 130 печатных работ. E-mail: tsapko@tpu.ru

Кузьминская Елена Вячеславовна, доцент отделения автоматизации и робототехники инженерной школы информационных технологий и робототехники Томского политехнического университета, доцент, кандидат технических наук. Основное направление научных исследований – метрология, стандартизация и сертификация, автоматизация управления жизненным циклом продукции, основы технического регулирования. Имеет более 30 печатных работ. E-mail: bedareva@tpu.ru

Командинов Павел Александрович, инженер отделения информационных технологий инженерной школы информационных технологий и робототехники Томского политехнического университета. Основное направление научных исследований – мехатроника, программирование. E-mail: pak37@tpu.ru

Khorev Vladimir S., Acting Head of the Information Technology Division, School of Information Technology and Robotics, Tomsk Polytechnic University; Associate Professor; PhD (Eng.). His primary research focusis non-destructive testing methods. He is the author of over 10 publications. E-mail: horevvs@tpu.ru

Tsapko Sergey G., Director of the Top Engineering Research and Education Center in Information Technology, School of Information Technology and Robotics, Tomsk Polytechnic University; Associate Professor; PhD (Eng.). His research interests include development of special-purpose fire-extinguishing agents, neural network algorithms for personal identification, prediction of emergencies in complex technical systems, development and synthesis of high-strength products based on new composite materials, automation of industrial processes, application of neural network algorithms in stock price dynamics forecasting, and simulation modeling of complex technical systems. He is the author of over 130 publications. E-mail: tsapko@tpu.ru

Kuzminskaya Elena V., Associate Professor, Division of Automation and Robotics, School of Information Technology and Robotics, Tomsk Polytechnic University; Associate Professor; PhD (Eng). Her research interests cover metrology, standardization and certification, automated product lifecycle management, and fundamentals of technical regulation. She is the author of over 30 publications. E-mail: bedareva@tpu.ru

Komandinov Pavel A., engineer, Information Technology Division, School of Information Technology and Robotics, Tomsk Polytechnic University. His research interests include mechatronics and programming. E-mail: pak37@tpu.ru

DOI: 10.17212/2782-2001-2026-3-131-150

Reliability of measuring rope wire dimensions using digital technologies and neural networks *

V.S. KHOREV [a], S.G. TSAPKO [b], E.V. KUZMINSKAYA [c], P.A. KOMANDINOV [d]

National Research Tomsk Polytechnic University, 30 Lenina Avenue, Tomsk, 634050, Russian Federation

[a] horevvs@tpu.ru [b] tsapko@tpu.ru [c] bedareva@tpu.ru [d] pak37@tpu.ru

Abstract

The relevance of the study is driven by the need to improve the reliability and efficiency of monitoring the geometric parameters of rope wire, since traditional contact methods do not allow real-time measurements on moving objects, while visual inspection is subjective and does not provide quantitative assessment. The transition to non-contact optical systems combined with intelligent data processing opens up the possibility of automated non-destructive testing with the required accuracy and repeatability of results.

The aim of the work is to develop and provide metrological substantiation for a method of measuring the diameter and detecting surface defects of rope wire based on machine vision and neural network algorithms. To achieve this goal, the following tasks were addressed: calibration of the optical system using working standards, development of image segmentation and analysis algorithms, and defect classification using a convolutional neural network.

As a result, an image processing algorithm has been proposed and implemented, including adaptive Otsu binarization, distance transformation for diameter determination, contour analysis for defect localization, and subsequent classification of defects using a neural network model. Experimental testing on wire with a diameter of 1.80 mm and a service life ranging from 500 to 3500 hours confirmed the viability of the approach: the measurement error of the diameter was ±0.022 mm (less than 1.5 %), ensuring result resolution down to 0.001 mm. The developed software package in Python enables automated monitoring in near real-time and can be recommended for implementation into non-destructive quality control systems.

Keywords: rope wire, computer vision, binary image, video recording, visualization, filtering, jamming, scoring, bending, neural network, defect, reference standard, python

REFERENCES

1. New laser micrometer from LaserLink company. *RusCable.ru*, 2016. (In Russian). Available at: https://www.ruscable.ru/news/2016/06/27/Novyj_lazernyj_mikrometr_ot_kompanii_LaserLink/ (accessed 19.02.2026).

2. Khodier S.A. Measurement of wire diameter by optical diffraction. *Optics & Laser Technology*, 2004, vol. 36 (1), pp. 63–67. DOI: 10.1016/S0030-3992(03)00134-8.

3. Harjuhahto J., Harjuhahto J., Lahti M., Hanhirova J., Sonerud B. Topography scanning as a part of process monitoring in power cable insulation process. *arXiv.org*, 2026, arXiv:2602.06519. Available at: https://arxiv.org/abs/2602.06519 (accessed 19.02.2026).

4. USA: LaserLink Inc. released scanning laser micrometer. *Ukrmach*, 2016. Available at: https://ukrmach.dp.ua/2016/07/04/ssha-laserlink-inc-vypustila-skaniruyushhij-lazernyj-mikrometr.html (accessed 19.02.2026).

5. West P. On-line gauge for wire and fibre diameter measurement. *Measurement and Control*, 1974, vol. 7 (2), pp. 45–46. DOI: 10.1177/002029407400700201.

* *Received 07 April 2026.*

6. Mishani I., Sintov A. Learning configurations of wires for real-time shape estimation and manipulation planning. *Engineering Applications of Artificial Intelligence*, 2023, vol. 121. DOI: 10.1016/j.engappai.2023.105967.

7. Laser micrometer TRUBACH Microline L360. *Changzhou Cathay Machinery Intelligent Technology Co., Ltd*. Website. (In Russian). Available at: https://czcathay.ru/equipment/lazernyj-mikrometr-trubach-microline-l360 (accessed 19.02.2026).

8. CORDIS. *Absolute diameter measurement of thin wire, thread and fibre*. Available at: https://cordis.europa.eu/project/id/SMT4952042 (accessed 19.02.2026).

9. As Good as it Gets: NIST Develops Its Fourth Generation Wire Micrometer that Rivals Best in the World. *National Institute of Standards and Technology*, 2021. Available at: https://www.nist.gov/news-events/news/2021/03/good-it-gets-nist-develops-its-fourth-generation-wire-micrometer-rivals (accessed 19.02.2026).

10. Tsapko S.G., Khorev V.S., Tsapko I.V., Druki A.A. Kompleksnyi podkhod k nerazrushayushchemu kontrolyu issledovatel'skoi provoloki s ispol'zovaniem tsifrovykh tekhnologii i neironnykh sistem [A comprehensive approach to non-destructive testing of research wire using digital technologies and neural systems]. *Perspektivy nauki = Prospects of Science*, 2025, no. 11 (194), pp. 89–96. Available at: https://moofrnk.com/assets/files/journals/science-prospects/194/science-prospect-11(194)-main.pdf (accessed 31.03.2026).

11. New system for measuring wire diameter with linear cameras. *Scuola universitaria professionale della Svizzera Italiana*: website, 2024. Available at: https://www.supsi.ch/en/nuovo-sistema-per-misurare-il-diametro-dei-fili-con-telecamere-lineari (accessed 31.03.2026).

12. VisionPlus: Next generation inspection system for wires and filaments. *ARIS: Accessible Research Information System of SUPSI*. Available at: https://aris.supsi.ch/entities/funding/239b9917-28b0-4028-af2f-479ea00d8cc9 (accessed 31.03.2026).

Для цитирования:

Достоверность измерений размеров канатной проволоки с использованием цифровых технологий и нейронных сетей / В.С. Хорев, С.Г. Цапко, А.В. Кузьминская, П.А. Командинов // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 131–150. – DOI: 10.17212/2782-2001-2026-3-131-150.

For citation:

Khorev V.S., Tsapko S.G., Kuzminskaya A.V., Komandinov P.A. Dostovernost' izmerenii razmerov kanatnoi provoloki s ispol'zovaniem tsifrovykh tekhnologii i neironnykh setei [Reliability of measuring rope wire dimensions using digital technologies and neural networks]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 131–150. DOI: 10.17212/2782-2001-2026-3-131-150.

ISSN 2782-2001
Системы анализа и обработки данных
том 103, № 3, 2026, с. 151–168

http://journals.nstu.ru/vestnik
Analysis and data processing systems
Vol. 103, No. 3, 2026, pp. 151–168

КОМПЬЮТЕРНЫЕ НАУКИ И ИНФОРМАТИКА

COMPUTER SCIENCE AND INFORMATION SCIENCE

УДК 519.872

DOI: 10.17212/2782-2001-2026-3-151-168

Характеристики потока обработанных задач модели вычислительного узла при конкуренции за ресурс*

И.Л. ЛАПАТИН [a], **А.А. НАЗАРОВ** [b], **С.В. ПАУЛЬ**[c], **Д.Д. САЛИМЗЯНОВА**[d]

РФ, 634050, г. Томск, пр. Ленина, 36, Национальный исследовательский Томский государственный университет

[a] *ilapatin@mail.ru* [b] *nazarov.tsu@gmail.com* [c] *paulsv82@mail.ru*
[d] *darya2001@inbox.ru*

Постоянный рост популярности удаленных и облачных вычислений порождает необходимость строить математические модели функционирования таких систем или отдельных ее элементов. Математическая модель позволяет рассчитывать определенные характеристики системы при различных параметрах ее функционирования. Расчет характеристик необходим на этапе проектирования вычислительного узла или оптимизации работы уже существующего. Теория массового обслуживания широко применяется для моделирования процессов, имеющих стохастическую природу. В работе рассматривается модель вычислительного узла в виде системы массового обслуживания с неограниченным числом приборов. Поступление задач на узел описывается точечным процессом Пуассона. Отличительной особенностью является зависимость скорости обслуживания (обработки) поступающих задач от общего числа задач на вычислительном узле. Чем больше задач одновременно обрабатывается, тем ниже скорость обработки каждой отдельной задачи. Этот эффект возникает при конкуренции за общий ресурс вычислительного узла. В настоящей работе объектом исследования модели является выходящий поток – моменты завершения обработки задач на вычислительном узле. Информация о характеристиках выходящего потока очень важна, так как после обработки задач (запросов) они должны быть переданы по некоторым телекоммуникационным каналам связи. Именно вероятностные характеристики потока решенных задач будут влиять на расчет ширины канала передачи данных для обеспечения заданных характеристик качества обслуживания. Для анализа выходящего потока модели вычислительного узла в работе применялся метод асимптотического анализа систем массового обслуживания. В результате были получены асимптотические характеристики выходящего потока рассматриваемой модели. Важно, что предложенный асимптотический подход в дальнейшем может быть применен для моделей с непуассоновским законом поступления запросов на узел.

Ключевые слова: передача данных, вычислительные системы, моделирование вычислительного узла, конкуренция за ресурс, теория массового обслуживания, выходящий поток, метод асимптотического анализа, простейший поток

* *Статья получена 03 февраля 2026 г.*

ВВЕДЕНИЕ

С развитием вычислительных технологий и ростом их популярности среди широкого круга пользователей постоянно растет потребность в использовании вычислительных мощностей различной конфигурации и способа доступа к ним. Это могут быть как обычные персональные устройства, так и огромные серверы или кластеры, доступ к которым пользователь может получить только удаленно.

Со стороны пользователей всё более актуальными становятся облачные центры обработки данных и хранилища [1, 2]. Их использование позволяет снизить финансовые затраты на покупку и обслуживание собственного оборудования необходимой мощности, которое требуется для обработки запрашиваемого объема данных.

Со стороны поставщиков облачных услуг существует проблема баланса как технического оснащения, так и эффективного распределения запросов клиентов по физическим ресурсам системы. Для того чтобы обеспечить определенное качество обслуживания своих клиентов, но при этом не покупать и не обслуживать избыточное оборудование, провайдерам необходимо прогнозировать производительность сервера при изменении характера и интенсивности нагрузки.

Проблеме прогнозирования производительности отдельного компьютера или сервера уделялось внимание еще при росте использования вычислительной техники во второй половине XX века. В работе [3] приведен достаточно подробный и систематизированный обзор работ по оценке производительности компьютеров. В ней авторы выделяют три основных подхода: измерение, аналитическое моделирование и имитационное моделирования. Каждый из этих подходов имеет свои сильные и слабые стороны. Подход, который предлагают авторы, относится к аналитическому моделированию функционирования вычислительных узлов. Под термином «вычислительный узел» будем понимать как обычный персональный компьютер, так и удаленный облачный сервер или его элемент.

Работа вычислительного узла не может быть описана детерминированными моделями, и изменение его состояний зависит от множества факторов. Для моделирования работы облачных узлов необходимо применение стохастических методов и моделей. Для этих целей прекрасно подходит математический аппарат теории массового обслуживания [4–6].

Первой особенностью функционирования вычислительного узла является стохастический характер поступления задач или запросов. Спонтанная нагрузка на узел требует выделять больше ресурсов для того, чтобы справиться с пиковой нагрузкой, хотя большую часть времени общая нагрузка может быть значительно ниже пиковой. Стохастические свойства поступления задач определяются моделью входящего потока [7] и его параметрами.

Второй особенностью работы вычислительного узла является ограниченность ресурсов. Следовательно, при росте одновременно выполняемых задач на узле начинается конкуренция за ресурсы. Это приводит к снижению производительности, снижению скорости обработки каждой отдельной задачи. Такой эффект называется деградацией скорости облуживания [8, 9]. Для моделирования такого эффекта необходимо использовать модели со специфи-

ческими дисциплинами обслуживания поступающих задач, такими как разделение процессора (Processor Sharing) и его обобщения.

Модели с разделением процессора (Processor Sharing, PS) стали известны благодаря работе [10]. Дисциплина PS подразумевает, что заявки на обслуживание принимаются одновременно, но скорость обслуживания каждой падает с увеличением заявок в системе, при этом общая скорость обработки пропорционально делится между всеми i заявками ($1/i$). В это же время появляются работы [11] с применением более общей дисциплины – «обобщенного совместного использования процессора» (Generalized Processor Sharing, GPS). В этом случае скорость обслуживания заявок может меняться не согласно последовательности $1/i$, а согласно произвольной заданной неотрицательной последовательности.

Исследование моделей с дисциплиной GPS на узлах представлено в основополагающей статье [12]. В ней рассмотрены как закрытые, так и открытые сети массового обслуживания и формулы для их характеристик.

В последние десятилетия необходимо отметить работы [13, 14] голландского ученого B. Zwart, которые посвящены моделям с зависимостью от состояния (state-dependent). Модели с дисциплинами PS и GPS являются примерами таких систем.

Кроме того, в последние годы уделяется большое внимание проблеме конкуренции за ресурсы в туманных и граничных вычислениях (Fog/Edge Computing), в которых эта проблема стоит более остро, чем в классических облачных вычислениях. Для решения этой задачи применяют как методы ИИ [15], так и комбинированные подходы с использованием аппарата теории массового обслуживания, например [16].

Школа Томского государственного университета по теории массового обслуживания занимается исследованием подобных моделей с зависимостью скорости обработки от состояния системы различной структуры [17, 18]. Прежде всего это модели с неограниченным числом приборов на одной или нескольких фазах и деградацией скорости обслуживания. Основной особенностью исследований является разработка и применение асимптотических методов исследования [19–21]. Такой подход позволяет получать приближенные аналитические результаты для моделей, допредельное аналитическое исследование которых затруднительно или невозможно. В частности, предлагаемые асимптотические подходы в ближайшее время дадут возможность получать результаты для моделей с непуассоновским входящим потоком, что позволит значительно обобщить и расширить границы применимости полученных ранее результатов для таких моделей.

В большинстве вышеуказанных работ объектом исследования является случайный процесс изменения числа одновременно обрабатываемых задач на узле. Именно этот процесс и его вероятностно-временные характеристики определяют характеристики скорости обработки поступающих задач.

В текущей же работе объектом исследования является выходящий поток [22, 23] системы, т. е. поток обработанных задач на вычислительном узле. Мотивация изучения свойств выходящих потоков обусловлена тем, что выходящий поток одного узла может являться входящим для другого. Задачи, которые прошли обработку на процессоре, могут требовать отправления результатов своей обработки по каналам передачи данных на другой узел или

конечному потребителю. А для определения необходимой пропускной способности канала передачи данных необходимо знать стохастические свойства поступления пакетов для передачи. Поэтому исследование свойств выходящего потока безусловно является актуальной задачей.

В работе предлагается исследование выходящего потока системы с неограниченным числом приборов, на вход которой поступает простейший поток заявок с учетом зависимости скорости их обслуживания от общего числа заявок в системе.

Аналитические результаты для такой марковской модели без учета деградации скорости обслуживания были получены еще в середине XX века. В работе [24] было показано, что поток обслуженных заявок в системе является простейшим. Немного позже этот результат был обобщен для случая, когда время обслуживания заявок задано произвольной функцией распределения [25]. Это позволяет сделать вывод о том, что в системе с неограниченным числом приборов и простейшим входящим потоком выходящий поток является простейшим независимо от закона распределения времени обслуживания.

В предложенной работе ставится задача определения влияния зависимости скорости обслуживания от числа заявок в системе на свойства выходящего потока. Целью настоящей работы является адаптация и применение метода асимптотического анализа [17] для исследования выходящих потоков в таких моделях. Предложенный подход асимптотического исследования выходящего потока в дальнейшем планируется применять к системе с более общей моделью входящего потока, например MAP (Markovian Arrival Process) [26].

1. ПОСТАНОВКА ЗАДАЧИ

Для описания функционирования вычислительного узла будем применять аппарат теории массового обслуживания [27]. Поэтому обратим внимание на некоторые особенности использования терминов. Запросы и задачи, которые поступают на вычислительный узел, будем называть заявками. Моменты времени поступления или обращения заявок образуют входящий поток. Заметим, что входящий поток является случайным и может быть описан одной из моделей случайного потока событий. Модель входящего потока определяет интенсивность нагрузки на узел и ее характер (равномерный или более «скачкообразный»). Каждая заявка входящего потока поступает на обработку, которую будем называть обслуживанием. Моменты окончания обслуживания образуют события выходящего потока. Описания того, каким образом происходит обслуживание, будет дано ниже.

Предлагается рассмотреть функционирование некоторого вычислительного узла в виде системы массового обслуживания с неограниченным числом приборов [27]. Неограниченное число приборов означает, что заявки (задачи) могут обслуживаться параллельно. В теоретической модели предполагается, что возможности принятия к параллельному обслуживанию не ограничены, но всегда можно оценить вероятность превышения некоторого заданного уровня параллельного обслуживания.

Задание времени обслуживания поступающих заявок будет отличаться от классического подхода, в котором задается закон распределения величины случайного времени обслуживания каждой заявки. В предлагаемой модели скорость обслуживания каждой заявки будет определяться двумя факторами: объемом (размером) поступающей заявки и скоростью (интенсивностью) ее обслуживания. Заметим, что в нашей модели скорость обслуживания не является постоянной величиной и изменяется в зависимости от числа одновременно обслуживаемых задач.

Для целей аналитического исследования модели будем полагать, что объем поступающих заявок будет иметь экспоненциальное распределение с параметром γ. Получается, что средний объем заявки равен γ^{-1} единиц объема. Базовой скоростью обслуживания заявок будем называть скорость обработки заявки в системе. Заявка в системе одна и поэтому не конкурирует с другими заявками за ресурс. Базовую скорость будем обозначать через μ единиц объема за единицу времени. Для снижения громоздкости выкладок будем полагать $\gamma = 1$, что соответствует нормировке единицы измерения объема на средний объем поступающих заявок.

Текущая скорость обслуживания определяется количеством заявок в системе: чем больше заявок одновременно обслуживается, тем выше конкуренция за ресурс и ниже скорость обслуживания. Обозначим случайный процесс $i(t)$ как число заявок в системе в момент времени t. Этот процесс может принимать значения $i = 0,\ 1,\ 2,\ldots$ Каждому значению процесса $i(t)$ будет соответствовать своя скорость обслуживания $\mu(i)$, которая определяется равенством

$$\mu(i) = \mu f(i),$$

где μ – базовая скорость обслуживания, а $f(i)$ – безразмерная функция, которая при целочисленных значениях аргумента определяет коэффициент снижения базовой скорости обслуживания. Функцию $f(i)$ будем называть функцией деградации скорости обслуживания.

В модели предполагается, что моменты обращения заявок к системе формируют стационарный пуассоновский поток (простейший поток) событий, который определяется параметром λ [28]. Эта модель потока обладает полезным свойством – независимостью длин интервалов между моментами наступления событий (поступления заявок), которые распределены по экспоненциальному закону с заданным параметром λ.

Для стабильного функционирования системы необходимо выполнение условия

$$\lambda < \lim_{i \to \infty} \mu i f(i),$$

чтобы при росте числа запросов система не переполнялась.

Теперь логика функционирования модели и все ее параметры определены. Перейдем к формальному описанию объекта исследования и приведем схему предложенной модели.

Объектом исследования является выходящий поток системы, то есть моменты завершения обслуживания заявок. Поскольку поступление заявок и их объем являются случайными, то и моменты окончания обслуживания тоже будут случайными. Соответственно, выходящий поток будет являться случайным

потоком событий. Для задания, анализа таких потоков и вычисления их характеристик используют два подхода: через длины интервалов между моментами наступления событий и через число событий, наступивших в потоке за некоторое время.

В работе используется второй подход описания потока. Для этого вводится стационарный случайный процесс $m(t)$ – число событий выходящего потока рассматриваемой системы за некоторое время t. Целью исследования является нахождение распределения вероятностей этого процесса или его числовых характеристик.

2. ИССЛЕДОВАНИЕ ВЫХОДЯЩЕГО ПОТОКА В СИСТЕМЕ С НЕОГРАНИЧЕННЫМ ЧИСЛОМ ПРИБОРОВ И ДЕГРАДАЦИЕЙ СКОРОСТИ ОБСЛУЖИВАНИЯ

Моменты наступления событий в выходящем потоке определяются текущим состоянием системы. Поэтому для исследования числа событий выходящего потока применяется метод расширения фазового пространства, при котором вводится дополнительный случайный процесс, определяющий текущее состояние системы.

Обозначим $P(i,\ m,\ t) = P\{i(t) = i,\ m(t) = m\}$ – вероятность того, что в момент времени t в системе будет занято i приборов, а в выходящем потоке произойдет m событий к моменту времени t. Здесь предполагается, что в момент начала наблюдения $t = 0$ система функционировала в стационарном режиме. Двумерный процесс $\{i(t),\ m(t)\}$ является марковским, для распределения вероятностей его состояний составим систему дифференциальных уравнений Колмогорова:

$$\frac{\partial P(i,\ m,\ t)}{\partial t} = -\big(\lambda + i\mu(i)\big)P(i,\ m,\ t) +$$
$$+ \lambda P(i-1,\ m,\ t) + (i+1)\mu(i+1)P(i+1,\ m-1,\ t).$$

Полученная система является дифференциально-конечноразностной системой уравнений с переменными коэффициентами. Для ее решения предлагается реализовать метод асимптотического анализа и найти предельное решение системы.

Необходимо сделать переход от распределения вероятностей к характеристическим функциям по переменной m:

$$H(i,\ u,\ t) = \sum_{m=0}^{\infty} e^{jum} P(i,\ m,\ t), \text{ где } j = \sqrt{-1}.$$

При подстановке последнего равенства в систему уравнений Колмогорова и последующем суммировании получим уравнение для функций $H(i,\ u,\ t)$:

$$\frac{\partial H(i,\ u,\ t)}{\partial t} = -\big(\lambda + i\mu(i)\big)H(i,\ u,\ t) + \lambda H(i-1,\ u,\ t) +$$
$$+ e^{ju}(i+1)\mu(i+1)H(i+1,\ u,\ t). \qquad (1)$$

Определим полную характеристическую функцию путем суммирования частичных $H(i,\ u,\ t)$ по всем состояниям i:

$$H(u,\ t) = \sum_{i=0}^{\infty} H(i,\ u,\ t).$$

Просуммировав уравнение (1) по всем i от 0 до ∞, получим уравнение для полной характеристической функции:

$$\frac{\partial H(u,\ t)}{\partial t} = (e^{ju} - 1) \sum_{i=0}^{\infty} i\mu(i) H(i,\ u,\ t). \tag{2}$$

Система уравнений (1) – (2) будет основной в дальнейших исследованиях.

2.1. АСИМПТОТИКА ПЕРВОГО ПОРЯДКА

Для решения системы (1) – (2) вводим предельное условие растущего времени наблюдения за выходящим потоком при определении бесконечно большого параметра T. Вводим следующие замены:

$$T = \frac{1}{\varepsilon},\ \ \tau = \varepsilon t,\ \ u = \varepsilon w,\ \ H(i,\ u,\ t) = F(i,\ w,\ \tau,\ \varepsilon),$$

где ε – малый параметр. Условие $T \to \infty$ будем называть предельным условием растущего времени наблюдения за выходящим потоком. Параметр T определяет зависимость времени t от «медленного времени» τ. Предполагается, что для любого фиксированного значения $\tau > 0$ значение t неограниченно возрастает. Используя введенные замены, уравнения (1) и (2) перепишем в следующем виде:

$$\varepsilon \frac{\partial F(i,\ w,\ \tau,\ \varepsilon)}{\partial \tau} = -\big(\lambda + i\mu(i)\big) F(i,\ w,\ \tau,\ \varepsilon) + \lambda F(i-1,\ w,\ \tau,\ \varepsilon) +$$

$$+ e^{j\varepsilon w}(i+1)\mu(i+1) F(i+1,\ w,\ \tau,\ \varepsilon), \tag{3}$$

$$\varepsilon \frac{\partial F(w,\ \tau,\ \varepsilon)}{\partial \tau} = (e^{j\varepsilon w} - 1) \sum_{i=0}^{\infty} i\mu(i) F(i,\ w,\ \tau,\ \varepsilon), \tag{4}$$

Устремляя в (3) и (4) $\varepsilon \to 0$, получим

$$\lambda F(i-1,\ w,\ \tau) - \big(\lambda + i\mu(i)\big) F(i,\ w,\ \tau) + (i+1)\mu(i+1) F(i+1,\ w,\ \tau) = 0, \tag{5}$$

$$\frac{\partial F(w,\ \tau)}{\partial \tau} = jw \sum_{i=0}^{\infty} i\mu(i) F(i,\ w,\ \tau). \tag{6}$$

Найдем решение уравнения (5) в виде произведения двух функций:

$$F(i,\ w,\ \tau) = R(i)\Phi(w,\ \tau), \tag{7}$$

где $R(i)$ – стационарное распределение вероятностей числа заявок в системе, а функция $\Phi(w, \tau)$ определяет предельное число событий, наступивших в выходящем потоке. При подстановке решения (7) в (5) получим

$$\Phi(w, \tau)\left[\lambda R(i-1) - (\lambda + i\mu(i))R(i) + (i+1)\mu(i+1)R(i+1)\right] = 0.$$

Так как функция $\Phi(w, \tau)$ тождественно не равна нулю, выражение в квадратных скобках должно обращаться в ноль для всех i. Это дает уравнение для стационарного распределения числа заявок $R(i)$ в системе:

$$\lambda R(i-1) - (\lambda + i\mu(i))R(i) + (i+1)\mu(i+1)R(i+1) = 0. \quad (8)$$

Из уравнения (8) можно выразить $R(i)$ через $R(i-1)$. Перенесем часть слагаемых:

$$\lambda R(i-1) - i\mu(i)R(i) = \lambda R(i) - (i+1)\mu(i+1)R(i+1).$$

Разности $\lambda R(i-1) - i\mu(i)R(i)$ должны быть равны между собой для всех i, что означает постоянство величины $C_1 = \lambda R(i-1) - i\mu(i)R(i)$. При $C_1 = 0$ получим

$$\lambda R(i-1) = i\mu(i)R(i), \ i \geq 1, \quad (9)$$

отсюда выразим $R(i)$:

$$R(i) = \frac{\lambda}{i\mu(i)} R(i-1).$$

Применяя эту формулу последовательно, можно записать явное выражение

$$R(i) = R(0)\prod_{n=1}^{i} \frac{\lambda}{n\mu(n)}.$$

Константа $R(0)$ – вероятность того, что в системе нет заявок, – находится из условия нормировки.

После подстановки (7) в (6) получим

$$\frac{\partial \Phi(w, \tau)}{\partial \tau} = jw \sum_{i=0}^{\infty} i\mu(i)R(i)\Phi(w, \tau),$$

Обозначим сумму в правой части как производную некоторой функции $x(\tau)$:

$$x'(\tau) = \sum_{i=0}^{\infty} i\mu(i)R(i).$$

Вычислим эту сумму, используя явный вид $R(i)$:

$$\sum_{i=0}^{\infty} i\mu(i)R(i) = \sum_{i=1}^{\infty} i\mu(i)\frac{\lambda}{i\mu(i)}R(i-1) = \lambda\sum_{i=1}^{\infty} R(i-1) = \lambda, \quad (10)$$

отсюда имеем $x'(\tau) = \lambda,$ следовательно, $x(\tau) = \lambda\tau.$

Таким образом, решение для характеристической функции имеет вид

$$F(i,w,\tau) = R(i)\Phi(w,\tau) = R(i)e^{jwx(\tau)} = R(i)e^{jw\lambda\tau},$$

Вернемся к исходным переменным через обратную замену:

$$H(i,u,t) \approx R(i)e^{ju\lambda t}. \tag{11}$$

Из формулы (11) следует, что среднее число событий в выходящем потоке, наступивших за время t, равно λt. Формулу (11) будем называть асимптотикой первого порядка.

2.2. АСИМПТОТИКА ВТОРОГО ПОРЯДКА

Для уточнения характеристик выходящего потока рассмотрим асимптотику второго порядка. Она позволяет определить предельный закон распределения числа событий выходящего потока.

Введем следующую замену:

$$H(i,\ u,\ t) = H_2(i,\ u,\ t)e^{ju\lambda t}.$$

При подстановке ее в исходные уравнения (1) и (2) получим уравнения для $H_2(i,\ u,\ t)$:

$$\frac{\partial H_2(i,\ u,\ t)}{\partial t} + ju\lambda H_2(i,\ u,\ t) = -\big(\lambda + i\mu(i)\big)H_2(i,\ u,\ t) +$$

$$+ \lambda H_2(i-1,\ u,\ t) + e^{ju}(i+1)\mu(i+1)H_2(i+1,\ u,\ t), \tag{12}$$

$$\frac{\partial H_2(u,\ t)}{\partial t} + ju\lambda H_2(u,\ t) = (e^{ju}-1)\sum_{i=0}^{\infty} i\mu(i)H_2(i,\ u,\ t). \tag{13}$$

Перейдем к предельному условию растущего времени наблюдения за выходящим потоком и введем следующие замены:

$$T = \frac{1}{\varepsilon^2},\ \ \tau = \varepsilon^2 t,\ \ u = \varepsilon w,\ \ H_2(i,\ u,\ t) = F_2(i,\ w,\ \tau,\ \varepsilon).$$

Перепишем уравнения (12) и (13) с учетом замены:

$$\varepsilon^2\frac{\partial F_2(i,\ w,\ \tau,\ \varepsilon)}{\partial\tau} + j\varepsilon w\lambda F_2(i,\ w,\ \tau,\ \varepsilon) = -\big(\lambda + i\mu(i)\big)F_2(i,\ w,\ \tau,\ \varepsilon) +$$

$$+ \lambda F_2(i-1,\ w,\ \tau,\ \varepsilon) + e^{j\varepsilon w}(i+1)\mu(i+1)F_2(i+1,\ w,\ \tau,\ \varepsilon), \tag{14}$$

$$\varepsilon^2\frac{\partial F_2(w,\ \tau,\ \varepsilon)}{\partial\tau} + j\varepsilon w\lambda F_2(w,\ \tau,\ \varepsilon) = (e^{j\varepsilon w}-1)\sum_{i=0}^{\infty} i\mu(i)F_2(i,\ w,\ \tau,\ \varepsilon).$$

Запишем первое уравнение системы (14) с точностью до $O(\varepsilon^2)$:

$$j\varepsilon w\lambda F_2(i,\ w,\ \tau,\ \varepsilon) = -(\lambda + i\mu(i))F_2(i,\ w,\ \tau,\ \varepsilon) + \lambda F_2(i-1,\ w,\ \tau,\ \varepsilon) +$$

$$+ (1 + j\varepsilon w)(i+1)\mu(i+1)F_2(i+1,\ w,\ \tau,\ \varepsilon) + O(\varepsilon^2). \qquad (15)$$

Решение $F(i,\ w,\ \tau,\ \varepsilon)$ будем искать в виде разложения:

$$F_2(i,\ w,\ \tau,\ \varepsilon) = \Phi_2(w,\ \tau)\{R(i) + j\varepsilon w g(i)\} + O(\varepsilon^2). \qquad (16)$$

Перепишем уравнение (15) с учетом равенства (16):

$$j\varepsilon w\lambda R(i) = -(\lambda + i\mu(i))\{R(i) + j\varepsilon w g(i)\} + \lambda\{R(i-1) + j\varepsilon w g(i-1)\} +$$

$$+ (i+1)\mu(i+1)\{R(i+1) + j\varepsilon w g(i+1)\} + j\varepsilon w(i+1)\mu(i+1)R(i+1) + O(\varepsilon^2), \qquad (17)$$

Разделим уравнение (17) на $jw\varepsilon$ и устремим $\varepsilon \to 0$:

$$\lambda R(i) = -(\lambda + i\mu(i))g(i) + \lambda f(i-1) + (i+1)\mu(i+1)g(i+1) + (i+1)\mu(i+1)R(i+1),$$

Выполняя преобразования с учетом (9), запишем

$$\lambda g(i-1) - (\lambda + i\mu(i))g(i) + (i+1)\mu(i+1)g(i+1) = 0. \qquad (18)$$

Анализируя уравнение (18), можно заметить, что его структура совпадает с общим видом уравнений, описывающих стационарное распределение вероятностей числа занятых приборов в системах массового обслуживания. Из этого следует, что распределение вероятностей числа занятых приборов выступает в качестве мультипликативного сомножителя в исследуемой модели, то есть отделяется от остальных компонент системы и входит в итоговое выражение в виде независимого коэффициента:

$$g(i) = C_2 R(i), \qquad (19)$$

где C_2 – константа.

Запишем согласованное уравнение системы (14) с точностью до $O(\varepsilon^2)$:

$$\varepsilon^2 \frac{\partial F_2(w,\ \tau,\ \varepsilon)}{\partial \tau} + j\varepsilon w\lambda F_2(w,\ \tau,\ \varepsilon) =$$

$$= \left[j\varepsilon w + \frac{(j\varepsilon w)^2}{2} \right] \sum_{i=0}^{\infty} i\mu(i)F_2(i,\ w,\ \tau,\ \varepsilon) + O(\varepsilon^2).$$

Решение будем искать в виде (16):

$$\varepsilon^2 \frac{\partial \Phi_2(w,\ \tau)}{\partial \tau} + j\varepsilon w\lambda \Phi_2(w,\ \tau)\left\{1 + j\varepsilon w \sum_{i=0}^{\infty} g(i)\right\} =$$

$$= j\varepsilon w\Phi_2(w,\ \tau)\sum_{i=0}^{\infty} i\mu(i)\{R(i) + j\varepsilon w g(i)\} + \frac{(j\varepsilon w)^2}{2}\sum_{i=0}^{\infty} i\mu(i)\Phi_2(w,\ \tau)R(i) + O(\varepsilon^2).$$

Разделим уравнение на ε^2 и устремим $\varepsilon \to 0$. Учитывая (10), запишем

$$\frac{\partial \Phi_2(w,\ \tau)}{\partial \tau} + (jw)^2 \lambda \Phi_2(w,\ \tau) \sum_{i=0}^{\infty} g(i) =$$

$$= (jw)^2 \Phi_2(w,\ \tau) \sum_{i=0}^{\infty} i\mu(i)\, g(i) + \frac{(jw)^2}{2} \Phi_2(w,\ \tau) \sum_{i=0}^{\infty} i\mu(i) R(i). \qquad (20)$$

Используя замену (19) и подставив ее в (20), получим уравнение

$$\frac{\partial \Phi_2(w,\ \tau)}{\partial \tau} + (jw)^2 \lambda \Phi_2(w,\ \tau) C_2 =$$

$$= (jw)^2 \Phi_2(w,\ \tau) C_2 \sum_{i=0}^{\infty} i\mu(i)\, R(i) + \frac{(jw)^2}{2} \lambda \Phi_2(w,\ \tau).$$

Применим преобразования и получим дифференциальное уравнение для $\Phi_2(w,\ \tau)$:

$$\frac{\partial \Phi_2(w,\ \tau)}{\partial \tau} = \frac{(jw)^2}{2} \lambda \Phi_2(w,\ \tau).$$

Это уравнение является уравнением с разделяющимися переменными, решение которого имеет вид

$$\Phi_2(w,\ \tau) = \exp\left\{ \frac{(jw)^2}{2} \lambda \tau \right\}.$$

Выполним обратный переход к исходным функциям:

$$H_2(u,\ t) = \exp\left\{ \frac{(ju)^2}{2} \lambda t \right\}.$$

Вернемся к замене для $H(i,\ u,\ t)$:

$$H(i,\ u,\ t) = H_2(i,\ u,\ t) e^{ju\lambda t}.$$

Для $H(u,\ t)$ запишем

$$H(u,\ t) = \sum_{i=0}^{\infty} H(i,\ u,\ t) = e^{iu\lambda t} \sum_{i=0}^{\infty} H_2(i,\ u,\ t) =$$

$$= e^{iu\lambda t} H(u,\ t) = \exp\left\{ ju\lambda t + \frac{(ju)^2}{2} \lambda t \right\}. \qquad (21)$$

Равенство (21) – это формула характеристической функции нормального распределения $N(m, \sigma^2)$, где

$$m = \lambda t, \quad \sigma = \sqrt{\lambda t}.$$

Полученное решение (21) показывает, что распределение числа обслуженных заявок за некоторое время в системе массового обслуживания с неограниченным числом приборов и интенсивностью обслуживания, зависящей от числа заявок в системе, является асимптотически нормальным, среднее и дисперсия которого совпадают и пропорциональны интенсивности входящего потока. Заметим, что таким же свойством обладает поток Пуассона, поэтому полученный асимптотический выходящий поток будем называть квазипуассоновским.

2.3. РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ЭКСПЕРИМЕНТОВ

Приведем несколько иллюстрационных численных примеров, которые были сделаны на основе результатов имитационного моделирования.

Сначала сравним полученные результаты асимптотического анализа (*asim*) и имитационного моделирования (*simul*) при помощи расстояния Колмогорова для функций распределения:

$$\Delta = \max \left| F(x)_{asimp} - F(x)_{simul} \right|.$$

Результаты сравнения для двух видов функции деградации и разных значений времени наблюдения T представлены в табл. 1.

Таблица 1

Table 1

Расстояние Колмогорова между имитационным и асимптотическим распределением вероятностей числа событий в выходящем потоке ($\lambda = 0{,}7$; $\mu = 1$)

Kolmogorov distance between simulated and asymptotic probability distributions of the number of requests in the outgoing flow ($\lambda = 0{,}7$; $\mu = 1$)

Вид функции деградации	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 10$
$f(i) = \frac{1}{\sqrt{1+i}}$	0,078	0,051	0,035	0,030	0,024
$f(i) = \frac{1}{i}$	0,074	0,052	0,033	0,031	0,023

Заметим, что большие отклонения при небольших значениях T наблюдались именно в начале распределения, что частично объясняется смещением моды распределения ближе к нулю. В этом случае усеченное нормальное распределение не очень хорошо аппроксимирует допредельное распределение, даже если оно является пуассоновским.

Из таблицы видно, что гауссовским распределением с равными математическим ожиданием и дисперсией можно аппроксимировать распределение

числа событий в выходящем потоке рассматриваемой модели, когда среднее число событий за интервал 5 или более.

Дополнительно были проведены эксперименты, когда в имитационной модели рассматривалось не только экспоненциальное распределение для объема заявок, но также равномерное и гамма-распределение.

Таблица 2

Table 2

Расстояние Колмогорова между имитационным и асимптотическим распределением вероятностей числа событий в выходящем потоке для разных функций деградации и закона распределения объема заявок ($\lambda = 0{,}7$; $\mu = 1$; $T = 10$)

Kolmogorov distance between the simulated and asymptotic probability distributions of the number of requests in the outgoing flow for different degradation functions and request volume distributions ($\lambda = 0{,}7$; $\mu = 1$; $T = 10$)

Распределение объема заявок	$f(i) = \frac{1}{\sqrt{1+i}}$	$f(i) = \frac{1}{i}$
Экспоненциальное ($\mu = 1$)	0,024	0,023
Равномерное ($a = 0$, $b = 2$)	0,025	0,018
Гамма ($\alpha = \beta = 0{,}5$)	0,024	0,015

По результатам численных экспериментов, представленных в табл. 2, можно наблюдать отсутствие чувствительности результата к виду распределения объема заявок. Такой результат наблюдался для случая экспоненциального, равномерного и гамма-распределения с одинаковыми средними, равными единице.

ЗАКЛЮЧЕНИЕ

В работе рассмотрена модель вычислительного узла в виде системы массового обслуживания с неограниченным числом приборов и зависимостью скорости обслуживания от общего числа заявок в системе. Объект исследования – выходящий поток системы. В результате применения метода асимптотического анализа получены характеристики распределения числа событий, закончивших обслуживание за некоторое время t, которые и образуют события выходящего потока. Для предложенной модели удалось показать, что математическое ожидание и дисперсия числа совпадают. В этом случае выходящий поток можно называть квазипуассоновским.

С практической точки зрения это означает, что, несмотря на зависимость скорости обработки запросов от общего числа запросов в системе, характеристики потока обработанных запросов близки к характеристикам потока поступления запросов. Также важно отметить, что характеристики выходящего потока рассматриваемой системы не зависят от типа зависимости скорости обслуживания $f(i)$, что видно из итоговой формулы (21).

Проведенные численные эксперименты показали, что предложенную гауссовскую аппроксимацию можно применять, когда среднее число событий

в выходящем потоке за интервал T равно пяти или более. При равномерном и гамма-распределении объема заявок распределение числа событий в выходящем потоке почти не отличается от случая экспоненциального распределения объема заявок.

Предложенный в работе подход для исследования выходящих потоков систем с неограниченным числом приборов и деградацией скорости обслуживания может быть применен для исследования моделей с более сложными моделями входящего потока (MMPP, MAP, рекуррентный), который характеризует поступление запросов на вычислительный узел.

СПИСОК ЛИТЕРАТУРЫ

1. Analysis of the performance of a cloud computing processing queue with correlated reneging of tasks and resubmission / G.S. Kuaban, B.S. Soodan, R. Kumar, P. Czekalski // 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET). – IEEE, 2021. – P. 1–8.

2. *Hanini M.*, *Kafhali S.E.*, *Salah K.* Dynamic VM allocation and traffic control to manage QoS and energy consumption in cloud computing environment // International Journal of Computer Applications in Technology. – 2019. – Vol. 60 (4). – P. 307–316.

3. *Heidelberger P.*, *Lavenberg S.S.* Computer performance evaluation methodology // IEEE Transactions on Computers. – 1984. – Vol. 100 (12). – P. 1195–1220.

4. *Vetha S.*, *Devi K.V.* Dynamic resource allocation in cloud using queuing model // Journal of Industrial Pollution Control. – 2017. – Vol. 33 (2). – P. 1547–1554.

5. A queuing theory model for cloud computing / J. Vilaplana, F. Solsona, I. Teixidó, J. Mateo, F. Abella, J. Rius // The Journal of Supercomputing. – 2014. – Vol. 69 (1). – P. 492–507.

6. *Liu X.*, *Li S.*, *Tong W.* A queuing model considering resources sharing for cloud service performance // The Journal of Supercomputing. – 2015. – Vol. 71 (11). – P. 4042–4055.

7. *Вишневский В.М.*, *Дудин А.Н.*, *Клименок В.И.* Стохастические системы с коррелированными потоками: теория и применение в телекоммуникационных сетях. – М.: Техносфера, 2018. – 564 с.

8. Queueing system with two phases of service and service rate degradation / E. Fedorova, I. Lapatin, O. Lizyura, A. Moiseev, A. Nazarov, S. Paul // Axioms. – 2023. – Vol. 12 (2). – P. 104.

9. *Bermejo B.*, *Juiz C.* A general method for evaluating the overhead when consolidating servers: performance degradation in virtual machines and containers // The Journal of Supercomputing. – 2022. – Vol. 78. – P. 11345–11372.

10. *Kleinrock L.* Queueing systems. Vol. 1. – New York: Wiley, 1975.

11. *Kelly F.P.* Networks of queues // Advances in Applied Probability. – 1976. – Vol. 8 (2). – P. 416–432.

12. *Cohen J.W.* The multiple phase service network with generalized processor sharing // Acta Informatica. – 1979. – Vol. 12. – P. 245–284.

13. *Guillemin F.*, *Robert P.*, *Zwart B.* Tail asymptotics for processor-sharing queues // Advances in Applied Probability. – 2004. – Vol. 36 (2). – P. 525–543.

14. *Zwart A.P.*, *Boxma O.J.* Sojourn time asymptotics in the M/G/1 processor sharing queue. – Centrum voor Wiskunde en Informatica, 1998. – 27 p. – (CWI report: Probability, Networks and Algorithms; vol. 9802).

15. *Alsadie D*. A comprehensive review of AI Techniques for resource management in fog computing: Trends, challenges, and future directions // IEEE Access. – 2024. – Vol. 12. – P. 118007–118059.

16. *Ibrahim A.S., Al-Mahdi H., Nassar H.* Characterization of task response time in a fog-enabled IoT network using queueing models with general service times. Journal of King Saud University // Computer and Information Sciences. – 2022. – Vol. 34 (9). – P. 7089–7100.

17. Asymptotic analysis of two-phase queueing system with service rate degradation and heterogeneous customers / E.A. Fedorova, I.L. Lapatin, O.D. Lizyura, A.N. Moiseev, A. Nazarov, S. Paul // 2023 5th International Conference on Problems of Cybernetics and Informatics (PCI), 28–30 august 2023. – IEEE, 2023. – P. 1–5.

18. Mathematical model of cloud node using closed queueing system with service rate degradation / A.A. Nazarov, A.N. Moiseev, S.V. Paul, I.L. Lapatin, E. Fedorova, O. Lizyura, R. Salimzyanov,

D. Salimzyanova // 21st International Asian School-Seminar on Optimization Problems of Complex Systems (OPCS). – IEEE, 2025. – P. 11219374.

19. *Fedorova E.A.*, *Nazarov A.A.*, *Nikolaeva D.I.* Asymptotic analysis of sojourn time in retrial queueing system with non-persistent customers and feedback // Communications in Computer and Information Science. – 2025. – Vol. 2472. – P. 98–107.

20. *Nazarov A.A., Fedorova E.A.* Method of the marginal asymptotic-diffusion analysis for multiclass retrial queue Mn/GIn/1 // Automation and Remote Control. – 2025. – Vol. 86 (3). – P. 235–248.

21. Performance modeling of cloud systems by an infinite-server queue operating in rarely changing random environment / S. Moiseeva, E. Polin, A. Moiseev, J. Sztrik // Future Internet. – 2025. – Vol. 17 (10). – P. 462.

22. *Daley D.J.* Notes on queueing output processes // Mathematical methods in queueing theory / ed. by A.B. Clarke. – Berlin; Heidelberg: Springer, 1974. – (Lecture Notes in Economics and Mathematical Systems; vol. 98). – P. 211–262.

23. *Daley D.J.* Queueing output processes // Advances in Applied Probability. – 1976. – Vol. 8 (2). – P. 395–415.

24. *Burke P.J.* The output of a queuing system // Operations Research. – 1956. – Vol. 4 (6). – P. 699–704.

25. *Mirasol N.M.* Letter to the editor – The output of an M / G /(infinity) queuing system is Poisson // Operations Research. – 1963. – Vol. 11 (2). – P. 282–284.

26. Matrix and analytical methods for performance analysis of telecommunication systems / V. Naumov, Y. Gaidamaka, N. Yarkina, K. Samouylov. – Berlin; Heidelberg: Springer, 2021. – 305 p. – DOI: 10.1007/978-3-030-83132-5.

27. *Моисеев А.Н.*, *Назаров А.А.* Бесконечнолинейные системы и сети массового обслуживания. – Томск: Изд-во НТЛ, 2015. – 240 с.

28. *Назаров А.А.*, *Терпугов А.Ф.* Теория массового обслуживания: учебное пособие. – 2-е изд., испр. – Томск: Изд-во НТЛ, 2010. – 228 с.

Лапатин Иван Леонидович, кандидат физико-математических наук, доцент кафедры теории вероятностей и математической статистики Национального исследовательского Томского государственного университета. Основное направление научных исследований – теория массового обслуживания, теория случайных процессов, математическая теория телетрафика. Имеет более 50 печатных работ. E-mail: ilapatin@mail.ru

Назаров Анатолий Андреевич, доктор технических наук, профессор кафедры теории вероятностей и математической статистики Национального исследовательского Томского государственного университета. Основное направление научных исследований – теория массового обслуживания, теория случайных процессов, математическая теория телетрафика. Имеет более 700 печатных работ. E-mail: nazarov.tsu@gmail.com

Пауль Светлана Владимировна, доктор физико-математических наук, профессор кафедры теории вероятностей и математической статистики Национального исследовательского Томского государственного университета. Основное направление научных исследований – теория массового обслуживания, теория случайных процессов, математическая теория телетрафика. Имеет более 100 печатных работ. E-mail: paulsv82@mail.ru

Салимзянова Дарья Дмитриевна, ассистент кафедры теории вероятностей и математической статистики Национального исследовательского Томского государственного университета. Основное направление научных исследований – математическое моделирование, теория массового обслуживания, теория случайных процессов, математическая теория телетрафика. Имеет более 10 печатных работ. E-mail: darya2001@inbox.ru

Lapatin Ivan L., PhD (Phys.-Math.), associate professor at the Department of Probability Theory and Mathematical Statistics, National Research Tomsk State University. His research interests are currently focused on the queueing theory, the theory of stochastic processes, and the mathematical teletraffic theory. He has more than 50 publications. E-mail: ilapatin@mail.ru

Nazarov Anatoly A., D. Sc. (Eng.), professor at the Department of Probability Theory and Mathematical Statistics, National Research Tomsk State University. His research interests are currently focused on the queueing theory, the theory of stochastic processes, and the mathematical teletraffic theory. He has more than 700 publications. E-mail: nazarov.tsu@gmail.com

Paul Svetlana V., D. Sc. (Phys.&Math.), professor at the Department of Probability Theory and Mathematical Statistics, National Research Tomsk State University. Her research interests are currently focused on the queueing theory, the theory of stochastic processes, and the mathematical teletraffic theory. She has more than 100 publications. E-mail: paulsv82@mail.ru

Salimzyanova Daria D., assistant lecturer at the Department of Probability Theory and Mathematical Statistics, National Research Tomsk State University. Her research interests are currently focused on mathematical modeling, the queueing theory, the theory of stochastic processes, and the mathematical teletraffic theory. She has more than 10 publications. E-mail: darya2001@inbox.ru

DOI: 10.17212/2782-2001-2026-3-151-168

Characteristics of the output flow of processed tasks in a computational node model under resource competition[*]

I.L. LAPATIN [a], A.A. NAZAROV [b], S.V. PAUL [c], D.D. SALIMZYANOVA [d]

National Research Tomsk State University, 36 Lenin Avenue, Tomsk, 634050, Russian Federation

[a] ilapatin@mail.ru [b] nazarov.tsu@gmail.com [c] paulsv82@mail.ru [d] darya2001@inbox.ru

Abstract

The continuous growth in the popularity of remote and cloud computing creates the need for developing mathematical models describing the operation of such systems or their individual components. Mathematical models make it possible to evaluate various performance characteristics of a system under different operating parameters. The calculation of these characteristics is required at the design stage of a computational node or for optimizing the operation of an existing one. The queueing theory is widely used to model processes of a stochastic nature.

This paper considers a computational node model represented as a queueing system with an unlimited number of servers. The arrival of tasks to the node is described by a Poisson point process. A distinctive feature of the proposed model compared to the classical one is the dependence of the service (processing) rate of incoming tasks on the total number of tasks present at the computational node. The greater the number of tasks being processed simultaneously, the lower the processing rate of each individual task. This effect arises due to competition for a shared resource of the computational node.

In this study, the object of investigation is the output flow of the model, i.e., the moments when task is completed on the computational node. Information about the characteristics of the output flow is of great importance since after processing, tasks (requests) must be transmitted via telecommunication channels. It is the probabilistic characteristics of the serviced task flow that affect the calculation of the required data transmission bandwidth to ensure specified quality-of-service parameters.

To analyze the output flow of the computational node model, asymptotic methods of queueing theory are employed. As a result, asymptotic characteristics of the output flow of the proposed model are obtained. It is important to note that the proposed asymptotic approach can potentially be extended to models with non-Poisson task arrival processes.

Keywords: data transmission, computing systems, computational node modeling, resource competition, queueing theory, output flow, asymptotic analysis method, Poisson flow

[*] *Received 03 February 2026.*

REFERENCES

1. Kuaban G.S., Soodan B.S., Kumar R., Czekalski P. Analysis of the performance of a cloud computing processing queue with correlated reneging of tasks and resubmission. *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET)*. IEEE, 2021, pp. 1–8.

2. Hanini M., Kafhali S.E., Salah K. Dynamic VM allocation and traffic control to manage QoS and energy consumption in cloud computing environment. *International Journal of Computer Applications in Technology*, 2019, vol. 60 (4), pp. 307–316.

3. Heidelberger P., Lavenberg S.S. Computer performance evaluation methodology. *IEEE Transactions on Computers*, 1984, vol. 100 (12), pp. 1195–1220.

4. Vetha S., Devi K.V. Dynamic resource allocation in cloud using queuing model. *Journal of Industrial Pollution Control*, 2017, vol. 33 (2), pp. 1547–1554.

5. Vilaplana J., Solsona F., Teixidó I., Mateo J., Abella F., Rius J. A queuing theory model for cloud computing. *The Journal of Supercomputing*, 2014, vol. 69 (1), pp. 492–507.

6. Liu X., Li S., Tong W. A queuing model considering resources sharing for cloud service performance. *The Journal of Supercomputing*, 2015, vol. 71 (11), pp. 4042–4055.

7. Vishnevskii V.M., Dudin A.N., Klimenok V.I. *Stokhasticheskie sistemy s korrelirovannymi potokami: teoriya i primenenie v telekommunikatsionnykh setyakh* [Stochastic systems with correlated flows: Theory and application in telecommunication networks]. Moscow, Tekhnosfera Publ., 2018. 564 p.

8. Fedorova E., Lapatin I., Lizyura O., Moiseev A., Nazarov A., Pau S. Queueing system with two phases of service and service rate degradation. *Axioms*, 2023, vol. 12 (2), p. 104.

9. Bermejo B., Juiz C. A general method for evaluating the overhead when consolidating servers: performance degradation in virtual machines and containers. *The Journal of Supercomputing*, 2022, vol. 78, pp. 11345–11372.

10. Kleinrock L. *Queueing systems*. Vol. 1. New York, Wiley, 1975.

11. Kelly F.P. Networks of queues. *Advances in Applied Probability*, 1976, vol. 8 (2), pp. 416–432.

12. Cohen J.W. The multiple phase service network with generalized processor sharing. *Acta Informatica*, 1979, vol. 12, pp. 245–284.

13. Guillemin F., Robert P., Zwart B. Tail asymptotics for processor-sharing queues. *Advances in Applied Probability*, 2004, vol. 36 (2), pp. 525–543.

14. Zwart A.P., Boxma O.J. *Sojourn time asymptotics in the M/G/1 processor sharing queue*. *CWI report: Probability, Networks and Algorithms*, vol. 9802. Centrum voor Wiskunde en Informatica, 1998. 27 p.

15. Alsadie D. A comprehensive review of AI Techniques for resource management in fog computing: Trends, challenges, and future directions. *IEEE Access*, 2024, vol. 12, pp. 118007–118059.

16. Ibrahim A.S., Al-Mahdi H., Nassar H. Characterization of task response time in a fog-enabled IoT network using queueing models with general service times. *Journal of King Saud University – Computer and Information Sciences*, 2022, vol. 34 (9), pp. 7089–7100.

17. Fedorova E.A., Lapatin I.L., Lizyura O.D., Moiseev A.N., Nazarov A., Paul S. Asymptotic analysis of two-phase queueing system with service rate degradation and heterogeneous customers. *2023 5th International Conference on Problems of Cybernetics and Informatics (PCI)*, 28–30 august 2023, pp. 1–5.

18. Nazarov A., Moiseev A., Paul S., Lapatin I., Fedorova E., Lizyura O., Salimzyanov R., Salimzyanova D. Mathematical model of cloud node using closed queueing system with service rate degradation. *21st International Asian School-Seminar on Optimization Problems of Complex Systems (OPCS)*. IEEE, 2025, p. 11219374.

19. Fedorova E.A., Nazarov A.A., Nikolaeva D.I. Asymptotic analysis of sojourn time in retrial queueing system with non-persistent customers and feedback. *Communications in Computer and Information Science*, 2025, vol. 2472, pp. 98–107.

20. Nazarov A.A., Fedorova E.A. Method of the marginal asymptotic-diffusion analysis for multiclass retrial queue Mn/GIn/1. *Automation and Remote Control*, 2025, vol. 86 (3), pp. 235–248.

21. Moiseeva S., Polin E., Moiseev A., Sztrik J. Performance modeling of cloud systems by an infinite-server queue operating in rarely changing random environment. *Future Internet*, 2025, vol. 17 (10), p. 462.

22. Daley D.J. Notes on queueing output processes. *Mathematical methods in queueing theory*. *Lecture Notes in Economics and Mathematical Systems*, vol. 98. Ed. by A.B. Clarke. Berlin, Heidelberg, Springer, 1974, pp. 211–262.

23. Daley D.J. Queueing output processes. *Advances in Applied Probability*, 1976, vol. 8 (2), pp. 395–415.

24. Burke P.J. The output of a queuing system. *Operations Research*, 1956, vol. 4 (6), pp. 699–704.

25. Mirasol N.M. Letter to the editor – The output of an M / G /(infinity) queuing system is Poisson. *Operations Research*, 1963, vol. 11 (2), pp. 282–284.

26. Naumov V., Gaidamaka Y., Yarkina N., Samouylov K. *Matrix and analytical methods for performance analysis of telecommunication systems*. Berlin, Heidelberg, Springer, 2021. 305 p. DOI: 10.1007/978-3-030-83132-5.

27. Moiseev A.N., Nazarov A.A. *Beskonechnolineinye sistemy i seti massovogo obsluzhivaniya* [Infinite linear systems and queuing networks]. Tomsk, NTL Publ., 2015. 240 p.

28. Nazarov A.A., Terpugov A.F. *Teoriya massovogo obsluzhivaniya* [Queueing theory]. 2nd ed., rev. Tomsk, NTL Publ., 2010. 228 p.

Для цитирования:

Характеристики потока обработанных задач модели вычислительного узла при конкуренции за ресурс / И.Л. Лапатин, А.А. Назаров, С.В. Пауль, Д.Д. Салимзянова // Системы анализа и обработки данных. – 2026. – № 3 (103). – С. 151–168. – DOI: 10.17212/2782-2001-2026-3-151-168.

For citation:

Lapatin I.L., Nazarov A.A., Paul S.V., Salimzyanova D.D. Kharakteristiki potoka obrabotannykh zadach modeli vychislitel'nogo uzla pri konkurentsii za resurs [Characteristics of the output flow of processed tasks in a computational node model under resource competition]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2026, no. 3 (103), pp. 151–168. DOI: 10.17212/2782-2001-2026-3-151-168.

Правила для авторов

Журнал «Научный вестник НГТУ» основан в 1995 году на базе Новосибирского государственного технического университета. Печатная версия журнала зарегистрирована в Министерстве РФ по делам печати, телевещания и средств массовых коммуникаций в 2000 году. Свидетельство о регистрации ПИ № 77-1599 от 10.02.2000. Периодичность выхода издания – один раз в три месяца (четыре номера в год). В 2021 году был переименован в «Системы анализа и обработки данных». Научно-технические статьи, направленные в адрес журнала, проходят рецензирование и редактирование.

В журнале «Системы анализа и обработки данных» публикуются оригинальные статьи по следующим группам специальностей:

1.2.2 – Математическое моделирование, численные методы и комплексы программ (технические науки), *ранее* 05.13.18;

2.2.6 – Оптические и оптико-электронные приборы и комплексы (технические науки), *ранее* 05.11.07;

2.2.11 – Информационно-измерительные и управляющие системы (технические науки), *ранее* 05.11.16;

2.2.12 – Приборы, системы и изделия медицинского назначения (технические науки), *ранее* 05.11.17;

2.3.1 – Системный анализ, управление и обработка информации, статистика (технические науки), *ранее* 05.13.01;

2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки), *ранее* 05.13.11.

Разделы журнала: информационные технологии и телекоммуникации; компьютерные науки и информатика; электроника, фотоника, приборостроение и связь.

Рекомендуемый объем статьи 10–16 страниц. Статьи объемом, не превышающим 8 страниц, могут быть размещены в разделе «Краткие сообщения». Возможна публикация дискуссионного материала в разделе «Дискуссии». В информационном разделе возможно размещение сообщений о состоявшихся и планируемых конференциях, семинарах и симпозиумах, сведений о факультетах и специальностях, научных разработках и научных школах. Статьи обзорного характера, как правило, не публикуются.

Website: http://journals.nstu.ru/vestnik

Условия приема статей

В редакцию журнала представляются следующие материалы.

1. Статья, подготовленная в соответствии с правилами оформления, – печатная версия, два экземпляра, подписанных авторами.

2. Контактная информация (телефоны рабочий и сотовый, адреса электронной почты, место работы, адрес места работы, должность, ученая степень, ученое звание автора, ORCID) – печатная версия, два экземпляра.

3. Описание статьи для базы данных «Российский индекс научного цитирования (РИНЦ)», подготовленное в соответствии с правилами оформления, – печатная версия, один экземпляр.

4. Лицензионный договор, заполненный и подписанный, отправляется со всем пакетом документов; сканированный заполненный договор записывается на диск.

5. Электронная версия статьи, контактной информации, описания статьи для базы данных РИНЦ и сканированный лицензионный договор – в отдельных файлах.

6. Экспертное заключение о возможности опубликования.

Плата за публикацию рукописей не взимается.

Требования к оформлению статьи

Формат оригинала **– A4**. Шрифт **– Times New Roman**. Текст набирается в русифицированном редакторе **Microsoft Word**. Параметры страницы (см):

левое поле – 3,3 верхнее поле – 4,0

правое поле – 3,3 нижнее поле – 3,9.

текстовое поле 22 × 14,4

Межстрочные интервалы одинарные. Перед УДК оставляется две пустые строки размером 12, УДК набирается прописными буквами кеглем 9. После УДК перед заголовком оставляется пустая строка размером 9. Название статьи набирается кеглем 14 полужирным строчным шрифтом с отступом слева 1,8 см, без переносов слов, с выравниванием влево. После заголовка оставляется пустая строка размером 12. Фамилии автора – прописным шрифтом кеглем 9. После фамилий пустая строка размером 7. Указатель сноски * после названия статьи устанавливается в верхнем индексе. В сноске дата поступления статьи указывается курсивным шрифтом кеглем 9. После ФИО авторов статьи набирается аннотация кеглем 9 с отступом первой строки на 0,8 см, межстрочный интервал с множителем 1,1. После аннотации оставляется пустая строка размером 4. **Ключевые слова** – кеглем 9.

Заголовки разделов обязательны (например, **Введение**, **Постановка задачи**, **Методы анализа**, **Основной результат**, **Выводы** или **Заключение**, **Список литературы**), набираются прописным полужирным прямым шрифтом кеглем 11 с отступом 1,8 см. Перед ними оставляется пустая строка размером 20, после них – пустая строка размером 10. Основной текст – кеглем 11.

Подзаголовки набираются прописным полужирным шрифтом кеглем 10 с отступом 1,8 см. Перед ними и после них – пустая строка размером 10.

Рисунки представляются в **Corel Draw 11** или **Microsoft Word (RUS)**, а также в формате TIFF, JPEG, PNG. Слово *Рис*. и его номер выделяются курсивом. Рисунки должны быть черно-белыми, в описании не должно быть ссылок на цвет.

Рисунки и таблицы нумеруются арабскими цифрами по порядку упоминания их в тексте. После номера должно следовать название **на русском и английском языке**. Подрисуночные подписи и названия таблиц – кеглем 10.

Все графы в таблицах должны иметь заголовки и разделяться вертикальными линиями. В головке таблицы по возможности указываются буквенные обозначения параметров и их единицы измерения.

Переменные в тексте набираются курсивным шрифтом кеглем 11. При написании формул использовать MathType версии не выше 6.0C. Формулы набирать по центру кеглем 11, надстрочные и подстрочные индексы – кеглем 9, знаки – кеглем 14 (латинские буквы – курсивным шрифтом, русские и греческие – прямым, цифры – прямым).

Например:

$$A(s) = \sum_{i=0}^{n^2+3} A_i s\lambda^{(i+1)^2} + \alpha .$$

Между текстом и формулами, а также между формулами оставляются пустые строки размером 4. Формулы не сжимать. Нумерацию формул указывать справа кеглем 11. Например:

$$D_f = D_c D + N_c N. \qquad (1)$$

Нумерация страниц осуществляется следующим образом: Вставка / Номера страниц...; установить **положение** «вверху страницы», **выравнивание** – «снаружи», без нумерации первой страницы.

Колонтитулы устанавливаются вверху (с отступом от верхнего края страницы 3 см). Чтобы колонтитулы на четной, нечетной и первой страницах отличались, нужно установить: Файл / Параметры страницы / Разметка; установить **разные четные и нечетные** и **первая отличается**.

Колонтитулы на четных страницах. Вид / Колонтитулы; набираются курсивным прописным шрифтом кеглем 8. Фамилия автора с инициалами указывается у правого края без отступа. Если авторов несколько, то указываются только два первых, далее «и др.».

Колонтитулы на нечетных страницах. Название статьи с прописной буквы курсивным шрифтом кеглем 8, номера страниц – кеглем 10.

Перед списком литературы пустая строка размером 20. После списка литературы пустая строка размером 10. Порядковый номер с красной строки кеглем 9 с интервалом 1,1. Фамилии и инициалы – кеглем 9 полужирным строчным шрифтом, название статьи (книги) – светлым шрифтом. Оформление библиографического списка по ГОСТ Р.7.0.5–2008. Библиографическая ссылка. Общие требования и правила составления. – М.: Стандартинформ, 2008. – 19 с.

После списка литературы идет пустая строка размером 24. Далее приводятся краткие сведения об авторах кеглем 9 с интервалом 1,1 и отступом от левого края 1 см. Фамилия, имя, отчество – с красной строки курсивом. Последующий текст (ученая степень, ученое звание, членство в академиях и т. д.; основное направление научных исследований; количество имеющихся публикаций, в том числе монографий, учебных пособий; e-mail) – прямым шрифтом. **Информация предоставляется на русском и английском языке**.

Далее дается перевод на английский язык фамилии, имени, отчества автора и названия статьи, а также аннотации и ключевых слов (обязательно).

Исправления, вносимые автором, при доработке статьи **выделять** желтым цветом.

Требования к оформлению РИНЦ

В описании статьи должна быть указана следующая информация.

1. Номера страниц, на которых размещена статья в сборнике.
2. Коды: УДК, и / или ББК, и / или DOI, и / или других классификационных индексов или систем регистрации.
3. Название статьи.
4. Фамилия, имя, отчество авторов.
5. Место работы каждого автора в именительном падеже. Указать должность автора по основному месту работы.
6. Контактная информация (почтовый адрес места работы, номер телефона, адрес электронной почты – обязательно).
7. Аннотация (не менее 320 и не более 640 знаков).
8. Ключевые слова: каждое слово или словосочетание отделяется от другого запятой (10–20 слов).
9. Пункты 3–8 на английском языке.
10. Список литературы.

Образец оформления приведен на сайте: http://journal.nstu.ru/vestnik.

Внимание!

Дополнительная информация к статье, опубликованной в № 2(102) за 2026 г.: Бритвин А.В., Копылова О.А., Никитенко Н.С., Поллер Б.В., Поллер А.Б., Хайретдинов М.С., Шастин Л.В. Свойства атмосферных оптических линий для систем акустооптического мониторинга опасных природных и антропогенных процессов (Britvin A.V., Kopylova O.A., Nikitenko N.S., Poller B.V., Poller A.B., Khayretdinov M.S., Shastin L.V. Properties of atmospheric optical lines for acousto-optical monitoring systems of hazardous natural and anthropogenic processes):

Исследования выполнены в рамках государственного задания ИВМиМГ СО РАН FWNM-2025-0004.

This work was carried out under state contract with ICMMG SB RAS FWNM-2025-0004.