

ISSN 2712-9942

УСПЕХИ КИБЕРНЕТИКИ

RUSSIAN JOURNAL OF CYBERNETICS

2025
T. 6 № 2

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ НАУЧНЫЙ ЦЕНТР
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ СИСТЕМНЫХ
ИССЛЕДОВАНИЙ НАЦИОНАЛЬНОГО ИССЛЕДОВАТЕЛЬСКОГО
ЦЕНТРА «КУРЧАТОВСКИЙ ИНСТИТУТ»**

Успехи кибернетики

Russian Journal of Cybernetics

**Том 6
№ 2**

**Vol. 6
No. 2**

Москва
2025

Учредитель и издатель

Федеральное государственное автономное учреждение «Федеральный научный центр
Научно-исследовательский институт системных исследований Национального исследовательского центра
«Курчатовский институт»
(НИЦ «Курчатовский институт» — НИИСИ)

«Успехи кибернетики» — это рецензируемый научный журнал, в котором публикуются научные статьи по следующим специальностям (физико-математические, технические науки):

- 1.2.2. Математическое моделирование, численные методы и комплексы программ;
- 2.2.2. Электронная компонентная база микро- и нанoeлектроники, квантовых устройств;
- 2.3.1. Системный анализ, управление и обработка информации.

Миссия журнала — развитие научных направлений по заявленной тематике в России и за рубежом, соответствующих приоритетным направлениям развития науки, технологий и техники в Российской Федерации, а также перечню критических технологий РФ.

Журнал ориентирован на пропаганду передовых идей в области физики, математики, технических наук, участие в реализации задач, сформулированных Президентом РФ в Указе от 01.12.2016 № 642, по научно-технологическому развитию РФ, а также импортозамещению по приоритетным направлениям стратегического развития страны, соответствующим тематике журнала, обеспечение печатными площадями высококвалифицированных кадров, повышение качества диссертационных исследований в данных отраслях путем развития механизма профессионального и общественного обсуждения их научных результатов, воспитание молодого поколения ученых.

Журнал принимает к публикации оригинальные статьи; переводы статей, опубликованных в зарубежных журналах (при согласии правообладателя на перевод и публикацию); обзоры; эссе; комментарии; другие информационные материалы.

Издание будет полезно ученым, работающим в соответствующих областях наук, а также аспирантам и студентам.

Главный редактор

Бетелин Владимир Борисович, академик РАН, д. ф.-м. н., профессор, научный руководитель, НИЦ «Курчатовский институт» — НИИСИ, Москва

Заместитель главного редактора

Галкин Валерий Алексеевич, д. ф.-м. н., профессор, директор, Сургутский филиал НИЦ «Курчатовский институт» — НИИСИ, Сургут

Члены редакционной коллегии

Власов Сергей Евгеньевич, д. т. н., директор, НИЦ «Курчатовский институт» — НИИСИ, Москва

Панченко Владислав Яковлевич, академик РАН, д. ф.-м. н., профессор, вице-президент, Российская академия наук; заведующий кафедрой медицинской физики, Московский государственный университет им. М. В. Ломоносова, Москва

Савин Геннадий Иванович, академик РАН, д. ф.-м. н., профессор, научный руководитель, МСЦ — филиал НИЦ «Курчатовский институт» — НИИСИ, Москва
Соифер Виктор Александрович, академик РАН, д. т. н., профессор, президент, главный научный сотрудник научно-исследовательской лаборатории автоматизированных систем научных исследований, Самарский национальный исследовательский университет им. ак. С. П. Королева, Институт систем обработки изображений РАН, Самара

Чубариков Владимир Николаевич, д. ф.-м. н., профессор, заведующий кафедрой математических и компьютерных методов анализа, Московский государственный университет им. М. В. Ломоносова, Москва

Чаплыгин Юрий Александрович, академик РАН, д. т. н., профессор, президент, Московский институт электронной техники, Москва

Петров Игорь Борисович, член-корреспондент РАН, д. ф.-м. н., профессор, заведующий кафедрой информатики, Московский физико-технический институт, Москва

Смирнов Николай Николаевич, д. ф.-м. н., профессор, заместитель декана механико-математического факультета, Московский государственный университет им. М. В. Ломоносова, заместитель директора по стратегическим информационным технологиям, НИЦ «Курчатовский институт» — НИИСИ, Москва

Тишкин Владимир Федорович, член-корреспондент РАН, д. ф.-м. н., профессор, профессор кафедры вычислительных методов, Московский государственный университет им. М. В. Ломоносова, Москва

Губайдуллин Амир Анварович, д. ф.-м. н., профессор, директор, Тюменский филиал Института теоретической и прикладной механики им. С. А. Христиановича Сибирского отделения РАН, Тюмень

Яковлевский Михаил Владимирович, член-корреспондент РАН, д. ф.-м. н., профессор, заместитель директора по научной работе, Институт прикладной математики им. М. В. Келдыша РАН, Москва

Шагалев Рашит Мирзагалиевич, член-корреспондент РАН, д. ф.-м. н., заместитель директора Федерального научного центра

Старков Сергей Олегович, д. ф.-м. н., профессор, начальник отделения интеллектуальных кибернетических систем, Обнинский институт атомной энергетики, Обнинск

Еськов Валерий Матвеевич, д. ф.-м. н., д. б. н., профессор, заведующий отделом, Сургутский филиал НИЦ «Курчатовский институт» — НИИСИ, Сургут

Лаврентьев Михаил Михайлович, д. ф.-м. н., профессор, заместитель директора по научной работе, Институт автоматизации и электротехники Сибирского отделения РАН, Новосибирск

Пятков Сергей Григорьевич, д. ф.-м. н., профессор, заведующий кафедрой высшей математики, Югорский государственный университет, Ханты-Мансийск

Ильин Валерий Павлович, д. ф.-м. н., профессор, профессор кафедры прикладной математики, Новосибирский государственный технический университет, Новосибирск

Крыжановский Борис Владимирович, член-корреспондент РАН, д. ф.-м. н., руководитель, Центр оптико-нейронных технологий НИЦ «Курчатовский институт» — НИИСИ, Москва

Нагорнов Олег Викторович, д. ф.-м. н., профессор, первый проректор, Национальный исследовательский ядерный университет «МИФИ», Москва

Земпо Ясунари, PhD (Physics), Университет Хосэй, Токио, Япония

Пападопулос Атанас, Doctorat d'Etat (Habilitation), mathematics, Университет Страсбурга, Национальный центр научных исследований Франции, Страсбург, Франция

Ответственный секретарь

Гавриленко Тарас Владимирович, к. т. н.

Ответственный редактор

Чалова Анна Петровна, к. филол. н.

Технические редакторы

Моргун Дмитрий Алексеевич, к. ф.-м. н., доцент

Амелин Максим Евгеньевич, Батуро Серафим Александрович, Бобровская Ольга Павловна, Гавриленко Михаил Тарасович, Горбунов Дмитрий Владимирович

Переводчик

Троицкий Дмитрий Игоревич, к. т. н., доцент

Художественное оформление обложки

© Горбунов Александр Сергеевич

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций.

Свидетельство о регистрации средства массовой информации ЭЛ № ФС 77-78857 от 04.08.2020.

Решением Высшей аттестационной комиссии с 27 июня 2022 года журнал включен в «Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук».

Издается с 2020 года. Выпускается 4 раза в год.

Адрес редакции:

НИЦ «Курчатовский институт» — НИИСИ, 117218, г. Москва, Нахимовский просп., дом 36, корпус 1.

Телефон: +7 9226-54-57-88, факс: +7 (495) 719-76-81, e-mail: taras.gavrilenko@gmail.com. Сайт: jcyb.ru.

© НИЦ «Курчатовский институт» — НИИСИ

© Авторы

Founder and Publisher

Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”
(NRC “Kurchatov Institute” – SRISA)

Russian Journal of Cybernetics is peer-reviewed.

The journal publishes papers on physics, mathematics, and engineering. The key areas of interest are:

- simulation, numerical methods, and simulation software
- electronic components for micro- and nanoelectronics, and quantum devices
- system analysis, control, and information processing.

Our mission is advancing these research areas in Russia and globally, supporting the national priority fields of research and development, and the critical technologies.

The Journal promotes advanced concepts in physics, mathematics, engineering, and facilitates the national research and technology development under the Russian President Decree No. 642 dated 01.12.2016, and phasing out strategic imports. We offer an opportunity for professional researchers to publish. We improve the doctorate research quality through public peer review and educate the new generation of researchers.

The Journal accepts original papers; translations of papers from international journals (consent of the right holders for the translation and publication is required); reviews; essays; remarks, and event reports.

The Journal is intended for researchers, university, and doctorate students.

Chief Editor

Prof. *Vladimir B. Betelin*, member of the Russian Academy of Sciences (RAS),
Doctor of Science (Phys&Math), Academic Director, NRC “Kurchatov Institute” – SRISA, Moscow

Vice Chief Editor

Prof. *Valery A. Galkin*, Doctor of Science (Phys&Math), Director, Surgut Branch of NRC “Kurchatov Institute” – SRISA, Surgut

Editorial Board

Prof. *Sergey E. Vlasov*, Doctor of Science (Engineering), Director, NRC “Kurchatov Institute” – SRISA, Moscow

Prof. *Vladislav Ya. Panchenko*, member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math), Vice President, Russian Academy of Sciences (RAS), Chairperson of the
Department of Medical Physics, Lomonosov Moscow State University, Moscow

Prof. *Gennady I. Savin*, member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math),

Academic Director, Joint Supercomputer Center – Branch of Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, Moscow

Prof. *Viktor A. Soifer*, member of the Russian Academy of Sciences (RAS), Doctor of Science (Engineering), President, Chief Researcher Lab for Automated Research Systems, Korolev Samara
National Research University, Academic Director, Institute of Image Processing Systems, RAS, Samara

Prof. *Vladimir N. Chubarikov*, Doctor of Science (Phys&Math), Chairperson, Department of Mathematical and Computer Analysis Methods, Lomonosov Moscow State University, Moscow

Prof. *Yuri A. Chaplygin*, member of the Russian Academy of Sciences (RAS), Doctor of Science (Engineering), President, National Research University of Electronic Technology, Moscow

Prof. *Igor B. Petrov*, corresponding member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math), Chairperson, Department of Informatics,

Moscow Institute of Physics and Technology, Moscow

Prof. *Nikolay N. Smirnov*, Doctor of Science (Phys&Math), Deputy Dean, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Deputy Director for Strategic Information
Technology, NRC “Kurchatov Institute” – SRISA, Moscow

Prof. *Vladimir F. Tishkin*, corresponding member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math), Professor, Department of Computational Methods, Lomonosov
Moscow State University, Moscow

Prof. *Amir A. Gubaidullin*, Doctor of Science (Phys&Math), Director, Tyumen Division, Khristianovich Institute of Theoretical and Applied Mechanics, RAS Siberian Branch, Tyumen

Prof. *Mikhail V. Iakobovskii*, corresponding member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math), Deputy Director for Research, Keldysh Institute of Applied
Mathematics, RAS, Moscow

Prof. *Rashit M. Shagaliyev*, corresponding member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math), Deputy Director, Federal Research Center

Prof. *Sergey O. Starkov*, Doctor of Science (Phys&Math), Head, Division of Intelligent Cybernetic Systems, Obninsk Institute of Atomic Energy, Obninsk

Prof. *Valery M. Eskov*, Doctor of Science (Phys&Math), Doctor of Science (Biology), Head of Division, Surgut Branch of NRC “Kurchatov Institute” – SRISA, Surgut

Prof. *Mikhail M. Lavrentiev*, Doctor of Science (Phys&Math), Deputy Director for Research, Institute of Automation and Electrometry, RAS Siberian Branch, Novosibirsk

Prof. *Sergey G. Pyatkov*, Doctor of Science (Phys&Math), Chairperson, Department of Further Mathematics, Yurga State University, Khanty-Mansiysk

Prof. *Valerii P. Ilyin*, Doctor of Science (Phys&Math), Professor, Department of Applied Mathematics, Novosibirsk State Technical University, Novosibirsk

Prof. *Boris V. Kryzhanovskii*, corresponding member of the Russian Academy of Sciences (RAS), Doctor of Science (Phys&Math), Director, Opto-neural Technologies Center, NRC “Kurchatov
Institute” – SRISA, Moscow

Prof. *Oleg V. Nagornov*, Doctor of Science (Phys&Math), First Vice-Rector, National Research Nuclear University MEPhI, Moscow

Prof. *Yasunari Zempo*, PhD (Physics), Hosei University, Tokyo, Japan

Prof. *Athanase Papadopoulos*, Doctorat d'Etat (Habilitation), Mathematics, University of Strasbourg, French National Centre for Scientific Research, Strasbourg, France

Executive Editor

Taras V. Gavrilenko, Ph.D. (Engineering)

Managing Editor

Anna P. Chalova, Ph.D. (Linguistics)

Technical Editors

Dmitry A. Morgun, Ph.D. (Phys&Math), Assoc. Prof.

Maksim E. Amelin

Olga P. Bobrovskaya

Serafim A. Batur

Mikhail T. Gavrilenko

Dmitry V. Gorbunov

Translator

Dmitry I. Troitsky, Ph.D. (Engineering), Assoc. Prof.

Cover Design

© Alexander S. Gorbunov

The journal is registered with the Federal Service for Supervision of Communications, Information Technology and Mass Media.

Mass media registration certificate EL No. FS 77-78857 dated on 04.08.2020.

On June 27, 2022 the National Board for Academic Degrees and Titles put the Russian Journal of Cybernetics on the list of leading journals for publishing papers by doctorate students.

Published since 2020. 4 issues per year.

Editorial Board Address:

NRC “Kurchatov Institute” – SRISA, Russia 117218, Moscow, Nakhimovsky pr., 36/1.
Tel.: +7(922)654-5788, fax: +7(495)719-76-81, e-mail: taras.gavrilenko@gmail.com. Web: jcyb.ru.

СОДЕРЖАНИЕ

В. Б. Бетелин	
Поколение романтиков знаний	6
С. Г. Вольпин, И. В. Афанаскин, А. К. Пономарев, Ю. М. Штейнберг, Д. В. Солопов	
Изучение пространственной неоднородности карбонатного трещиновато-порового пласта на основе гидродинамических исследований и сейсмических данных	8
В. А. Галкин	
О динамике примеси в поле течения стратифицированной вязкой несжимаемой жидкости	22
И. В. Афанаскин, С. Г. Вольпин, А. В. Королев, М. Ю. Ахапкин	
Математические модели для анализа и прогнозирования разработки нефтяных месторождений: емкостно-резистивная модель без учета забойного давления и логарифмическая модель обводнения скважины	27
В. П. Кощеев	
К задаче построения уравнения Эйлера–Лагранжа с помощью леммы Пуанкаре	40
В. П. Кощеев	
Классические и квантовомеханические поправки к решению уравнения Эйлера–Лагранжа	44
Я. Д. Бакуменко, А. Н. Таджибова	
Современные проблемы и решения в области машинного перевода	47
Г. Е. Деев, С. В. Ермаков, S. Starkloff	
Линейные разностные уравнения и структуры Фибоначчи	60
Э. А. Юнусов, Д. В. Горбунов	
Проблемы и обзор алгоритмических решений в технологии FFF 3D-печати	67
Э. М. Исмаилов, Т. В. Гавриленко	
Проблема устойчивости критической инфраструктуры в условиях высокого уровня автоматизации	75
С. В. Мещеряков, В. Ф. Шуршев	
Децентрализованные технологии в кадастровом учете недвижимости	80
Ш. И. Мутаиров, К. И. Бушмелева	
Анализ методов извлечения ключевых слов из текстовых данных	86
И. П. Саблин, М. В. Михайлюк, Д. В. Омельченко, Д. А. Кононов	
Моделирование звука в системах виртуального окружения	92
В. А. Солозобов, С. А. Лысенкова	
Методика оценки и выбора оптимальной модели, прогнозирующей окончание периода заморозков	100
Э. Г. Тунян, Р. С. Сазиков, С. А. Харламов	
Обогащение базы знаний с помощью автоматического извлечения ключевых слов и аннотаций	108
М. В. Юрчишина	
Функциональная модель системы поддержки принятия решений «Оптимальный учебный план»	114

CONTENTS

V. B. Betelin	
The Generation of Those Who Were Romancing Knowledge	6
S. G. Volpin, I. V. Afanaskin, A. K. Ponomarev, Yu. M. Shteinberg, D. V. Solopov	
Spatial Heterogeneity of Carbonate Fractured-Porous Formations from Well Testing and Seismic Data	8
V. A. Galkin	
Dynamic Behavior of Impurities in the Flow Field of a Stratified Viscous Incompressible Fluid	22
I. V. Afanaskin, S. G. Volpin, A. V. Korolev, M. Yu. Ahapkin	
Mathematical Models for the Analysis and Forecasting of Oil Field Development: A Capacitive-Resistive Model Without Bottomhole Pressure and a Logarithmic Model of Well Water Breakthrough	27
V. P. Koshcheev	
Constructing the Euler–Lagrange Equation Using the Poincaré Lemma	40
V. P. Koshcheev	
Classical and Quantum-Mechanical Corrections to the Solution of the Euler–Lagrange Equation	44
Ya. D. Bakumenko, A. N. Tadzhibova	
Current Challenges and Approaches in Machine Translation	47
G. E. Deev, S. V. Ermakov, S. Starkloff	
Linear Difference Equations and Fibonacci Structures	60
E. A. Yunusov, D. V. Gorbunov	
FFF 3D Printing Technology Algorithms: Problems and Review	67
E. M. Ismailov, T. V. Gavrilenko	
Sustainability of Highly Automated Critical Infrastructure	75
S. V. Meshcheryakov, V. F. Shurshev	
Decentralized Technologies for Cadastral Systems	80
Sh. I. Mutairov, K. I. Bushmeleva	
Comparative Analysis of Keyword Extraction Strategies	86
I. P. Sablin, M. V. Mikhailyuk, D. V. Omelchenko, D. A. Kononov	
Sound Modeling in Virtual Environment Systems	92
V. A. Solozobov, S. A. Lysenkova	
Assessment and Selection of an Optimal Model for Forecasting the End of the Frost Period	100
E. G. Tunyan, R. S. Sazikov, S. A. Kharlamov	
Automatic Extraction of Keywords and Summaries for Knowledge Base Population	108
M. V. Yurchishina	
Functional Model of the Optimal Curriculum Decision Support System	114

ПОКОЛЕНИЕ РОМАНТИКОВ ЗНАНИЙ

В. Б. Бетелин

*Федеральное государственное автономное учреждение «Федеральный научный центр
Научно-исследовательский институт системных исследований Национального исследовательского
центра «Курчатовский институт», Москва, Российская Федерация
ORCID: <http://orcid.org/0000-0001-6646-2660>, ✉ betelin@inbox.ru*

Для цитирования: Бетелин В. Б. Поколение романтиков знаний. Успехи кибернетики. 2025;6(2):6–7.

THE GENERATION OF THOSE WHO WERE ROMANCING KNOWLEDGE

V. B. Betelin

*Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”,
Moscow, Russian Federation
ORCID: <http://orcid.org/0000-0001-6646-2660>, ✉ betelin@inbox.ru*

Cite this article: Betelin V. B. The Generation of Those Who Were Romancing Knowledge. Russian Journal of Cybernetics. 2025;6(2):6–7.



Сергей Григорьевич Вольпин, 17.07.1946 — 04.01.2025

Согласно толковому словарю С. И. Ожегова романтика — это идеи и чувства, эмоционально возвышающие человека. Порождать эти идеи и чувства могут условия жизни людей, обстановка, содействующие их эмоционально-возвышенному мироощущению. Формирование таких идей и чувств нашего поколения пришлось на время, когда в СССР на государственном уровне господствовал культ знаний, то есть в условиях нашей жизни, окружающей нас обстановке именно стремление к знаниям, понимание их непреходящей ценности, которые заложили в нас эти условия и обстановка, и являлись

основой эмоционально-возвышенного мироощущения нашего поколения — поколения романтиков знаний, творческих исканий и трудных путей.

Художественное описание сути и образа жизни романтиков знаний — сказка для научных сотрудников младшего возраста «Понедельник начинается в субботу» Аркадия и Бориса Стругацких. Это была настольная книга нашего поколения, а ее название — наш жизненный девиз.

К этому поколению принадлежал до его ухода из жизни и мой многолетний друг и соратник Сергей Григорьевич Вольпин, с которым мы более шестидесяти лет назад учились в одном классе в московской физико-математической школе. Девиз «Понедельник начинается в субботу» тогда еще незримо, но уже витал над нами в эти школьные годы, которые требовали и от Сергея, и от всех нас усвоения большого объема теоретических и практических знаний по математике, физике и программированию, при большой увлеченности поэзией и песнями Окуджавы и Визбора, стихами Вознесенского и Ахмадуллиной, спектаклями театра на Таганке. Именно незримое присутствие этого девиза в школе научило и Сергея, и всех нас не только уметь знать, но и уметь думать, и уметь работать, то есть, по сути дела, научило ответу на вопрос Обломова «зачем жить».

Понедельник начинался в субботу для Сергея и в годы учебы в Институте нефтехимической и газовой промышленности им. Губкина, и после окончания в годы практической и исследовательской работы непосредственно на месторождениях нефти. Понедельник начинался для Сергея в субботу и тогда, когда он вел научно-исследовательскую работу и в научно-исследовательских институтах нефтяной отрасли, и в последние семнадцать лет в НИИСИ РАН, то есть вся его практическая и творческая жизнь прошла под этим девизом.

Сергей ушел от нас в понедельник вечности, который начался для него в субботу 4 января 2025 года. Это невосполнимая потеря для всех нас, тех, кто его знал, любил и работал с ним.

ИЗУЧЕНИЕ ПРОСТРАНСТВЕННОЙ НЕОДНОРОДНОСТИ КАРБОНАТНОГО ТРЕЩИНОВАТО-ПОРОВОГО ПЛАСТА НА ОСНОВЕ ГИДРОДИНАМИЧЕСКИХ ИССЛЕДОВАНИЙ И СЕЙСМИЧЕСКИХ ДАННЫХ

С. Г. Вольпин^{1,a}, И. В. Афанаскин^{2,b}, А. К. Пономарев^{1,c}, Ю. М. Штейнберг^{1,d}, Д. В. Солопов^{1,e}

¹ Федеральное государственное автономное учреждение «Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Москва, Российская Федерация

² Государственное автономное образовательное учреждение высшего образования Альметьевский государственный технологический университет «Высшая школа нефти», г. Альметьевск, Российская Федерация

^a ORCID: <http://orcid.org/0000-0002-3591-2910>, sergvolpin@gmail.com

^b ORCID: <http://orcid.org/0000-0003-0617-221X>, ivan@afanaskin.ru

^c ORCID: <http://orcid.org/0000-0003-1924-6759>, akponomarev@mail.ru

^d ORCID: <http://orcid.org/0000-0002-3159-0408>, yurash22@gmail.com

^e ORCID: <http://orcid.org/0000-0002-4357-7462>, soldi99@gmail.com

Аннотация: в статье представлены результаты комплексной интерпретации материалов гидродинамических исследований скважин, эксплуатирующих нефтяную залежь с карбонатным трещиновато-поровым коллектором. Целью работ являлось определение фильтрационно-емкостных свойств пласта и выявление высокопроницаемых путей фильтрации. Исследования проведены методами восстановления давления, падения давления и длительных кривых изменения давления и дебита. Интерпретация осуществлялась по методу наилучшего совмещения расчетных и фактических кривых с использованием диагностических графиков Бурде и Блэйсингейма. Дополнительно выполнено комплексирование результатов гидродинамических исследований скважин с данными геофизических и сейсмических исследований (карты аномалий дуплексных волн). Получены количественные оценки проницаемости, пьезопроводности, скин-факторов и расстояний до границ. Установлены зоны повышенной трещиноватости, подтвержденные аномалиями на картах сейсмических атрибутов. Показано, что комплексный подход позволяет повысить достоверность интерпретации и выявить латеральные флюидоупоры.

Ключевые слова: гидродинамические исследования скважин, метод восстановления давления, метод падения давления, длительные кривые, гидропрослушивание, трещиновато-поровый коллектор, проницаемость, анизотропия, скин-фактор, дуплексные волны, флюидоупоры.

Благодарности: публикация выполнена в рамках государственного задания НИЦ «Курчатовский институт» — НИИСИ по теме № FNEF-2024-0001 «Создание и реализация доверенных систем искусственного интеллекта, основанных на новых математических и алгоритмических методах, моделях быстрых вычислений, реализуемых на отечественных вычислительных системах».

Для цитирования: Вольпин С. Г., Афанаскин И. В., Пономарев А. К., Штейнберг Ю. М., Солопов Д. В. Изучение пространственной неоднородности карбонатного трещиновато-порового пласта на основе гидродинамических исследований и сейсмических данных. *Успехи кибернетики*. 2025;6(2):8–21.

Поступила в редакцию: 01.06.2025.

В окончательном варианте: 17.06.2025.

SPATIAL HETEROGENEITY OF CARBONATE FRACTURED-POROUS FORMATIONS FROM WELL TESTING AND SEISMIC DATA

S. G. Volpin^{1,a}, I. V. Afanaskin^{2,b}, A. K. Ponomarev^{1,c}, Yu. M. Shteinberg^{1,d}, D. V. Solopov^{1,e}

¹ Scientific Research Institute for System Analysis of the National Research Centre "Kurchatov Institute", Moscow, Russian Federation

² Almet'yevsk State Technological University Petroleum Higher School, Almet'yevsk, Russian Federation

^a ORCID: <http://orcid.org/0000-0002-3591-2910>, sergvolpin@gmail.com

^b ORCID: <http://orcid.org/0000-0003-0617-221X>, ivan@afanaskin.ru

^c ORCID: <http://orcid.org/0000-0003-1924-6759>, akponomarev@mail.ru

^d ORCID: <http://orcid.org/0000-0002-3159-0408>, yurash22@gmail.com

^e ORCID: <http://orcid.org/0000-0002-4357-7462>, soldi99@gmail.com

Abstract: we studied the spatial heterogeneity of carbonate fractured-porous reservoirs using a comprehensive interpretation of well-test data. Our goal was to evaluate the reservoir's flow and storage properties and to identify high-permeability flow channels. We analyzed pressure build-up tests, pressure fall-off tests, and long-term pressure and rate histories. We interpreted the data by matching calculated and measured curves, and applying Burdine and Blasingame diagnostic plots. We also integrated the well test results with geophysical and seismic data, including duplex wave anomaly maps. The analysis provided quantitative estimates of permeability, diffusivity, skin factor, and distance to flow boundaries. We identified highly fractured zones, which correlated with anomalies on seismic attribute maps. The results show that combining dynamic and geophysical data improves interpretation reliability and helps delineate lateral barriers.

Keywords: hydrodynamic well testing, pressure recovery, pressure drop, long-period curves, wireline logging, porous fractured reservoir, permeability, anisotropy, skin factor, duplex waves, reservoir caprock.

Acknowledgements: this is a part of the FNEF-2024-0001 Development and Deployment of Trusted AI Systems based on New Mathematical and Algorithmic Approaches and Fast Computing Models Compatible with Domestic Computer Hardware government contract.

Cite this article: Volpin S. G., Afanaskin I. V., Ponomarev A. K., Shteinberg Yu. M., Solopov D. V. Spatial Heterogeneity of Carbonate Fractured-Porous Formations from Well Testing and Seismic Data. *Russian Journal of Cybernetics*. 2025;6(2):8–21.

Original article submitted: 01.06.2025.

Revision submitted: 17.06.2025.

Введение

Гидродинамические исследования скважин (ГДИС) представляют собой один из важнейших методов изучения продуктивных пластов, применяемый на различных этапах разработки нефтяных месторождений. Целью ГДИС является количественная и качественная оценка фильтрационно-емкостных свойств пласта, идентификация границ залежи и дренируемой области, а также диагностика состояния призабойной зоны скважины и определение энергетического состояния пласта.

Среди методов ГДИС особое значение занимают классические подходы: метод восстановления давления (при котором регистрируется кривая восстановления давления — КВД) и метод падения давления (при котором регистрируется кривая падения давления — КПД). КВД регистрируется при исследовании добывающих скважин, анализируются изменения забойного давления после остановки добычи. КПД регистрируется при исследовании нагнетательных скважин, анализируются изменения забойного давления после остановки закачки. Оба метода ориентированы на интерпретацию переходных процессов в пласте и позволяют оценивать пластовое давление, гидропроводность, пьезопроводность, скин-фактор и расстояния до границ, а также некоторые другие специальные параметры интерпретационных моделей (например, параметры модели двойной пористости, длину и проводимость трещины гидроразрыва пласта, работающую длину горизонтальной скважины и др.). Особое значение этих двух методов обусловлено двумя причинами. Во-первых, они позволяют достаточно точно определить пластовое давление. Более точным является только прямой замер в пьезометрических скважинах. Во-вторых, при регистрации КВД и КПД точно известен дебит (расход), так как он равен нулю, а кривая забойного давления более гладкая, чем при работе скважины (она содержит меньше шумов, а значит, из нее можно извлечь больше полезной информации и ее легче анализировать).

Развитие технологий последних 15 лет дало импульс применению методов безостановочных исследований — длительных кривых изменения давления и дебита (ДКИД). Эти исследования выполняются в работающей скважине и позволяют извлекать информацию из показателей нормальной эксплуатации скважины, что особенно ценно в условиях, когда остановка скважины невозможна или экономически нецелесообразна. Они позволяют существенно расширить охват пласта исследованиями.

Еще одним важным методом является гидропрослушивание — исследование, основанное на регистрации давления в реагирующих скважинах при искусственном возмущении (например, резкой остановке или изменении режима) возмущающей скважины. Этот метод применяется для оценки связей между скважинами и свойств пласта в межскважинном пространстве.

Комплексное применение КВД, КПД, ДКИД и гидропрослушивания позволяет не только детализировать локальные свойства пласта вокруг скважин, но и выявить латеральные и вертикальные неоднородности, трещиноватость и наличие различного рода границ (как непроницаемых, так и частично проницаемых). Наиболее важным такой подход становится в условиях сложнопостроенных

карбонатных трещиновато-пористых коллекторов, где изменчивость свойства пласта особенно велика.

Одной из приоритетных задач анализа разработки месторождений с трещиновато-поровыми коллекторами является получение достоверной информации о пространственном распределении фильтрационно-емкостных свойств (ФЕС) пласта [1–3]. Это особенно актуально для карбонатных отложений, характеризующихся высокой геологической неоднородностью, наличием систем направленных трещин и латеральных экранов. Для достижения данной цели необходимо интеграционное использование методов гидродинамических исследований скважин, геофизических и сейсмических исследований.

Настоящее исследование выполнено на участке нефтяной залежи одного из нефтяных месторождений Тимано-Печерской нефтегазоносной провинции с карбонатным трещиновато-пористым коллектором. Объектом анализа являются данные по скв. №№ 11Г, 13Г, 1001, 1002 и 1006 (номера скважин изменены), включающие результаты комплексных ГДИС: метод восстановления давления, метод падения давления, исследования на режимах с непрерывной регистрацией давления и дебита, а также результаты гидропрослушивания. Целью работы является определение ФЕС, идентификация высокопроницаемых путей фильтрации, а также уточнение положения латеральных флюидоупоров [4–7] с учетом сейсмических данных.

Методика проведения исследования

Гидродинамические исследования проводились в соответствии с методическими указаниями [8]. КВД и КПД, ДКИД интерпретировались в соответствующем программном комплексе. Применялись аналитические модели однородного пласта с различными граничными условиями, включая бесконечный пласт, пласт с линейной границей постоянного давления и пласт с двумя непроницаемыми границами. Интерпретация осуществлялась методом наилучшего совмещения расчетных и фактических кривых, дополнительно использовались диагностические графики (Бурде, Блэйсингейм) [2, 9–12].

При интерпретации КВД и КПД предварительно выполнялось построение диагностических графиков для идентификации модели пласта, после чего осуществлялся итерационный подбор параметров: гидропроводности, пьезопроводности, абсолютной и фазовой проницаемости, скин-фактора, расстояний до границ. В зонах нагнетания оценивались параметры текущей проницаемости по воде, в зонах добычи — по нефти. Расчеты проводились с учетом относительных фазовых проницаемостей, полученных из лабораторных данных и их аппроксимации.

Методика интерпретации ГДИС базировалась на подходах, изложенных в работах Кульпина Л. Г. и Мясникова Ю. А. [1], а также Bourdet D. [9, 10].

Поток жидкости в нефтяном пласте при ГДИС на неустановившихся режимах фильтрации (КВД, КПД, ДКИД, гидропрослушивание) описывается теорией упругого режима. Рассматривается фильтрация слабосжимаемой жидкости в упругой среде. Такое течение описывает уравнение пьезопроводности [13]:

$$\frac{\partial p}{\partial t} = \chi \Delta p, \quad (1)$$

где p — давление в точке пласта, t — время, χ — пьезопроводность, Δ — оператор Лапласа.

Пьезопроводность зависит от свойств пласта и жидкости:

$$\chi = \frac{k}{m\mu c_t}, \quad (2)$$

где k — проницаемость пласта, m — пористость пласта, μ — динамическая вязкость жидкости, c_t — сжимаемость системы пласт + флюид. Если фильтрация двухфазная (нефть и вода), но за время исследования водонасыщенность пласта и, соответственно, обводненность добывающих скважин меняются мало (не меняются), то уравнение (1) все еще можно использовать, однако нужно рассчитать эффективную вязкость с учетом относительных фазовых проницаемостей (ОФП), а также учесть двухфазный характер насыщения при расчете сжимаемости системы.

Уравнение (1) называется основным дифференциальным уравнением теории упругого режима фильтрации. Оно записано для однородного пласта. С его помощью можно решать задачи для кусочно-неоднородного пласта различной геометрии. Однако возможны и более сложные формы записи уравнения пьезопроводности.

Современным методом интерпретации ГДИС является метод наилучшего совмещения расчетной и фактической кривых забойного давления. Для его реализации с учетом сложных интерпретационных моделей (которые позволяют описывать различную конструкцию скважины, несколько вариантов

моделей неоднородности пласта по свойствам и различные конфигурации границ пласта) используют решение уравнения пьезопроводности в изображениях. Однако визуальный анализ такого решения и понимание взаимосвязей решения и параметров модели сложны, поэтому для иллюстрации часто используют аналитические решения уравнения (1). В данной работе применялись модели вертикальной скважины и горизонтальной скважины как в однородном бесконечном пласте, так и в пласте с одной границей постоянного давления и двумя параллельными непроницаемыми границами. Такие модели описываются двумя режимами притока жидкости к скважине: радиальный поток и линейный поток. Рассмотрим два решения уравнения (1): прямолинейно-параллельный фильтрационный поток упругой жидкости и плоскорадиальный фильтрационный поток упругой жидкости [13]. Эти два решения описывают необходимые нам режимы притока жидкости к скважине. Будем рассматривать только решения для постоянного дебита, поскольку решения для постоянного давления на границах не представляет для нас в рамках этой статьи практического интереса.

Рассмотрим прямолинейно-параллельный фильтрационный поток упругой жидкости. Пусть в полубесконечном горизонтальном пласте постоянной толщины h и ширины B начальное давление всюду постоянно и равно p_0 [13]. В удаленных точках пласта $x \rightarrow \infty$ давление в любой момент времени равно p_0 . В момент времени $x = 0$ пускается в работу галерея с постоянным объемным дебитом q . В пласте образуется неустановившийся прямолинейно-параллельный поток жидкости. Давление в любой точке пласта x и в любой момент времени t можно определить, интегрируя уравнение (1). В общем виде решение можно записать как:

$$p(x, t) = p_0 + \frac{q}{B} \frac{\mu}{kh} \left\{ x \left[1 - \operatorname{erf} \left(\frac{x}{2\sqrt{\chi t}} \right) \right] - \frac{2\sqrt{\chi t}}{\sqrt{\pi}} \exp \left(-\frac{x^2}{4\chi t} \right) \right\}, \quad (3)$$

где $\operatorname{erf}(x)$ — специальная табулированная функция, называемая интегралом вероятности.

Рассмотрим плоскорадиальный фильтрационный поток упругой жидкости. Пусть в неограниченном горизонтальном пласте постоянной толщины h имеется добывающая скважина нулевого радиуса (точечный сток) [13]. Начальное пластовое давление во всем пласте одинаково и равно p_0 . В удаленных точках пласта $r \rightarrow \infty$ давление в любой момент времени равно p_0 . В момент времени $t = 0$ скважина пускается в работу с постоянным объемным дебитом q . В пласте образуется неустановившийся плоскорадиальный поток жидкости. Распределение давления в пласте (в любой его точке r в каждый момент времени t) определяется интегрированием уравнения (1). В общем виде решение можно записать как:

$$p(r, t) = p_0 - \frac{q}{4\chi} \frac{\mu}{kh} \left[-Ei \left(-\frac{r^2}{4\chi t} \right) \right], \quad (4)$$

где $-Ei(-x)$ — специальная табулированная функция, называемая интегральной показательной функцией.

Решение (4) называется основной формулой теории упругого режима фильтрации.

Дифференциальное уравнение упругого режима (1) является линейным. Значит, к его решению применим метод суперпозиции (сумма решений также является решением) [13]. Это позволяет учитывать работу скважин с переменным дебитом, рассчитывать интерференцию скважин, а также конструировать решения для задач с непроницаемыми границами и границами постоянного давления с помощью метода зеркального отображения источника (стока).

Результаты интерпретации исследований

В данном разделе представлены результаты исследования скв. №№ 11Г, 13Г, 1001, 1002, 1006 методом восстановления (падения) давления и исследования скважин на режимах с непрерывной регистрацией изменения давления (без остановки скважин).

Для получения представления о районе работ на рис. 1 приведена выкопировка из структурной карты по кровле проницаемых карбонатов для рассматриваемой залежи.

Для сопоставления результатов ГДИС и сейсмических исследований на рис. 2 приведена выкопировка из карты аномалий дуплексных волн для рассматриваемой залежи [14]. Сейсмические построения проводились группой Хромовой И. Ю.

Использованные в процессе интерпретации ГДИС ОФП в системе нефть-вода представлены на рис. 3 (приняты по действующему проектному документу). При известной обводненности продукции

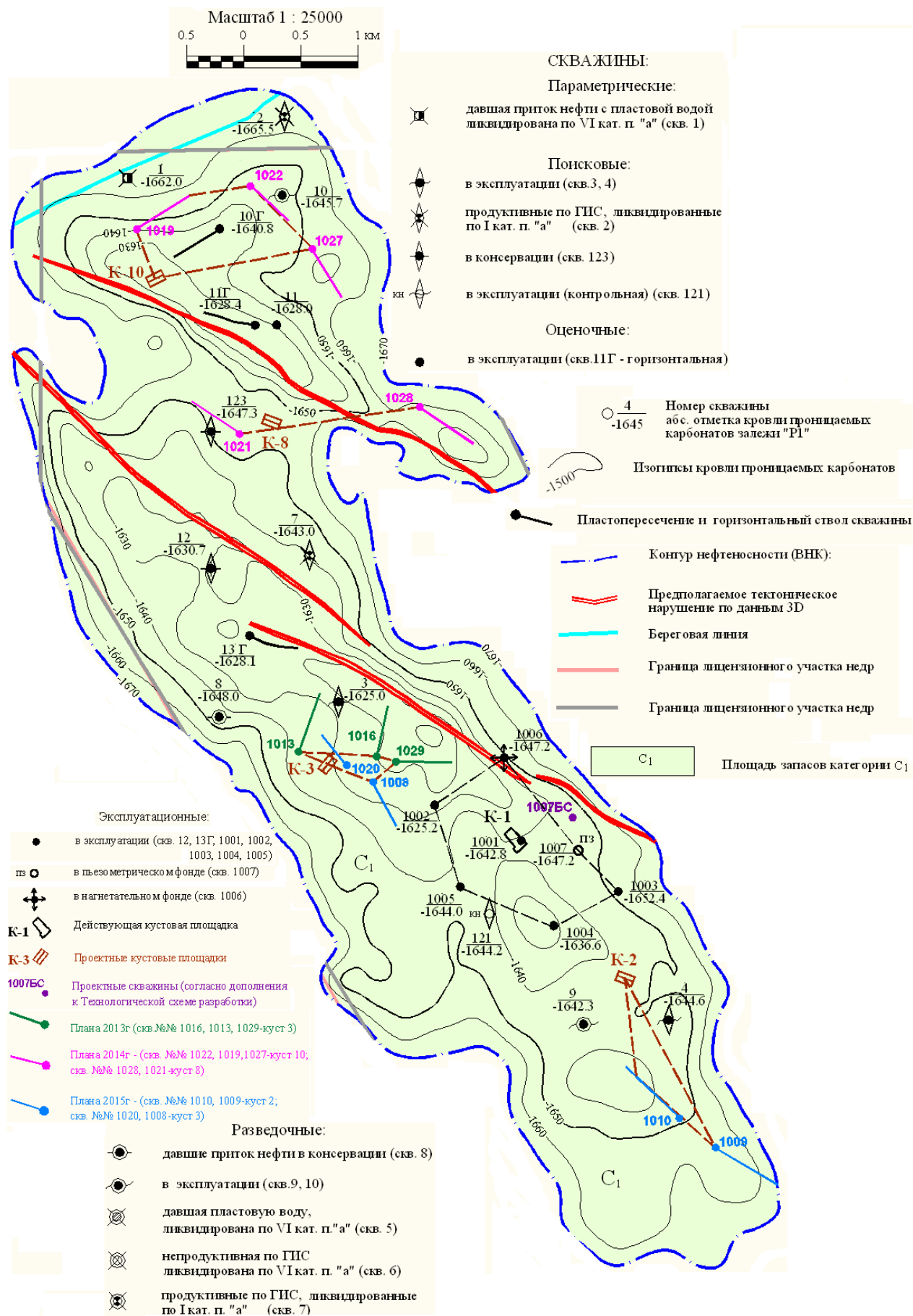


Рис. 1. Выкопировка из структурной карты по кровле проницаемых карбонатов [15]

добывающих скважин по зависимости доли воды в потоке в поверхностных условиях от водонасыщенности, приведенной на рис. 3, оценивалась водонасыщенность пласта, а по водонасыщенности рассчитывались текущие значения ОФП по нефти и воде в окрестности добывающих скважин. В окрестности нагнетательной скважины водонасыщенность рассчитана из значения нефтенасыщенности, а нефтенасыщенность принята равной остаточной насыщенности нефти при вытеснении нефти водой.

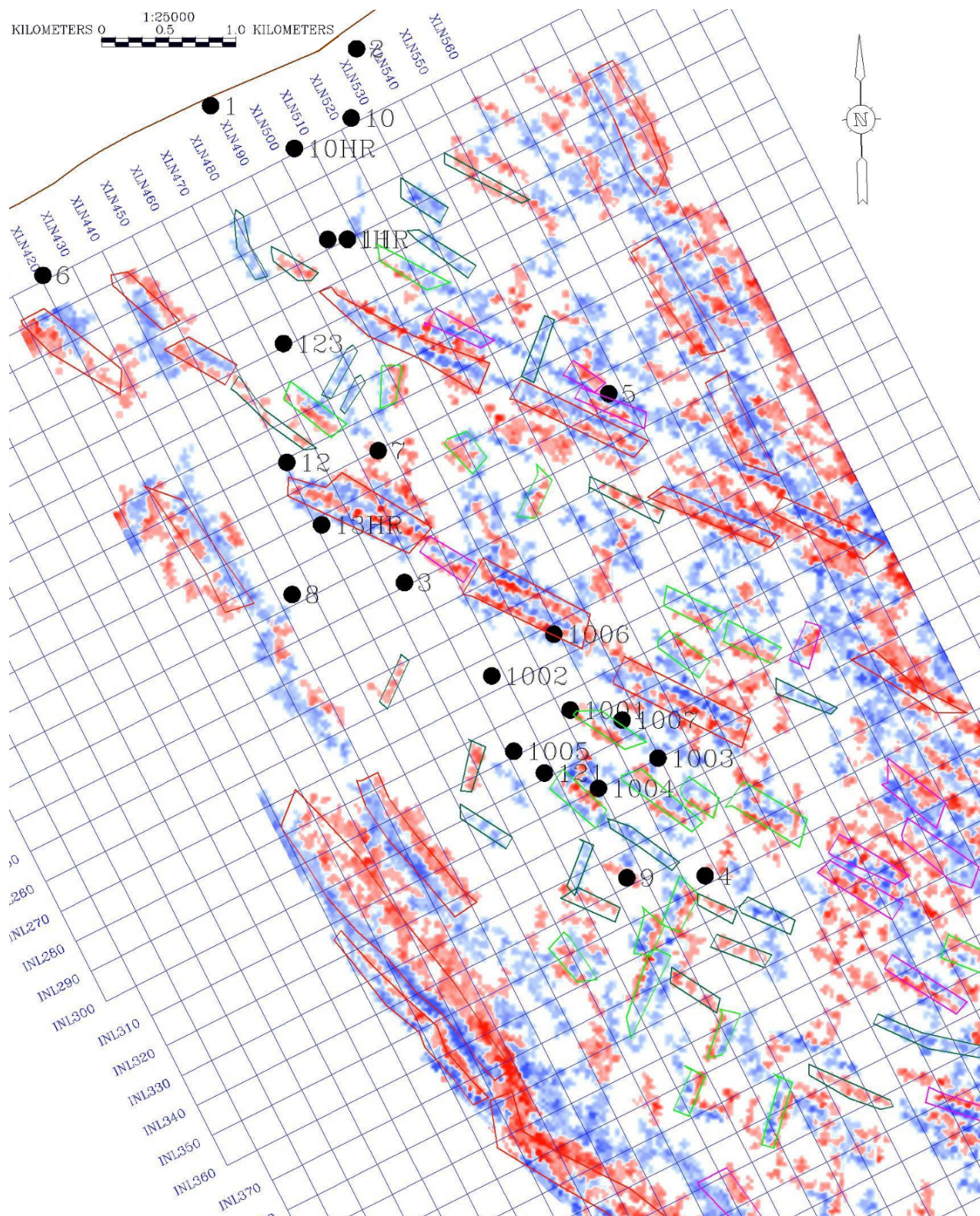


Рис. 2. Выкопировка из карты аномалий дуплексных волн (по результатам Хромовой И. Ю.) [15]

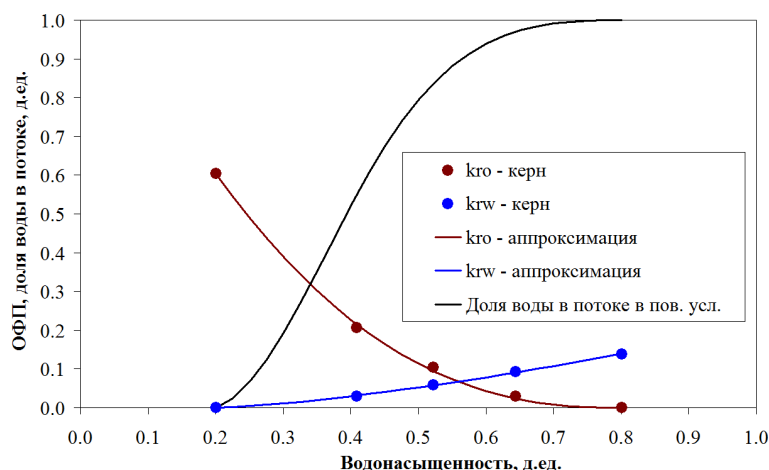


Рис. 3. ОФП в системе вода-нефть, использованные при обработке ГДИС.
kro — ОФП по нефти, krgw — ОФП по воде

Было исследовано 5 скважин (№№ 11Г, 13Г, 1001, 1002, 1006). При этом было зарегистрировано 6 КВД, 3 КПД, 3 длительных кривых изменения дебита и забойного давления (режимы) и проведено гидропрослушивание по 3 направлениям. Все исследования обработаны по методу наилучшего совмещения.

На рис. 4 представлены основные результаты ГДИС на выкопировке из карты аномалий дуплексных волн. В табл. 1 приведены результаты интерпретации ГДИС методом восстановления давления и методом исследования на режимах с непрерывной регистрацией изменения давления. В табл. 2 приведены результаты интерпретации ГДИС методом гидропрослушивания.

Перейдем к анализу результатов ГДИС по скважинам.

В скв. № 11Г зарегистрированы 2 КВД и 1 ДКИД. Приняты 2 интерпретационные модели — модель горизонтальной скважины, вскрывающей однородный бесконечный пласт с подошвенной водой (что обосновано тем, что нефтяная залежь подстилается подошвенной водой), и модель горизонтальной скважины, вскрывающей однородный пласт с линейной границей постоянного давления (что обосновывается тем, что согласно карте аномалий дуплексных волн вблизи скв. № 11Г находится зона аномальных фильтрационных свойств). С учетом результатов интерпретации КПД, зарегистрированной в скв. № 1006, и результатов гидропрослушивания более предпочтительной является модель горизонтальной скважины, вскрывающей однородный пласт с линейной границей постоянного давления. Расстояние до границы постоянного давления по результатам ГДИС составляет 47–56 м, что соответствует возможному расположению высокопроницаемого коридора, выявленного по карте аномалий дуплексных волн. Коэффициент вертикальной анизотропии проницаемости составляет 0.12 д.ед., что говорит о слабо развитой системе трещин, плохо сообщающихся между собой. Учитывая все вышесказанное и невысокое значение абсолютной проницаемости пласта, можно сделать вывод о том, что скв. № 11Г находится вблизи высокопроницаемого коридора в слабопроницаемой зоне.

В скв. № 13Г зарегистрированы 1 КВД и 1 ДКИД. Принятая интерпретационная модель — модель горизонтальной скважины, вскрывающей однородный пласт с линейной границей постоянного давления (что обосновывается тем, что согласно карте аномалий дуплексных волн вблизи скв. № 13Г находится зона аномальных фильтрационных свойств, параллельная горизонтальному стволу). Расстояние до границы постоянного давления по результатам ГДИС составляет 48–54 м, что соответствует возможному расположению высокопроницаемого коридора, выявленного по карте аномалий дуплексных волн. Коэффициент вертикальной анизотропии проницаемости составляет 0.004 д.ед., что говорит о крайне слабо развитой системе трещин, очень плохо сообщающихся между собой. Учитывая все вышесказанное и невысокое значение абсолютной проницаемости пласта, можно сделать вывод о том, что скв. № 13Г находится вблизи высокопроницаемого коридора в слабопроницаемой зоне.

В скв. № 1001 зарегистрированы 1 КВД и 1 ДКИД. Принятая интерпретационная модель — модель вертикальной скважины, вскрывающей однородный бесконечный пласт. Отсутствие влияния высокопроницаемого коридора на зарегистрированную КВД объясняется тем, что за время регистрации

КВД (почти 200 ч.) скважина только начала выходить на радиальный режим фильтрации, что, в свою очередь, объясняется высокой обводненностью продукции и низкой текущей фазовой проницаемостью как по нефти, так и по воде. Отсутствие влияния высокопроницаемого коридора на зарегистрированную ДКИД объясняется малым дебитом скважины по жидкости (малым возмущением пласта) и тем, что средняя часть кривой изменения давления зарегистрирована телеметрической системой в составе установки электрического центробежного насоса с разрешающей способностью 1 кгс/см². Кроме того, для уверенного выделения границ на ДКИД необходим замер дебита по фазам 3–5 раз в сутки, а в процессе исследования скв. № 1001 по техническим причинам замер дебита производился не чаще 1 раза в сутки. Учитывая все вышесказанное и невысокое значение абсолютной проницаемости пласта, можно сделать вывод о том, что скв. № 1001 находится в слабопроницаемой зоне. Согласно результатам гидропрослушивания скв. № 1001 находится вблизи высокопроницаемого коридора.

В скв. № 1002 зарегистрированы 2 КВД. Принято 2 интерпретационные модели — модель вертикальной скважины, вскрывающей однородный бесконечный пласт (что является самой простой моделью), и модель вертикальной скважины, вскрывающей однородный пласт с линейной границей постоянного давления (что обосновывается тем, что согласно карте аномалий дуплексных волн на некотором расстоянии от скв. № 1002 находится зона аномальных фильтрационных свойств). С учетом результатов интерпретации КПД, зарегистрированной в скв. № 1006, и результатов гидропрослушивания более предпочтительной является модель вертикальной скважины, вскрывающей однородный пласт с линейной границей постоянного давления. Расстояние до границы постоянного давления по результатам ГДИС составляет 655–682 м, что соответствует возможному расположению высокопроницаемого коридора, выявленного по карте аномалий дуплексных волн. Учитывая все вышесказанное и невысокое значение абсолютной проницаемости пласта, можно сделать вывод о том, что скв. № 1002 находится вдали от высокопроницаемого коридора в слабопроницаемой зоне.

В скв. № 1006 зарегистрированы 3 КПД. Принятая интерпретационная модель — модель вер-



Рис. 4. Основные результаты ГДИС на выкопировке из карты аномалий дуплексных волн.

Таблица 1

Результаты интерпретации ГДИС методом восстановления давления и методом исследования на режимах с непрерывной регистрацией изменения давления [15]

№ п/п	№ скв.	Вид исследования	Интерпретационная модель	Гидропроводность, Д, см/сПа	Проницаемость абсолютная/ текущая по нефти / текущая по воде, мД	Пьезопроводность, см ² /с	Скин-фактор	Расстояние до границы (границ), м
1	11Г	КВД1	Горизонтальная скважина, вскрывающая однородный бесконечный пласт с подошвенной водой	3.74	10.6/ 0.6/ 0.8	42.8	0.0 ¹	-
2		КВД2	Горизонтальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	3.75	10.7/ 0.6/ 0.8	43.5	0.0 ¹	51
3	11Г	КВД1	Горизонтальная скважина, вскрывающая однородный бесконечный пласт с подошвенной водой	3.65	10.4/ 0.6/ 0.7	42.4	0.0 ¹	-
4		КВД2	Горизонтальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	3.65	10.4/ 0.6/ 0.7	42.4	0.0 ¹	47
5	11Г	ДКИД1 ²	Горизонтальная скважина, вскрывающая однородный бесконечный пласт с подошвенной водой	-	25.4/ 1.5/ 1.8	-	0.0 ¹	-
6		ДКИД2 ²	Горизонтальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	-	27.1/ 1.6/ 1.9	-	0.0 ¹	56
7	13Г	КВД	Горизонтальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	12.51	42.2/ 5.2/ 2.1	170.9	0.3 ¹	54
8	13Г	ДКИД ²	Горизонтальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	-	71.5/ 8.8/ 3.6	-	0.7 ¹	48
9	1001	КВД	Вертикальная скважина, вскрывающая однородный бесконечный пласт	11.41	35.7/ 3.0/ 2.2	253.9	2.5	-
10	1001	ДКИД ²	Вертикальная скважина, вскрывающая однородный бесконечный пласт	-	34.1/ 2.9/ 2.0	-	1.7	-
11	1002	КВД1	Вертикальная скважина, вскрывающая однородный бесконечный пласт	10.38	49.2/ 12.0 1.3	140.4	- 0.1	-
12		КВД2	Вертикальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	11.35	53.8/ 13.1/ 1.5	153.6	0.6	655
13	1002	КВД1	Вертикальная скважина, вскрывающая однородный бесконечный пласт	11.17	53.0/ 12.9/ 1.4	151.1	- 0.2	-
14		КВД2	Вертикальная скважина, вскрывающая однородный пласт с линейной границей постоянного давления	11.31	53.7/ 13.1/ 1.4	153.0	- 0.1	682
15	1006	КПД	Вертикальная скважина, вскрывающая однородный пласт с двумя параллельными непроницаемыми границами	6054	28935/ 0/ 4022	248334	23.3	930, 1600
16	1006	КПД	Вертикальная скважина, вскрывающая однородный пласт с двумя параллельными непроницаемыми границами	5991	28633/ 0/ 3980	245740	22.6	930, 1400
17	1006	КПД	Вертикальная скважина, вскрывающая однородный бесконечный пласт	6100	29151/ 0/ 4052	250186	24.3	-
18	1006	КПД	Вертикальная скважина, вскрывающая однородный пласт с двумя параллельными непроницаемыми границами	6055	101295/ 0/ 14080	869353	22.7	1750, 3060
19	1006	КПД	Вертикальная скважина, вскрывающая однородный пласт с двумя параллельными непроницаемыми границами	5991	100216/ 0/ 13930	860092	22.0	1750, 2620
20	1006	КПД	Вертикальная скважина, вскрывающая однородный бесконечный пласт	6101	102050/ 0/ 14185	875836	23.7	-

¹ – скин-фактор на стенке горизонтального ствола.

² – длительная кривая изменения дебита и забойного давления, полученная путем исследования скважины на режимах с непрерывной регистрацией изменения давления (без остановки скважины).

тикальной скважины, вскрывающей однородный пласт с двумя параллельными непроницаемыми границами. Всего принято 2 варианта интерпретации результатов исследований скв. № 1006 по методу падения давления, отличающихся эффективной толщиной пласта. Рассмотрены следующие варианты:

Таблица 2

Результаты интерпретации ГДИС методом гидропрослушивания [15]

№ п/п	1	2	3	4	5
Возмущающая скважина	1006				
Реагирующая скважина	13Г		1001		1002
Расстояние между скважинами, м	2600		740		710
Средняя толщина пласта в межскважинном пространстве, м	15.4	10.4	18.7	13.7	Реакция не выявлена
Коэффициент гидропроводности, Д.см/сПз	127.8	127.8	84.2	84.2	
Пьезопроводность пласта, см ² /с	4487	6645	2461	3359	
Абсолютная проницаемость пласта, мД	806	1193	437	596	
Пористость, д.ед.	0.02	0.03	0.15	0.21	

1. Эффективная толщина пласта 14 м (согласно интервалу поглощения при бурении).
2. Эффективная толщина пласта 4 м (зона дробления по каверномеру).

Определенная по КПД проницаемость пласта в районе скв. № 1006 очень высока и составляет от нескольких десятков до сотни Д в зависимости от принимаемой эффективной толщины пласта. При эффективной толщине пласта 14 м расстояние до параллельных непроницаемых границ составляет 930 и 1400–1600 м. При эффективной толщине пласта 4 м расстояние до параллельных непроницаемых границ составляет 1750 и 2620–3060 м. Это соответствует возможному расположению границ нескольких близко расположенных высокопроницаемых коридоров, выявленных по карте аномалий дуплексных волн. Учитывая все вышесказанное и очень высокое значение абсолютной проницаемости пласта, можно сделать вывод о том, что скв. № 1006 находится в высокопроницаемом коридоре.

При интерпретации КВД (КПД) и ДКИД, зарегистрированных в скв. №№ 13Г, 1001, 1002 и 1006, получены следующие значения гидропроводности и пьезопроводности пласта:

1. Сква. № 13Г: гидропроводность 12.5 Д*см/сПз, пьезопроводность — 170.9 см²/с.
2. Сква. № 1001: гидропроводность 11.4 Д*см/сПз, пьезопроводность — 253.9 см²/с.
3. Сква. № 1002: гидропроводность 11.4 Д*см/сПз, пьезопроводность — 153.0 см²/с.
4. Сква. № 1006: гидропроводность 6054 Д*см/сПз, пьезопроводность — 248334 см²/с (при эффективной толщине 14 м) и 869353 см²/с (при эффективной толщине 4 м).

При интерпретации гидропрослушивания по схеме скв. № 1006 (возмущающая) – скв. №№ 13Г, 1001, 1002 (реагирующие) получены следующие значения гидропроводности и пьезопроводности пласта:

1. При гидропрослушивании по направлению 1006→13Г: гидропроводность 127.8 Д*см/сПз, пьезопроводность — 4487 см²/с (при эффективной толщине пласта в районе скважины № 1006 — 14 м) и 6645 см²/с (при эффективной толщине пласта в районе скважины № 1006 — 4 м).
2. При гидропрослушивании по направлению 1006→1001: гидропроводность 84.2 Д*см/сПз, пьезопроводность — 2461 см²/с (при эффективной толщине пласта в районе скважины № 1006 — 14 м) и 3359 см²/с (при эффективной толщине пласта в районе скважины № 1006 — 4 м).
3. При гидропрослушивании по направлению 1006→1002 реакции выявить не удалось.

Таким образом, значения гидропроводности и пьезопроводности в районе скв. №№ 13Г, 1001 и 1002, определенные по КВД и ДКИД, очень близки. Значения гидропроводности и пьезопроводности в районе скв. № 1006, определенные по КПД, на несколько порядков превосходят значения гидропроводности и пьезопроводности в районе скв. №№ 13Г, 1001 и 1002, определенные по КВД и ДКИД. Значения гидропроводности и пьезопроводности, определенные по результатам гидропрослушивания, находятся между значениями, полученными по КВД (ДКИД), зарегистрированными в скв. №№ 13Г, 1001, 1002, и значениями, полученными по КПД, зарегистрированными в скв. № 1006. Отсюда можно сделать вывод о том, что скв. № 13Г, 1001 и 1002 находятся в слабопроницаемой части пласта, скв. № 1006 находится внутри высокопроницаемого коридора, при этом скв. №№ 13Г, 1001 находятся вблизи

высокопроницаемого коридора.

Из приведенного анализа ГДИС можно сделать следующие выводы. Добывающие скв. №№ 11Г, 13Г и 1001 находятся вблизи высокопроницаемых коридоров в слабопроницаемой зоне, добывающая скв. № 1002 находится далеко от высокопроницаемых коридоров в слабопроницаемой зоне, нагнетательная скв. № 1006 находится внутри высокопроницаемого коридора. По результатам гидропрослушивания установлено, что направления развития высокопроницаемых коридоров в районе скв. №№ 13Г, 1001, 1002, 1006 соответствуют расположению зон с измененными свойствами на карте аномалий дуплексных волн.

Сопоставление КВД, ДКИД и гидропрослушивания

Совместная интерпретация КВД и ДКИД позволяет оценить согласованность получаемых фильтрационно-емкостных свойств. В целом, при правильном проведении исследований, основанном на предварительно рассчитанном дизайне, параметры проницаемости и проницаемости, полученные методами КВД и ДКИД, обычно находятся в пределах 15–20% расхождения. Однако ДКИД дают более устойчивые оценки при длительной фильтрации, особенно в скважинах с выраженным динамическим изменением дебита, в то время как КВД были чувствительнее к локальным нарушениям и позволяли точнее оценить скин-фактор и определить пластовое давление.

Сравнение результатов исследований скважин методом регистрации ДКИД и исследований методом восстановления давления для скв. №№ 11Г, 13Г и 1001 приведено в табл. 3.

Таблица 3

Сравнение результатов исследований скважин методом регистрации ДКИД и исследований методом регистрации КВД

Исследование	КВД	ДКИД	КВД	ДКИД	КВД	ДКИД
Скважина	11Г		13Г		1001	
Дебит жидкости, м ³ /сут	65		50		23	
Абсолютная проницаемость, мД	10.4	27.1	42.2	71.5	35.7	34.1
Расстояние до границы, м	47	56	54	45	-	-

Из табл. 3 видно, что значения расстояния до границ, определенные по результатам интерпретации КВД и ДКИД, близки, а значения абсолютной проницаемости существенно различаются, причем проницаемость, определенная по результатам интерпретации ДКИД, больше, что может объясняться большим радиусом исследования и попаданием в зону исследования значительной части высокопроницаемого коридора.

Сравнение с результатами гидропрослушивания подтвердило достоверность интерпретации. Зоны высокоамплитудных откликов гидропрослушивания совпадают с областями, в которых по ГДИС наблюдались направленные каналы фильтрации (особенно по скв. № 1006). Это усиливает уверенность в корректности выявления зоны как высокопроницаемого направления, вызванного трещиноватостью.

Таким образом, совместное использование КВД, ДКИД и данных гидропрослушивания обеспечивает согласованную картину распределения фильтрационно-емкостных свойств и позволяет уточнить пространственную неоднородность коллекторов, включая влияние тектонических нарушений, трещиноватости и латеральных барьеров.

Комплексирование с сейсмическими данными

Сопоставление интерпретированных по ГДИС фильтрационно-емкостных свойств с геофизическими и сейсмическими материалами позволило существенно повысить достоверность и детальность полученных моделей. Основным инструментом сопоставления служили карты аномалий дуплексных волн, построенные по результатам 3D-сеймики. Эти карты отражают зоны с повышенной трещиноватостью и возможными тектоническими нарушениями.

Наиболее яркий пример — скв. № 1006, расположенная в зоне аномалии дуплексных волн. По данным сеймики, здесь выявлены линейные аномалии протяженностью более 400 м, ориентированные в северо-западном направлении. Сопоставление с результатами ГДИС подтверждает сильно

повышенные ФЕС в районе скв. № 1006 по результатам КПД. Гидропрослушивание также выделяет высокопроницаемый канал от скв. № 1006 к скв. № 13Г. Таким образом, трещиноватость оказывает определяющее влияние на формирование направленных высокопроницаемых каналов.

По скв. №№ 11Г и 13Г наблюдается частичное совпадение расстояния до линейных границ постоянного давления, выявленных по КВД и ДКИД, с расстоянием до границы высокопроницаемого канала, выделенного по результатам построения карты аномалий дуплексных волн.

Таким образом, комплексирование ГДИС с сейсмическими материалами позволило повысить разрешающую способность интерпретации, уточнить геометрические параметры каналов фильтрации и подтвердить пространственное положение латеральных экранирующих объектов.

Комплексный подход к построению моделей

Комплексный подход к построению геолого-гидродинамических моделей нефтяных залежей становится обязательным в условиях высокой геологической сложности, особенно при работе с трещиновато-поровыми карбонатными коллекторами. Такие коллекторы характеризуются выраженной неоднородностью, наличием направленных трещин, латеральных экранов, зон повышенной или пониженной проницаемости и значительной анизотропией. В этих условиях информация, полученная по одному методу, зачастую оказывается недостаточной или интерпретируется неоднозначно. Поэтому интеграция различных источников данных — это не просто пожелание, а насущная необходимость.

Ключевыми компонентами для комплексирования являются: геофизические исследования скважин (ГИС), данные керн, результаты сейсмических исследований, а также данные ГДИС и промыслово-геофизических исследований (ПГИ). Каждый из этих видов предоставляет специфическую информацию о строении и свойствах коллектора. ГИС позволяют локализовать границы пропластков по вертикали, выявить продуктивные пропластки, определить их пористость и насыщенность. Керн предоставляет уникальные данные о текстуре, фациальных особенностях и трещиноватости, а также используется для лабораторных определений абсолютной проницаемости и относительных фазовых проницаемостей, дает возможность связать пористость и абсолютную проницаемость. Сейсмика, особенно при использовании атрибутивного анализа и методов дуплексных волн, предоставляет информацию о глубинах отражающих горизонтов (сред с разными свойствами), латеральных изменениях свойств, выявляет флюидоупоры, зоны трещиноватости и экраны.

Важнейшую роль играет сопоставление с результатами ГДИС, которые дают информацию о пьезо- и гидропроводности, скин-факторе, расстояниях до границ и характере границ. Особенная ценность ГДИС состоит в том, что они позволяют определять свойства пласта непосредственно в пластовых условиях, полностью аналогичных условиям разработки. КВД позволяет диагностировать отклонения от радиального притока, что может свидетельствовать о наличии трещин или границ. КПД уточняет параметры зоны закачки и позволяет оценить эффективность воздействия. ДКИД отражают отклик системы в реальном времени и особенно полезны в условиях нестабильной эксплуатации. Метод гидропрослушивания выявляет межскважинные связи и эффективную проницаемость в межскважинном пространстве.

Комплексирование этих данных позволяет строить более точные геолого-гидродинамические модели. Например, наличие аномалий дуплексных волн может указывать на флюидоупоры или трещиноватые зоны, а их сопоставление с зонами повышенной пьезопроводности по результатам ГДИС позволяет подтвердить или откорректировать интерпретацию. Совмещение карты аномалий с картой распределения проницаемости, полученной по результатам интерпретации КВД, КПД и ДКИД, позволяет уточнить границы продуктивных зон, выявить каналы ускоренной фильтрации и определить зоны, требующие интенсификации, правильно спланировать заводнение.

Таким образом, комплексный подход к интерпретации данных, полученных разными методами, существенно увеличивает достоверность оценки параметров пласта, минимизирует неопределенности и повышает эффективность решений, принимаемых при проектировании и оптимизации разработки. Особенно это актуально для карбонатных трещиновато-поровых коллекторов, в которых без комплексирования невозможно надежно оценить основные пути фильтрации и локализовать барьеры.

Заключение

Полученные результаты демонстрируют эффективность применения комплексного подхода к анализу фильтрационно-емкостных свойств сложнопостроенного карбонатного коллектора. Сравнение

методов КВД, ДКИД и данных гидропрослушивания позволяет говорить о высокой степени внутренней согласованности результатов исследований при условии корректного выбора моделей фильтрации и предварительного многовариантного дизайна исследований. Особо важно, что разные методы дополняют друг друга: КВД позволяет точно определить скин-фактор и пластовое давление, оценить проницаемость ближней зоны, ДКИД — устойчиво интерпретировать процессы в длительном режиме работы скважин, оценить размер дренируемого объема и его проницаемость, а гидропрослушивание — подтвердить направления фильтрации и существование гидродинамической связи.

Важным методическим результатом является демонстрация чувствительности результатов интерпретации к границам фильтрации и латеральным экранам. Ряд скважин демонстрирует явные признаки экранирования, что подтверждается не только ГДИС, но и сейсмическими материалами. Это указывает на необходимость применения направленного заводнения и отказа от стандартных схем разработки.

С точки зрения анализа структуры пласта доказано, что зоны, выделенные по аномалиям дуплексных волн, действительно, совпадают с высокопроницаемыми участками. Это делает возможным использование сейсмики не только для качественной интерпретации, но и для количественного прогноза ФЕС при должной калибровке по результатам ГДИС. Такой подход может быть применен и для других месторождений с трещиновато-поровым строением.

Наконец, результаты подтверждают необходимость продолжения исследований по интеграции ГДИС с методами цифровой интерпретации сейсмических данных. Построение интегрированных моделей трещиноватости и проницаемости позволит значительно повысить эффективность заводнения и управления разработкой.

ЛИТЕРАТУРА

1. Кульпин Л. Г., Мясников Ю. А. *Гидродинамические методы исследования нефтегазоводоносных пластов*. М.: Недра; 1974. 200 с.
2. Деева Т. А., Камартдинов М. Р., Кулагина Т. Е. и др. *Гидродинамические исследования скважин: анализ и интерпретация данных*. Томск: ЦППС НД ТПУ; 2009. 243 с. ISBN: 5-98298-048-X.
3. Эрлагер Р. *Гидродинамические методы исследования скважин*. М.-Ижевск: ИКИ; 2006. 512 с. ISBN 5-93972-521-X.
4. Славкин В. С. *Геолого-геофизическое изучение нефтяных продуктивных отложений*. М.: Изд-во МГУ; 1999. 160 с.
5. Еременко Н. А., Чилингар Г. В. *Геология нефти и газа на рубеже веков*. М.: Наука; 1996. 176 с.
6. Yielding G., Freeman B., Needham D. Quantitative Fault Seal Prediction. *AAPG Bulletin*. 1997;81:897-917. DOI: 10.1306/522B498D-1727-11D7-8645000102C1865D.
7. Несмеянов С. А. *Инженерная геотектоника*. М.: Наука; 2012. 560 с. ISBN: 5-02-033051-5.
8. РД 153-39.0-109-01. Методические указания по комплексированию и этапности выполнения геофизических, гидродинамических и геохимических исследований нефтяных и нефтегазовых месторождений.
9. Bourdet D., Ayoub J. A., Pirard Y. M. Use of Pressure Derivative in Well-Test Interpretation. *SPE Formation Evaluation*. 1989;6:293-302. Режим доступа: [https://blasingame.engr.tamu.edu/z_zCourse_Archive/P648_19A/P648_19A_Reading_Portfolio/SPE_012777_\(Bourdet\)_Pressure_Derivative_for_PTA_\(OCR\)_pdf.pdf](https://blasingame.engr.tamu.edu/z_zCourse_Archive/P648_19A/P648_19A_Reading_Portfolio/SPE_012777_(Bourdet)_Pressure_Derivative_for_PTA_(OCR)_pdf.pdf).
10. Bourdet D. *Well Test Analysis: The Use of Advanced Interpretation Models*. Amsterdam: Elsevier; 2002. 436 p. ISBN: 0-444-50968-2.
11. Houze O., Viturat D., Fjaere O. S. et al. *Dynamic Data Analysis*. V5.60. Kappa Engineering; 2024. 788 p. Режим доступа: <https://www.kappaeng.com/downloads/ddabook>.
12. Кременецкий М. И., Ипатов А. И. *Применение промыслово-геофизического контроля для оптимизации разработки месторождений нефти и газа. Т. 1. Основы гидродинамико-геофизического контроля разработки и мониторинга добычи*. М.-Ижевск: ИКИ; 2020. 676 с. ISBN 978-5-4344-0886-8.
13. Басниев К. С., Кочина И. Н., Максимов В. М. *Подземная гидродинамика*. М.: Недра; 1993. 416 с.

14. Хромова И. Ю. Миграция дуплексных волн — метод картирования трещиноватых зон тектонического генезиса. *Геология нефти и газа*. 2008;3:37-47. Режим доступа: <https://www.elibrary.ru/item.asp?id=11765261>.
15. Суходанова С. С. *Создание 3D модели залежи с карбонатными трещиноватыми коллекторами на основе комплексирования гидродинамических, геофизических, сейсмических и промысловых данных (на примере нижнепермских отложений Варандейского месторождения)*: дис. . . . канд. техн. наук. М.: ИПНГ РАН; 2016. 157 с. Режим доступа: <https://archive.ipng.ru/council/dissertations/5846>.

О ДИНАМИКЕ ПРИМЕСИ В ПОЛЕ ТЕЧЕНИЯ СТРАТИФИЦИРОВАННОЙ ВЯЗКОЙ НЕСЖИМАЕМОЙ ЖИДКОСТИ

В. А. Галкин

¹ Сургутский филиал федерального государственного автономного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Сургут, Российская Федерация

² Сургутский государственный университет, г. Сургут, Российская Федерация

ORCID: <http://orcid.org/0000-0002-9721-4026>,  val-gal@yandex.ru

Посвящается светлой памяти
Сергея Григорьевича Вольпина

Аннотация: исследуется структура решения уравнения переноса примеси под действием вихревого решения уравнения Навье–Стокса в случае течения несжимаемой жидкости в плоскостях, ортогональных выделенной оси — «скважине». Приведено описание структуры таких решений в окрестности скважины, порожденных поворотами в инвариантных плоскостях течения. Установлено, что динамика примеси в таких течениях может порождать пространственную структуру, аналогичную явлению кавитации для быстро движущихся объектов вдоль выделенной оси, порождающих узкую зону разрежения примеси. Приведены примеры образования таких структур, обладающих осевой симметрией.

Ключевые слова: уравнения несжимаемой жидкости, точные решения системы Навье–Стокса, уравнение переноса.

Благодарности: работа выполнена при финансовой поддержке в рамках государственного задания НИЦ «Курчатовский институт» — НИИСИ (Выполнение фундаментальных научных исследований ГП 47) по теме № 0580-2021-0007 «Развитие методов математического моделирования распределенных систем и соответствующих методов вычисления». Автор выражает благодарность за обсуждение работы сотрудникам Сургутского филиала НИЦ «Курчатовский институт» — НИИСИ А. О. Дубовику, Т. В. Гавриленко, Д. А. Моргуну.

Для цитирования: Галкин В. А. О динамике примеси в поле течения стратифицированной вязкой несжимаемой жидкости. *Успехи кибернетики*. 2025;6(2):22–26.

Поступила в редакцию: 01.06.2025.

В окончательном варианте: 16.06.2025.

DYNAMIC BEHAVIOR OF IMPURITIES IN THE FLOW FIELD OF A STRATIFIED VISCOUS INCOMPRESSIBLE FLUID

V. A. Galkin

¹ Surgut Branch of Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, Surgut, Russian Federation

² Surgut State University, Surgut, Russian Federation

ORCID: <http://orcid.org/0000-0002-9721-4026>,  val-gal@yandex.ru

Abstract: we studied the structure of the solution to the impurity transport equation under the influence of a vortex-type solution of the Navier–Stokes equations for incompressible fluid flow in planes orthogonal to a distinguished axis, referred to as the “well”. We described the structure of these solutions in the vicinity of the well, which arises due to rotational motion in the invariant flow planes. Our analysis shows that the dynamics of impurities in such flows can generate a spatial structure resembling cavitation, observed in high-speed motion along the selected axis. This phenomenon results in the formation of a narrow region of impurity dilution. We also provided examples demonstrating the formation of such axially symmetric structures.

Keywords: incompressible fluid equations, exact solutions of the Navier-Stokes system, transport equation.

Acknowledgements: this study is a part of the government contract 47 GP with the Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, project No. 0580-2021-0007 Advancing Distribution System Simulation and Computation Methods. The author expresses his gratitude for the discussion of the work to A.O. Dubovik, T.V. Gavrilenko, and D.A. Morgun

of the Surgut Branch of Scientific Research Institute for System Analysis of the National Research Centre "Kurchatov Institute".

Cite this article: Galkin V. A. Dynamic Behavior of Impurities in the Flow Field of a Stratified Viscous Incompressible Fluid. *Russian Journal of Cybernetics*. 2025;6(2):22–26.

Original article submitted: 01.06.2025.

Revision submitted: 16.06.2025.

Введение

В [1–5] приведены содержательные результаты, относящиеся к теме настоящей работы.

В координатном пространстве $\mathbb{R}_3 = \{x = (x_1, x_2, x_3)\}$ рассматривается динамика во времени $t \in \mathbb{R}$ гладкого поля скоростей $V : \mathbb{R}_3 \times \mathbb{R} \rightarrow \mathbb{R}_3$ ($V \in C^2$), удовлетворяющего системе Навье–Стокса для несжимаемой жидкости

$$\frac{\partial V}{\partial t} + (V \cdot \nabla) V + \frac{1}{\bar{\rho}} \nabla P(x, t) = \varepsilon^2 \Delta V, \quad (1)$$

$$\operatorname{div} V = 0, \quad (2)$$

где $P : \mathbb{R}_3 \times \mathbb{R} \rightarrow \mathbb{R}$ — давление жидкости, $\bar{\rho}$ и ε^2 — постоянные, характеризующие плотность и вязкость жидкости. В настоящей работе рассматривается класс решений $\{V, P\}$, обладающих свойством плоской стратификации относительно некоторой оси $\mathcal{L}$, которую без потери общности направим вдоль вектора $\bar{e}_3 = (0, 0, 1)$ ($\{\bar{e}_i\}_{i=1}^3$ — ортонормированный базис в $\mathbb{R}_3$). Проекцию вектора x на ось $\mathcal{L}$ обозначим в дальнейшем $x_3 = (x, \bar{e}_3)$, положим $x_i \stackrel{\text{def}}{=} (x, \bar{e}_i)$. Свойство стратификации означает, что выполнено соотношение ортогональности

$$(V, \mathcal{L}) = 0. \quad (3)$$

Обозначим величину $r(x_1, x_2) \stackrel{\text{def}}{=} \sqrt{x_1^2 + x_2^2}$. В дальнейшем наряду с декартовыми координатами будем также использовать цилиндрическую систему координат (r, φ, z) .

Ниже решения уравнений (1), (2) рассматриваются в пространственной области $M = \{x \in \mathbb{R}_3 \setminus \mathcal{L} : a < x_3 < b \text{ при значениях времени } t > 0\}$. Отметим, что на оси $\mathcal{L}$, вообще говоря, могут быть сосредоточены особенности рассматриваемых течений, приводящие к обобщенным функциям Дирака. Поэтому, чтобы в рамках настоящей работы не выходить из пространства гладких классических функций, ось («скважина» $\mathcal{L}$) исключается из рассмотрения. Отметим, что расширенное исследование, включающее ось $\mathcal{L}$, требует привлечения теории функциональных решений [6].

Предположим, что в слое M распределена невесомая безынерционная примесь, увлекаемая полем гидродинамического течения V , которая описывается концентрацией $n(x, t)$, подчиняющейся линейному уравнению переноса

$$\frac{\partial n(x, t)}{\partial t} + (V, \nabla n) = 0, \quad x \in M, \quad t > 0. \quad (4)$$

Подчеркнем, что явление диффузии в уравнение (4) не включено, чтобы выявить наиболее существенные детали рассматриваемых ниже явлений.

Настоящая работа является продолжением исследований автора [1–5].

Вихревое стратифицированное гидродинамическое поле

Обозначим $T_{x_3} : \mathbb{R}_3 \rightarrow \mathbb{R}_3$, $x_3 \in \mathbb{R}$, однопараметрическое семейство поворотов вокруг оси $\mathcal{L}$, заданное формулами

$$T_{x_3}(x) = \begin{pmatrix} \cos(x_3) & \sin(x_3) & 0 \\ -\sin(x_3) & \cos(x_3) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \quad (5)$$

Теорема. Пусть векторное поле U на множестве $x \in M$ задано соотношением

$$U(x) = r^{-2}(x) T_{x_3}(x). \quad (6)$$

Тогда поле

$$V(x, t) = U(x) \exp(-\varepsilon^2 t) \quad (7)$$

удовлетворяет на M при $t > 0$ системе уравнений Навье–Стокса (1), (2) с условием стратификации (3), при этом значения давления $P(x, t)$ заданы формулой

$$P(x, t) = -\frac{\bar{\rho}}{2}(V(x, t), V(x, t)) + \beta(t), \quad (8)$$

где $\beta(t)$ — произвольная функция, зависящая от времени $t \in \mathbb{R}$.

Утверждение теоремы проверяется непосредственной подстановкой формул (6)–(8) в соотношения (1)–(3).

Поскольку преобразование $T_{x_3} : \mathbb{R}_3 \rightarrow \mathbb{R}_3$, $x_3 \in \mathbb{R}$ является периодическим по аргументу x_3 , то таким же свойством обладает гидродинамическое поле $V(x, t)$, и это поле является касательным к каждому слою $x_3 = \text{const}$.

Решения системы (1)–(4)

Так как решения уравнения переноса (4) сохраняют постоянное значение вдоль характеристик, то значения концентрации $n(x, t)$ получаются переносом начального распределения примеси $n(x, 0)$ в области $M = \mathbb{R}_3 \setminus \mathcal{L}$ вдоль траекторий системы обыкновенных дифференциальных уравнений

$$\frac{dx}{d\tau} = U(x), \quad \tau(t) = \varepsilon^{-2}(1 - \exp(-\varepsilon^2 t)). \quad (9)$$

Переходя в системе (9) к цилиндрическим координатам, получаем тождества

$$\begin{aligned} r^2(\tau) &= r^2(0) + 2\tau \cos(x_3); \\ \varphi(\tau) &= \varphi(0) - \frac{1}{2}tg(x_3) \ln \left(\frac{r^2(\tau)}{r^2(0)} \right), \quad x_3 \neq \pm \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}; \\ x_3(\tau) &= x_3(0). \end{aligned} \quad (10)$$

При значениях $x_3(0) = \pm \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$ траектории характеристик, задаваемых системой (9), в цилиндрической системе координат

$$\begin{aligned} r^2(\tau) &= r^2(0); \\ \varphi(\tau) &= \varphi(0) \mp \tau r^{-2}(0), \quad x_3 = \pm \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}; \\ x_3(\tau) &= x_3(0). \end{aligned} \quad (11)$$

Отметим, что формулы (10) определяют спиральную форму траекторий в фазовой плоскости (r, φ) , $x_3 = x_3(0)$, которые «раскручиваются» от оси $\mathcal{L}$ при положительных значениях $\cos(x_3)$ и, напротив, «накручиваются» на нее при отрицательных величинах $\cos(x_3)$. (В частном случае, когда $x_3 = k\pi$, $k \in \mathbb{Z}$, семейство спиралей превращается в пучок прямых исходящих из $\mathcal{L}$ либо сходящихся к $\mathcal{L}$.) Когда же $\cos(x_3) = 0$, то траектории лежат на концентрических окружностях с центром в начале координат фазовой плоскости. Это определяет специальные ограничения на начальные данные задачи Коши (4), ибо необходимы условия согласования для корректности задачи, поскольку должно выполняться требование постоянства начальной функции на упомянутых окружностях в этих плоскостях.

Таким образом, при достаточно малых положительных начальных значениях величины $r(0)$, когда $\cos(x_3) < 0$, траектория характеристики попадает за конечное время на ось $\mathcal{L}$. Тем самым в таких точках на оси $\mathcal{L}$ концентрация примеси переходит в обобщенную функцию Дирака в плоскости (x_1, x_2) , $x_3 = x_3(0)$, пропорциональную количеству вещества примеси, попавшему в указанную точку из рассматриваемой фазовой плоскости.

Для «раскручивающихся» траекторий в плоскостях, где $\cos(x_3) > 0$, из начального условия $n(x, 0)|_{x \in \mathcal{L}} = 0$ получаем, что $n(x, t) = 0$ в круге $r \leq \sqrt{2\tau \cos(x_3)}$. Таким образом, вдоль вертикальной оси $\mathcal{L}$ в плоскостях с условием $\cos(x_3) > 0$ формируются периодически расположенные «каверны», в которых отсутствует примесь. Указанное явление качественно весьма напоминает образование кавитационных пузырей в жидкости вдоль траектории быстро движущихся тел. См., например, рис. 1, 2, иллюстрирующие высокоэнергичные процессы переноса.



The Slow Mo Guys на Русском (Выпуск # 3) - Стрельба под водой

Рис. 1. Стрельба под водой (из открытых источников сети Интернет)



Рис. 2. Высокоэнергетический взрывной процесс в атмосфере (из открытых источников сети Интернет)

Заключение

Приведено описание структуры решений уравнения переноса в вихревом потоке стратифицированной вязкой несжимаемой жидкости в окрестности скважины, порожденных осевыми вращениями в плоскостях, ортогональных скважине. Установлено, что динамика примеси в таких течениях может порождать пространственную структуру, аналогичную явлению кавитации для быстро движущихся объектов вдоль выделенной оси, порождающих узкую зону разрежения примеси.

ЛИТЕРАТУРА

1. Галкин В. А. Об одном классе точных решений системы Навье–Стокса для несжимаемой жидкости в шаре и сферическом слое. *Ж. вычисл. матем. и матем. физ.* 2023;63(6):1000–1005.
2. Галкин В. А., Дубовик А. О. Об одном классе точных решений системы уравнений Навье–Стокса для несжимаемой жидкости. *Матем. моделирование.* 2023;35(8):3–13.
3. Галкин В. А. О структуре винтовых осесимметричных решений системы Навье–Стокса для несжимаемой жидкости. *Ж. вычисл. матем. и матем. физ.* 2024;64(5):780–790.
4. Галкин В. А., Смородинов А. Д., Моргун Д. А. Решение уравнения Навье–Стокса для сталкивающихся потоков. *Успехи кибернетики.* 2023;4(2):8–15.
5. Галкин В. А., Дубовик А. О. О моделировании слоистого течения вязкой проводящей жидкости в области, изменяющейся во времени. *Матем. моделирование.* 2020;32(4):31–42.
6. Галкин В. А. *Анализ математических моделей. Системы законов сохранения, уравнения Больцмана и Смолуховского.* М.: БИНОМ; 2009. 428 с.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЛЯ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ РАЗРАБОТКИ НЕФТЯНЫХ МЕСТОРОЖДЕНИЙ: ЕМКОСТНО-РЕЗИСТИВНАЯ МОДЕЛЬ БЕЗ УЧЕТА ЗАБОЙНОГО ДАВЛЕНИЯ И ЛОГАРИФМИЧЕСКАЯ МОДЕЛЬ ОБВОДНЕНИЯ СКВАЖИНЫ

И. В. Афанаскин^{1,а}, С. Г. Вольпин^{2,б}, А. В. Королев^{2,в}, М. Ю. Ахапкин^{2,г}

¹ Государственное автономное образовательное учреждение высшего образования Альметьевский государственный технологический университет «Высшая школа нефти», г. Альметьевск, Российская Федерация

² Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Москва, Российская Федерация

^а ORCID: <http://orcid.org/0000-0003-0617-221X>,  ivan@afanaskin.ru

^б ORCID: <http://orcid.org/0000-0002-3591-2910>, sergvolpin@gmail.com

^в ORCID: <http://orcid.org/0000-0001-9163-9106>, aleksandre.korolev@mail.ru

^г ORCID: <http://orcid.org/0000-0002-8482-3236>, makhapkin@niisi.ras.ru

Аннотация: в работе предлагается модифицированная CRM-модель (Capacitance-Resistive Model), исключающая использование забойного давления. Модель может применяться в условиях заводнения или при упруговодонапорном режиме разработки. В отличие от традиционных подходов, модель основана на учете накопленных объемов добычи и закачки и описывает межскважинное взаимодействие добывающих скважин с включением влияния аквифера. Модель является полуаналитической. Дополнительно введена эмпирическая логарифмическая модель обводнения, учитывающая зависимость водонасыщенности от накопленной добычи жидкости. Полученные уравнения позволяют рассчитывать дебиты жидкости и показатели обводнения без использования забойного давления, что существенно расширяет область применимости модели в условиях ограниченного промыслового мониторинга. Модель протестирована на данных по четырем добывающим скважинам месторождения с карбонатным коллектором, эксплуатируемого при упруговодонапорном режиме. Сравнение результатов расчетов по стандартной и модифицированной CRM-моделям показало приемлемое качество аппроксимации дебитов жидкости при исключении давления, а также высокую точность модели обводнения. Предложенный подход может быть применен для анализа разработки и прогнозирования без необходимости в труднодоступных параметрах, таких как забойное давление, а также для оперативной интерпретации промысловой информации.

Ключевые слова: математические модели, прогнозирование разработки нефтяных месторождений, CRM-модель, модель обводнения, упруговодонапорный режим разработки.

Благодарности: публикация выполнена в рамках государственного задания НИЦ «Курчатовский институт» — НИИСИ по теме № FNEF-2024-0001 «Создание и реализация доверенных систем искусственного интеллекта, основанных на новых математических и алгоритмических методах, моделях быстрых вычислений, реализуемых на отечественных вычислительных системах».

Для цитирования: Афанаскин И. В., Вольпин С. Г., Королев А. В., Ахапкин М. Ю. Математические модели для анализа и прогнозирования разработки нефтяных месторождений: емкостно-резистивная модель без учета забойного давления и логарифмическая модель обводнения скважины. *Успехи кибернетики*. 2025;6(2):27–39.

Поступила в редакцию: 23.05.2025.

В окончательном варианте: 11.06.2025.

MATHEMATICAL MODELS FOR THE ANALYSIS AND FORECASTING OF OIL FIELD DEVELOPMENT: A CAPACITIVE-RESISTIVE MODEL WITHOUT BOTTOMHOLE PRESSURE AND A LOGARITHMIC MODEL OF WELL WATER BREAKTHROUGH

I. V. Afanaskin^{1,а}, S. G. Volpin^{2,б}, A. V. Korolev^{2,в}, M. Yu. Ahapkin^{2,г}

¹ Almet'yevsk State Technological University, Petroleum Higher School, Almet'yevsk, Russian Federation

² Scientific Research Institute for System Analysis of the National Research Centre Kurchatov Institute, Moscow, Russian Federation

^а ORCID: <http://orcid.org/0000-0003-0617-221X>,  ivan@afanaskin.ru

^б ORCID: <http://orcid.org/0000-0002-3591-2910>, sergvolpin@gmail.com

^c ORCID: <http://orcid.org/0000-0001-9163-9106>, aleksandre.korolev@mail.ru

^d ORCID: <http://orcid.org/0000-0002-8482-3236>, makhapkin@niisi.ras.ru

Abstract: we developed a modified Capacitance-Resistive Model (CRM) that does not require bottomhole pressure data. This semi-analytical model applies to waterflooding and elastic-water-drive reservoir development. Unlike traditional CRM formulations, our approach relies on cumulative production and injection volumes and incorporates inter-well interactions along with the effects of aquifers.

We also introduced an empirical logarithmic model for water breakthrough, which relates water saturation to cumulative liquid production. The resulting equations enable the calculation of liquid flow rates and water breakthrough without bottomhole pressure, significantly broadening the model's applicability in fields with limited monitoring data.

We tested the model on four production wells in a carbonate reservoir under elastic-water-drive conditions. Comparisons with the standard CRM model demonstrated acceptable accuracy in flow rate prediction without pressure input and high precision in modeling water breakthrough dynamics. This approach supports both development forecasting and real-time field analysis without relying on hard-to-obtain parameters such as bottomhole pressure.

Keywords: mathematical models, oil field development forecasting, CRM model, water breakthrough model, elastic water drive.

Acknowledgements: this is a part of the FNEF-2024-0001 Development and Deployment of Trusted AI Systems based on New Mathematical and Algorithmic Approaches and Fast Computing Models Compatible with Domestic Computer Hardware government contract.

Cite this article: Afanaskin I. V., Volpin S. G., Korolev A. V., Ahapkin M. Yu. Mathematical Models for the Analysis and Forecasting of Oil Field Development: A Capacitive-Resistive Model Without Bottomhole Pressure and a Logarithmic Model of Well Water Breakthrough. *Russian Journal of Cybernetics*. 2025;6(2):27–39.

Original article submitted: 23.05.2025.

Revision submitted: 11.06.2025.

Введение

Capacitance-Resistive Model (CRM) — это класс математических моделей, состоящий из ёмкостно-резистивных моделей, предназначенных для описания взаимодействия между нагнетательными и добывающими скважинами в процессе разработки нефтяных месторождений. Эти модели основаны на принципах материального баланса и представляют пласт как систему, преобразующую закачку в отклик на добычу. Математическая структура CRM включает уравнение сохранения массы, отвечающее за ёмкость, и уравнение притока, описывающее сопротивление между скважинами. Такая формализация позволяет воспроизводить динамику межскважинных потоков без построения полной геологической модели, что делает CRM-методы особенно привлекательными для оперативной оценки состояния системы поддержания пластового давления (ППД).

В зависимости от постановки задачи и уровня детализации модели выделяют различные модификации CRM: CRMT [1, 2, 7, 8, 11–13]; CRMP [2, 6–9, 11–13]; CRMIP [3, 7, 8, 12]; CRM-Block [7, 8]; ML-CRM [7]; CRMT-M, CRMP-M и CRMIP-M [5, 8] и другие. Ключевое преимущество этих моделей — возможность обходиться без регулярных замеров пластового давления, которое в промышленной практике определяется редко и с большой неопределённостью. Благодаря этому CRM-модели широко применяются для оценки дренируемых объёмов, анализа межскважинной интерференции, краткосрочного прогнозирования и оптимизации системы заводнения.

Отдельное направление в развитии CRM-моделирования составляют подходы, полностью исключаящие из расчетов забойное давление. Такие модели становятся особенно актуальными в условиях, где замеры давления отсутствуют, нестабильны или нерепрезентативны. В ряде современных публикаций представлены примеры успешного применения CRM-моделей именно в таком формате. Так, в работе [15] показано, что классическая CRM-модель может быть эффективно адаптирована к реальным данным без учета забойных давлений, с получением высокой точности прогноза дебитов на основе только объёмов закачки и добычи.

В статье [16] предлагается модифицированная многоскважинная модель (MCCM), способная учитывать переменный скин-фактор и остановки добывающих скважин. Хотя исходная формулировка

модели включает забойное давление, авторы подчеркивают, что она может быть использована и без давления, что особенно важно для полевых данных с частыми остановками и неполными измерениями.

В отечественной практике примером такого подхода является метод Cost-БФА, представленный в [17], в котором применяется упрощенная CRM-модель, полностью исключая давление из расчетной схемы. Эта модель интегрирована с экономической оценкой и предназначена для оптимизации системы ППД при помощи анализа межскважинной связи на основе исторических данных закачки и добычи.

В настоящей работе рассматривается развитие этого направления: предложена CRM-модель, построенная без использования забойного давления, адаптированная к условиям упруговодонапорного режима и дополненная логарифмической моделью обводнения скважин. Рассматриваемый подход позволяет анализировать межскважинную связь и динамику обводненности без необходимости в труднодоступных параметрах, обеспечивая при этом достаточную точность прогноза и применимость в инженерной практике.

Математическая модель

Изучим CRM-модель для следующих специальных условий: рассматривается единый поровый объем для каждой добывающей скважины, на процессы в котором оказывают влияние окружающие скважины; учитывается влияние как добывающих, так и нагнетательных скважин; учитывается влияние аквифера (с каждой добывающей скважиной взаимодействует отдельный участок аквифера — водоносной области, взаимодействие участков аквифера не учитывается). Такую модель можно рассматривать как модификацию модели CRMP. Остальные условия — стандартные для CRMP-модели [2, 6–9, 11–13].

Запишем систему уравнений сохранения, описывающую такую CRM-модель:

$$c_{t,i} V_{p,i} \frac{dp_i}{dt} = -q_{l,i}(t) - \sum_{j=1}^{N_i} g_{ij} q_{l,j}(t) + \sum_{j=1}^{M_i} f_{ij} \omega_j(t) + q_{a,i}(t), \quad (1)$$

$$q_{l,i}(t) = PI_i [p_i(t) - p_{w,i}(t)], \quad (2)$$

$$q_{o,i}(t) = q_{l,i}(t) [1 - f_{w,i}(t)], \quad (3)$$

$$q_{w,i}(t) = q_{l,i}(t) f_{w,i}(t), \quad (4)$$

$$c_{ta,i} V_{pa,i} \frac{dp_{a,i}}{dt} = -q_{a,i}(t), \quad (5)$$

$$q_{a,i}(t) = PI_{a,i} [p_{a,i}(t) - p_i(t)], \quad (6)$$

где $c_{t,i}$ и $V_{p,i}$ — суммарная сжимаемость поровой среды и жидкости, поровый объем области дренирования i -ой скважины соответственно; $p_i(t)$ — пластовое давление в объеме дренирования i -ой скважины; $q_{l,i}(t)$, $q_{l,j}(t)$, $\omega_j(t)$ и $q_{a,i}(t)$ — дебит жидкости i -ой и j -ой добывающих скважин, расход закачиваемой воды j -ой нагнетательной скважины, приток воды из законтурной области (аквифера) в объем дренирования i -ой добывающей скважины соответственно; PI_i и $p_{w,i}(t)$ — коэффициент продуктивности по жидкости и забойное давление i -ой скважины соответственно; $q_{o,i}(t)$, $q_{w,i}(t)$ и $f_{w,i}(t)$ — дебит нефти и воды, обводненность i -ой скважины соответственно; $c_{ta,i}$ и $V_{pa,i}$ — полная сжимаемость системы пласт-вода и поровый объем участка аквифера, взаимодействующего с поровым объемом i -скважины, соответственно; $PI_{a,i}$ и $p_{a,i}(t)$ — аналог коэффициента продуктивности и пластовое давление участка аквифера, взаимодействующего с поровым объемом i -скважины, соответственно.

Здесь и ниже все дебиты, расходы, притоки и пр. записываются при пластовых давлении и температуре.

Уравнения сохранения объема жидкости и воды в пластовых условиях — уравнения (1) и (5) соответственно. Фактически это уравнения материального баланса для объема дренирования скважины и связанного с ним объема законтурной водоносной области. Уравнения притока жидкости, воды и нефти к скважине, уравнение притока воды из аквифера — уравнения (2)–(4) и (6) соответственно.

Система уравнений (1)–(6) дополняется начальными условиями:

$$p_{w,i}(t=0) = p_{a,i}(t=0) = p_{0,i}, \quad (7)$$

$$q_{l,i}(t=0) = 0. \quad (8)$$

Выразим из (2) $p_i(t)$ и продифференцируем:

$$\frac{dp_i}{dt} = \frac{dp_{w,i}}{dt} + \frac{1}{PI_i} \frac{dq_{l,i}}{dt}. \quad (9)$$

Подставим (9) в (1) и проинтегрируем по времени от нуля до момента t :

$$\begin{aligned} c_{t,i} V_{p,i} \left(\int_{p_{0,i}}^{p_{w,i}(t)} dp_{w,i} + \frac{1}{PI_i} \int_0^{q_{l,i}(t)} dq_{l,i} \right) = \\ = - \int_0^t q_{l,i}(\tau) d\tau - \sum_{j=1}^{Ni} \left(g_{ij} \int_0^t q_{l,j}(\tau) d\tau \right) + \sum_{j=1}^{Mi} \left(f_{ij} \int_0^t w_j(\tau) d\tau \right) \sum_{j=1}^{Mi} f_{ij} w_j(t) + \\ + \int_0^t q_{a,i}(\tau) d\tau. \end{aligned} \quad (10)$$

Совершая элементарные преобразования, получим выражение для $p_{w,i}(t)$:

$$p_{w,i}(t) = p_{0,i} - \frac{Q_{l,i}(t) + \sum_{j=1}^{Ni} g_{ij} Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij} W_j(t) - Q_{a,i}(t)}{c_{t,i} V_{p,i}} - \frac{q_{l,i}(t)}{PI_i}, \quad (11)$$

где $Q_{l,i}(t)$, $Q_{l,j}(t)$, $W_j(t)$ и $Q_{a,i}(t)$ — накопленные добыча жидкости, закачка воды и приток воды из аквифера по скважинам соответственно.

С учетом формул (3) и (4), считая $q_{\alpha,i}(t) \rightarrow q_{\alpha,i}(t - \Delta t)$ при $\Delta t \rightarrow 0$, где $\alpha = o, w$, перепишем (11):

$$p_{w,i}(t) = p_{0,i} - \frac{Q_{l,i}(t) + \sum_{j=1}^{Ni} g_{ij} Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij} W_j(t) - Q_{a,i}(t)}{c_{t,i} V_{p,i}} - \frac{q_{o,i}(t - \Delta t)}{2PI_{o,i}} - \frac{q_{w,i}(t - \Delta t)}{2PI_{w,i}}, \quad (12)$$

где $PI_{o,i}$ и $PI_{w,i}$ — коэффициенты продуктивности i -скважины по нефти и воде, которые для расчета забойного давления и дебита жидкости будем полагать постоянными и не зависящими от обводненности.

Считая $\Delta t \rightarrow 0$, примем $Q_{\alpha,i}(t) \rightarrow Q_{\alpha,i}(t - \Delta t)$. Кроме того:

$$Q_{l,i}(t) = Q_{l,i}(t - \Delta t) + \Delta t q_{l,i}(t). \quad (13)$$

Тогда выразим из (11) и (12) дебит жидкости i -ой скважины $q_{l,i}(t)$:

$$q_{l,i}(t) = \frac{p_{0,i} - \frac{Q_{l,i}(t - \Delta t) + \sum_{j=1}^{Ni} g_{ij} Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij} W_j(t) - Q_{a,i}(t - \Delta t)}{c_{t,i} V_{p,i}} - p_{w,i}(t)}{\frac{1}{PI_i} + \frac{\Delta t}{c_{t,i} V_{p,i}}}, \quad (14)$$

$$q_{l,i}(t) = \frac{p_{0,i} - \frac{Q_{l,i}(t - \Delta t) + \sum_{j=1}^{Ni} g_{ij} Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij} W_j(t) - Q_{a,i}(t - \Delta t)}{c_{t,i} V_{p,i}} - p_{w,i}(t)}{\frac{1 - f_{w,i}(t - \Delta t)}{PI_{o,i}} + \frac{f_{w,i}(t - \Delta t)}{PI_{w,i}} + \frac{\Delta t}{c_{t,i} V_{p,i}}}. \quad (15)$$

Проинтегрируем (5) по времени от 0 до t . Получим выражение для давления на участке законтурной водоносной области:

$$p_{a,i}(t) = p_{0,i} - \frac{Q_{a,i}(t)}{c_{ta,i} V_{pa,i}}. \quad (16)$$

С помощью (6) и (16) получим выражение для притока воды $q_{a,i}(t)$:

$$q_{a,i}(t) = \frac{\frac{PI_{a,i}}{c_{t,i}V_{p,i}} \left[Q_{l,i}(t) + \sum_{j=1}^{Ni} g_{ij}Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij}W_j(t) \right] - PI_{a,i}\Omega_i Q_{a,i}(t - \Delta t)}{1 + PI_{a,i}\Omega_i\Delta t}, \quad (17)$$

где

$$\Omega_i = \frac{1}{c_{ta,i}V_{pa,i}} + \frac{1}{c_{t,i}V_{p,i}}. \quad (18)$$

Подставим (11) в (14) и (12) в (15) и получим две формулы для дебита жидкости $q_{l,i}(t)$ без участия забойного давления $p_{w,i}(t)$:

$$q_{l,i}(t) = \frac{PI_i}{c_{t,i}V_{p,i} + \Delta t PI_i} \left[Q_{l,i}(t - \Delta t) + \sum_{j=1}^{Ni} g_{ij}Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij}W_j(t) - Q_{a,i}(t - \Delta t) \right], \quad (19)$$

$$q_{l,i}(t) = \frac{Q_{l,i}(t - \Delta t) + \sum_{j=1}^{Ni} g_{ij}Q_{l,j}(t) - \sum_{j=1}^{Mi} f_{ij}W_j(t) - Q_{a,i}(t - \Delta t)}{c_{t,i}V_{p,i} \left[\frac{1 - f_{w,i}(t - \Delta t)}{PI_{o,i}} + \frac{f_{w,i}(t - \Delta t)}{PI_{w,i}} + \frac{\Delta t}{c_{t,i}V_{p,i}} \right]}. \quad (20)$$

Теория Баклея–Левретта [14] позволяет связать объемную долю воды в потоке в вертикальном сечении пласта с водонасыщенностью в этом сечении (капиллярные и гравитационные силы при этом не учитываются). Тогда обводненность в пластовых условиях следующим образом зависит от водонасыщенности на стенке скважины:

$$wct_i(s_i) = \frac{q_{w,i}(t)}{q_{l,i}(t)} = \frac{q_{w,i}(t)}{q_{w,i}(t) + q_{o,i}(t)} = \frac{k_{rw,i}(s_i)}{k_{rw,i}(s_i) + k_{ro,i}(s_i) \frac{\mu_w}{\mu_o}}, \quad (21)$$

где $s_i = s_i(t)$ — водонасыщенность на стенке скважины, $k_{ro,i}(s_i)$ и $k_{rw,i}(s_i)$ — относительные фазовые проницаемости (ОФП) для нефти и воды, μ_o и μ_w — динамические вязкости для нефти и воды. Поскольку ранее дебиты были записаны в пластовых условиях, обводненность было решено записать в тех же условиях.

Для моделирования водонасыщенности могут быть использованы эмпирические зависимости водонасыщенности от накопленной добычи (обычно при 100 % компенсации отбора используется линейная зависимость от накопленной добычи нефти). В этой работе рассматривается эмпирическая логарифмическая зависимость водонасыщенности от накопленной добычи жидкости (интерференция скважин не учитывается):

$$s_i(t) = a_i + b_i \cdot \ln(Q_{l,i}(t)), \quad a_i < 0, \quad b_i > 0, \quad (22)$$

где a_i и b_i — константы, определяемые путем линейной регрессии.

Относительные фазовые проницаемости по нефти и воде описываются с помощью аналитической модели:

$$k_{rw}(s) = A \left(\frac{s(t) - s_{wcr}}{1 - s_{wcr}} \right)^\alpha, \quad k_{ro}(s) = B \left(\frac{1 - s_{owcr} - s(t)}{1 - s_{owcr} - s_{wcr}} \right)^\beta, \quad (23)$$

где s_{wcr} — насыщенность пласта связанной водой; s_{owcr} — остаточная нефтенасыщенность при вытеснении нефти водой; A , B , α , β — коэффициенты аппроксимации.

Практическое применение предложенной модели

Рассмотрим пример использования предложенной CRM-модели для анализа работы группы скважин на реальном месторождении. В течение анализируемого времени месторождение эксплуатировалось при упруговодонапорном режиме разработки, закачка воды отсутствовала.

Основные характеристики залежи нефти:

1. Глубина водонефтяного контакта (ВНК) — 2904 м.

2. Начальное пластовое давление — 311 бар.
3. Пористость — 0,12 д.ед.
4. Эффективная нефтенасыщенная толщина — 21 м.
5. Проницаемость — 350 мД.
6. Свойства нефти: вязкость — 5,88 мПа·с, объемный коэффициент — $1,08 \text{ м}^3/\text{м}^3$.
7. Свойства воды: вязкость — 1,1 мПа·с, объемный коэффициент — $1,01 \text{ м}^3/\text{м}^3$.
8. Суммарная сжимаемость системы пласт-флюид — $3,4 \cdot 10^{-5} \text{ 1/бар}$.

На рис. 1 приведена схема расположения скважин и границ. Анализируется работа скважин №№ 2, 10, 18, 24. Оценивается влияние на их работу скважин №№ 1302 и 1303, которые были запущены позже. Длительность анализируемого периода — 10,5 месяцев.

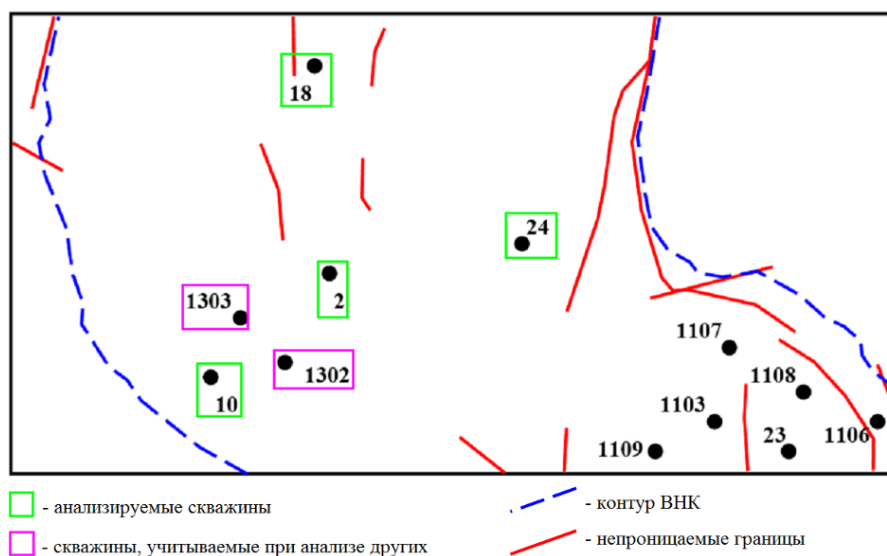


Рис. 1. Схема расположения скважин и границ

На рис. 2 приведены ОФП по нефти и воде и их аппроксимация с помощью формул (23).

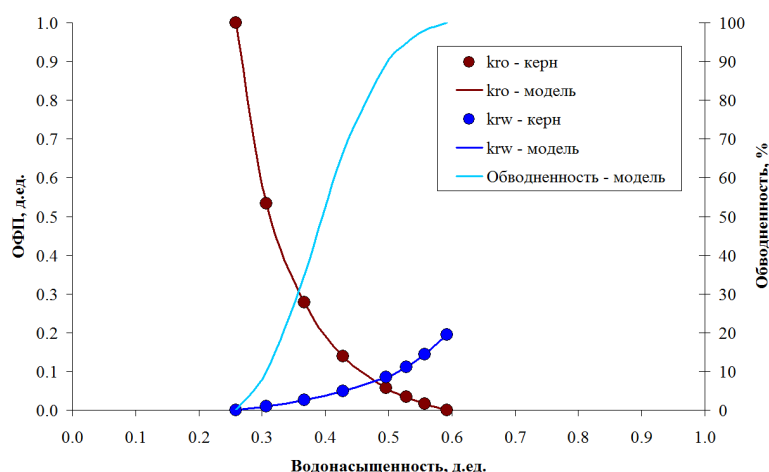


Рис. 2. ОФП по нефти и воде и их аппроксимация

На рис. 3, 4 и 5 представлены соответственно история забойного давления, дебита жидкости и обводненности по скважинам. Продолжительная история обводнения, которую можно анализировать, имеется только по скважинам №№ 2 и 24.

Далее на рисунках приведено сравнение модельных и фактических данных. Для оценки качества моделей и сравнения их между собой использован т.н. «скорректированный коэффициент детерминации» R_a^2 , который учитывает разное количество параметров моделей: $R_a^2 = 1 - (1 - R^2)(n - 1) / (n - k)$, где n — количество точек данных, k — количество параметров модели. Для

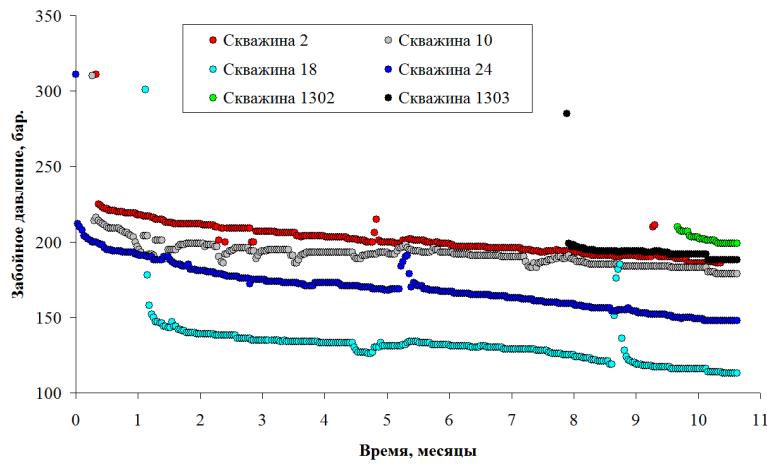


Рис. 3. История забойного давления по скважинам

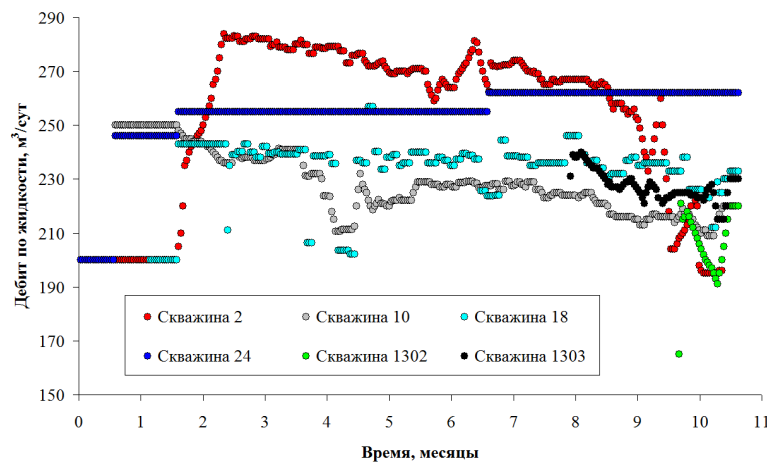


Рис. 4. История дебита жидкости по скважинам

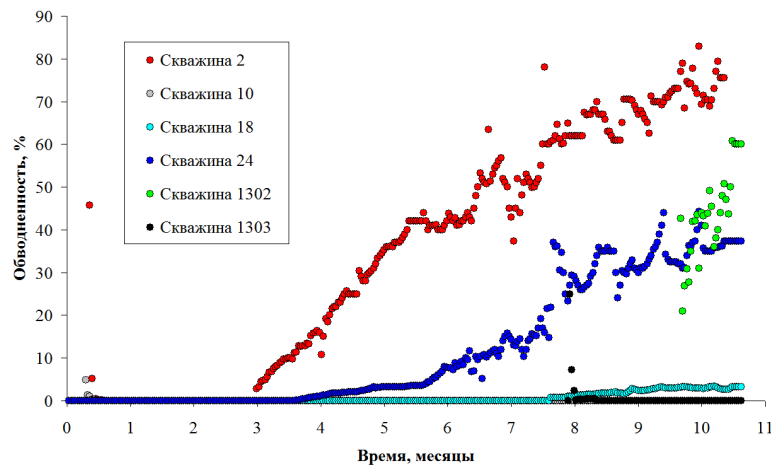


Рис. 5. История обводненности по скважинам

модели с одной независимой переменной коэффициент детерминации R^2 описывает долю дисперсии зависимой переменной, объясняемую рассматриваемой моделью.

На рис. 6–9 для скв. №№ 2, 10, 18 и 24 приведено сравнение стандартной CRM-модели для расчета дебита жидкости с учетом забойного давления и предлагаемой модифицированной CRM-модели без использования забойного давления. Для скв. № 2 получены следующие результаты: стандартная CRM-модель $R_a^2 = 0,9461$; модифицированная CRM-модель $R_a^2 = 0,7391$. Для скв. № 10 получены следу-

ющие результаты: стандартная CRM-модель $R_a^2 = 0,8454$; модифицированная CRM-модель $R_a^2 = 0,7924$. Для скв. № 18 получены следующие результаты: стандартная CRM-модель $R_a^2 = 0,9514$; модифицированная CRM-модель $R_a^2 = 0,9273$. Для скв. № 24 получены следующие результаты: стандартная CRM-модель $R_a^2 = 0,9505$; модифицированная CRM-модель $R_a^2 = 0,7493$. Итак, получен закономерный и логичный результат: пренебрежение забойным давлением при расчете дебита жидкости при помощи CRM-модели ухудшает ее качество. Однако ни в одном из рассмотренных случаев скорректированный коэффициент детерминации не упал ниже 0,7. Таким образом, при отсутствии данных о забойном давлении скважины можно пользоваться предлагаемой модифицированной CRM-моделью. Обе модели для дебита жидкости учитывают взаимодействие добывающих скважин, что косвенно отражает изменение объемов дренирования скважин за счет интерференции.

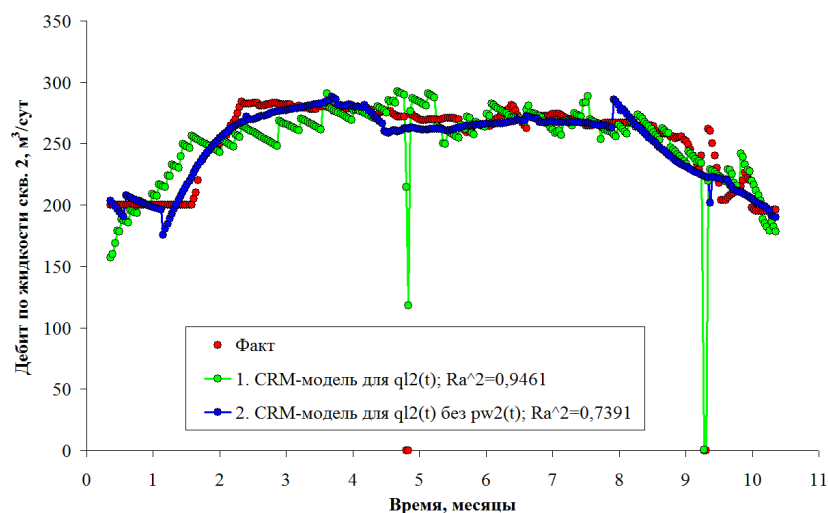


Рис. 6. Моделирование дебита жидкости скв. № 2. $ql2(t)$ и $pw2(t)$ — дебит жидкости и забойное давление скв. № 2, Ra^2 — скорректированный коэффициент детерминации

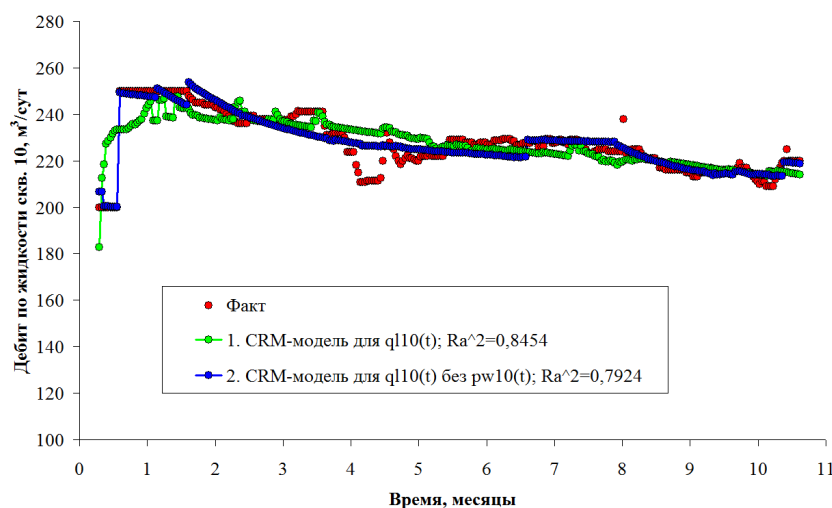


Рис. 7. Моделирование дебита жидкости скв. № 10. $ql10(t)$ и $pw10(t)$ — дебит жидкости и забойное давление скв. № 10, Ra^2 — скорректированный коэффициент детерминации

На рис. 10–13 приведены модели для водонасыщенности и обводненности по скважинам №№ 2 и 24. Несмотря на отсутствие в модели обводнения учета интерференции добывающих скважин, получено хорошее качество совмещения. Вывод: учет интерференции добывающих скважин для модели дебита жидкости важнее, чем для модели обводнения. Нужно отметить, что в модели обводнения интерференция скважин все-таки косвенно присутствует, так как при расчетах водонасыщенности используется накопленная добыча жидкости, которая рассчитывается из дебита жидкости, учитывающего интерференцию добывающих скважин.

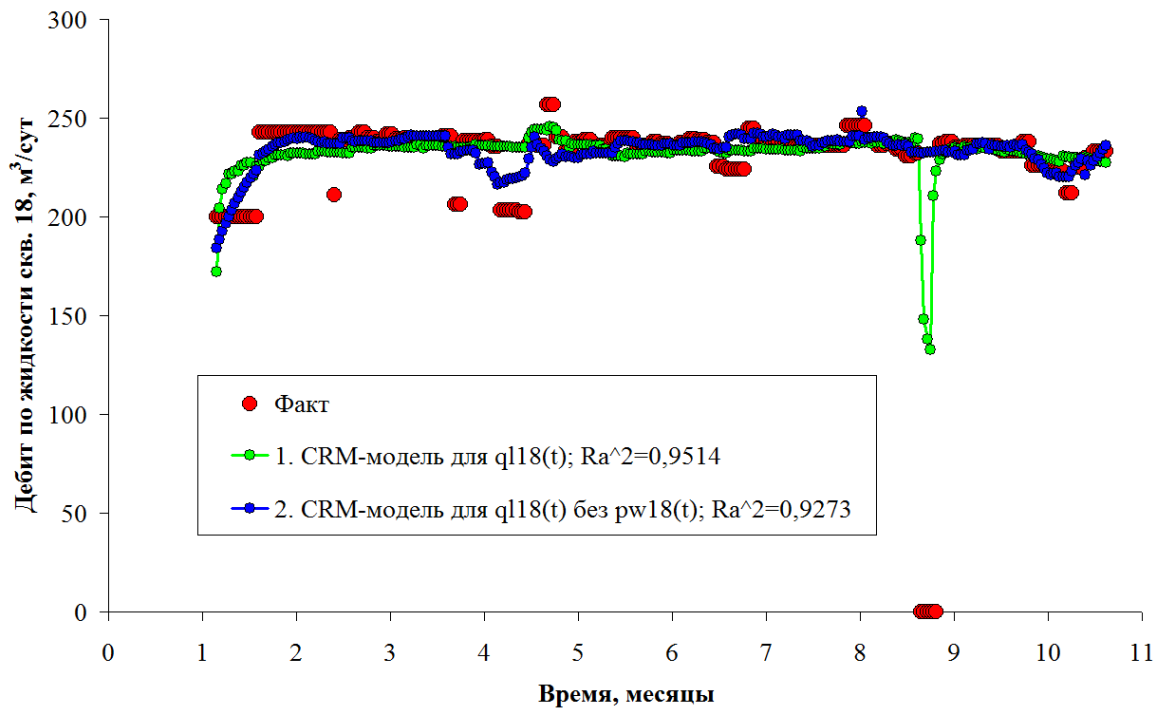


Рис. 8. Моделирование дебита жидкости скв. № 18. $q_{l18}(t)$ и $p_{w18}(t)$ — дебит жидкости и забойное давление скв. № 18, Ra^2 — скорректированный коэффициент детерминации

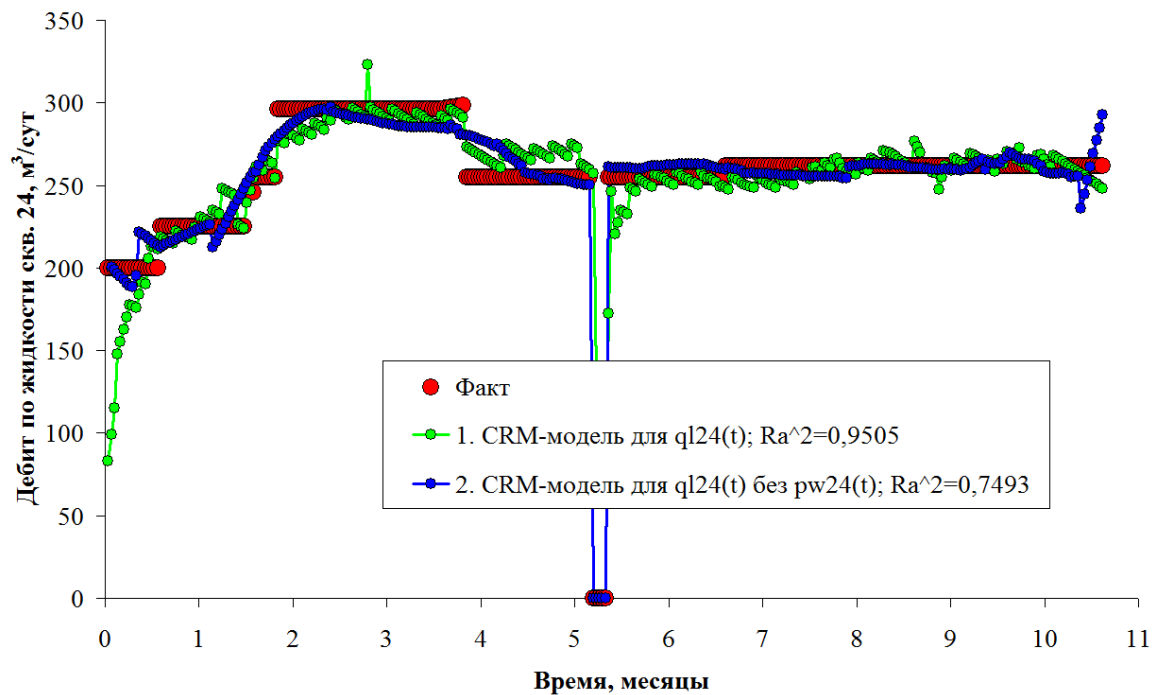


Рис. 9. Моделирование дебита жидкости скв. № 24. $q_{l24}(t)$ и $p_{w24}(t)$ — дебит жидкости и забойное давление скв. № 24, Ra^2 — скорректированный коэффициент детерминации

Обсуждение результатов

Результаты моделирования, представленные в работе, позволяют сделать целый ряд важных выводов о применимости модифицированной CRM-модели без использования забойного давления и предложенной логарифмической модели обводнения. Прежде всего, следует отметить, что адаптация емкостно-резистивной модели под условия ограниченного промыслового мониторинга, в частности при отсутствии данных о забойном давлении, является актуальной задачей, особенно в условиях старе-

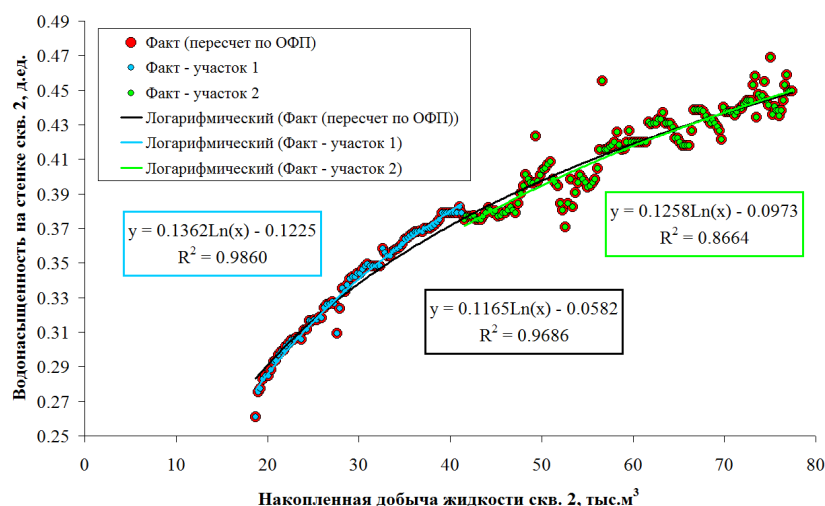


Рис. 10. Зависимость водонасыщенности на стенке скв. № 2 от накопленной добычи жидкости этой скважины

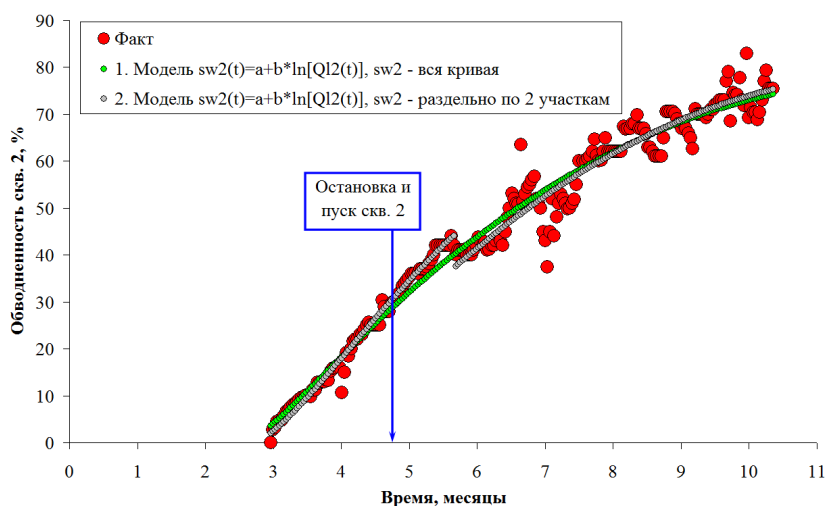


Рис. 11. Моделирование обводненности скв. № 2. $sw2(t)$ и $Ql2(t)$ — водонасыщенность на стенке скважины и накопленная добыча жидкости скв. № 2 соответственно

ющих месторождений, удаленных участков или малых предприятий, где регулярные замеры давления невозможны.

Анализ результатов по четырем скважинам показал, что предлагаемая модифицированная CRM-модель сохраняет приемлемую точность прогнозирования дебита жидкости, несмотря на отсутствие информации о забойном давлении. Наиболее высокий скорректированный коэффициент детерминации при использовании модели без забойного давления был получен для скважины № 18 — $R_a^2 = 0,9273$, что свидетельствует о высокой достоверности расчетов даже в упрощенной постановке. Минимальный скорректированный коэффициент детерминации — $R_a^2 = 0,7391$ (по скважине № 2), что также можно считать удовлетворительным результатом с точки зрения инженерного применения, особенно если учитывать, что модель не использует ни одного значения измерения давления.

Важно отметить, что в модели сохранен ключевой элемент классических CRM — межскважинное взаимодействие, что позволяет учесть влияние скважин-соседей, изменений в дренируемых объемах. Это отличает предложенный подход от более простых эмпирических методов, не способных отразить такую динамику. При этом модель избегает необходимости калибровки по давлению, что существенно снижает требования к качеству исходной информации и трудозатраты на ее сбор.

Сравнение результатов стандартной и модифицированной моделей продемонстрировало закономерное снижение точности при исключении давления, однако это снижение оказалось весьма уме-

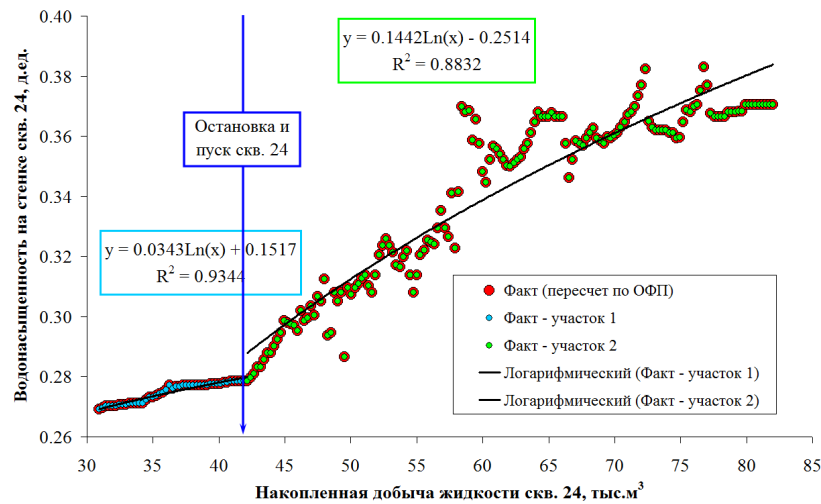


Рис. 12. Зависимость водонасыщенности на стенке скв. № 24 от накопленной добычи жидкости этой скважины

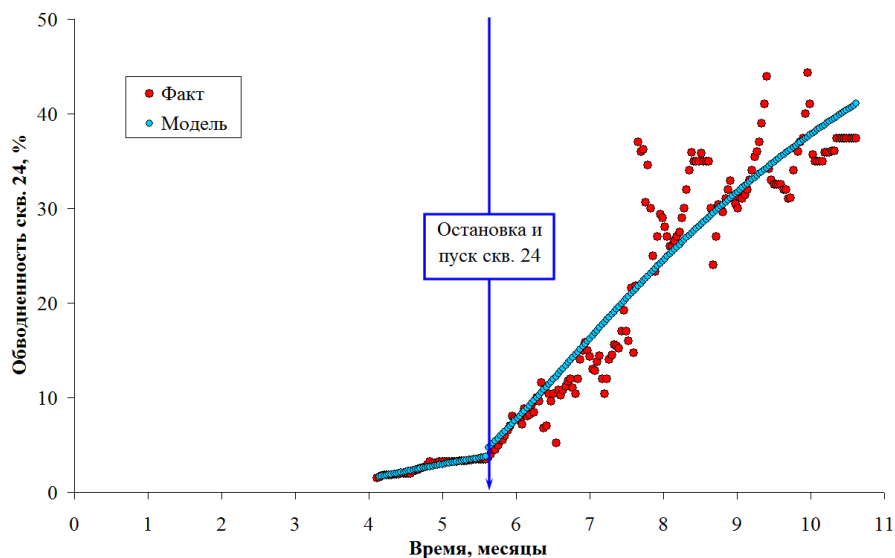


Рис. 13. Моделирование обводненности скв. № 24

ренным. Таким образом, можно утверждать, что модифицированная модель работает надежно и воспроизводит основные тренды добычи. Это делает ее особенно полезной в условиях, когда забойное давление отсутствует и необходимо получить быстрый приблизительный прогноз.

Не менее интересными являются результаты по модели обводнения. Использование логарифмической зависимости между водонасыщенностью и накопленной добычей жидкости позволило успешно аппроксимировать данные по скважинам № 2 и № 24. Несмотря на то, что модель обводнения не включает прямой учет интерференции, ее эффективность обусловлена тем, что накопленная добыча жидкости в ней выступает в роли «свертки» множества факторов, включая влияние соседних скважин, изменение объема дренирования, эволюцию фазовой проницаемости и др. Таким образом, модель водонасыщенности оказывается устойчивой и воспроизводит наблюдаемые тренды обводнения, при этом ее реализация остается простой и не требует большого количества параметров.

Следует подчеркнуть, что предложенная логарифмическая модель отличается от традиционных линейных зависимостей, применяемых, как правило, к накопленной добыче нефти. Такой переход к логарифму от общего объема добычи жидкости позволяет, по-видимому, отразить не только постепенное насыщение порового пространства водой, но и описать неоднородность пласта, что делает модель более реалистичной и подходящей для условий с закачкой воды и выраженным упруговодонапорным режимом.

Также заслуживает внимания тот факт, что во всех случаях расчетов использовались ОФП, полученные из керновых данных. Это позволило избежать необходимости в подборе эмпирических коэффициентов функций ОФП и повысило обоснованность расчетов. Таким образом, предложенные модели сочетают теоретическую строгость, инженерную реализуемость и устойчивость к неполноте данных, что делает их пригодными для широкого спектра практических задач.

Заключение

В представленной работе подготовлена и апробирована модифицированная версия CRM-модели, не использующая данные о забойном давлении и применимая для условий заводнения и упруговодонапорного режима разработки. Такая модель расширяет применимость ёмкостно-резистивных подходов, позволяя проводить анализ межскважинного взаимодействия, оценку откликов на закачку и построение прогноза добычи даже в условиях ограниченного промыслового мониторинга.

Дополнительным вкладом в развитие моделей такого класса является введение логарифмической зависимости между водонасыщенностью и накопленной добычей жидкости, позволяющей моделировать процессы обводнения скважин без использования полной информации о фазовом составе. Эта модель, несмотря на свою простоту, показала высокую точность аппроксимации и устойчивость к изменениям параметров.

Проведенные расчеты по четырем скважинам реального месторождения с карбонатным коллектором подтвердили, что модифицированная CRM-модель, несмотря на отсутствие давления, обеспечивает приемлемое качество прогноза дебита жидкости (скорректированный коэффициент детерминации $R_a^2 > 0,7$ во всех случаях), а логарифмическая модель водонасыщенности позволяет достоверно моделировать рост обводненности. Таким образом, полученные результаты подтверждают возможность использования представленных моделей в качестве инженерного инструмента для оперативного анализа и прогноза разработки.

Практическое значение данного подхода особенно велико в случаях дефицита данных, необходимости быстрого принятия решений и оценки сценариев разработки. Предложенные модели могут быть интегрированы в программные комплексы оперативного мониторинга, использованы в рамках упрощенного анализа истории разработки при отсутствии цифровой геолого-гидродинамической модели, а также применены на стадиях планирования добычи в малых нефтяных компаниях и на зрелых активах.

Таким образом, модифицированная CRM-модель и логарифмическая модель обводнения представляют собой перспективное направление развития прикладного моделирования разработки нефтяных месторождений, сочетая математическую строгость, адаптивность к данным и практическую применимость в современных условиях нефтедобычи.

ЛИТЕРАТУРА

1. Данько М. Ю., Бриллиант Л. С., Завьялов А. С. Применение метода динамического материального баланса и CRM-метода (capacitance-resistive models) к подсчету запасов ачимовских и баженовских коллекторов. *Недропользование XXI век*. 2019;4:76–85.
2. Ручкин А. А., Степанов С. В., Князев А. В. и др. Исследование особенностей оценки взаимовлияния скважин на примере модели CRM. *Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика*. 2018;4:148–168. DOI: 10.21684/2411-7978-2018-4-4-148-168.
3. Хатмуллин И. Ф., Андрианова А. М., Маргарит А. С. Полуаналитические модели расчета интерференции скважин на базе класса моделей CRM. *Нефтяное хозяйство*. 2018;12:38–41. DOI: 10.24887/0028-2448-2018-12-38-41.
4. Gladkov A., Sakhibgareev R., Salimov D. et al. Application of CRM for Production and Remaining Oil Reserves Reservoir Allocation in Mature West Siberian Waterflood Field. *The SPE Russian Petroleum Technology Conference*. 2017. DOI: 10.2118/187841-MS.
5. Yousefi S. H., Rashidi F., Sharifi M. et al. Prediction of Immiscible Gas Flooding Performance: a Modified Capacitance-Resistance Model and Sensitivity Analysis. *Petroleum Science*. 2019;16:1086–1104. DOI: 10.1007/s12182-019-0342-6.
6. Lasdon L., Shirzadi S., Ziegel E. Implementing CRM Models for Improved Oil Recovery in Large Oil Fields. *Optimization and Engineering*. 2017;18:87–103. DOI: 10.1007/s11081-017-9351-8.

7. Holanda R. W., Gildin E., Jensen J. L. et al. A State-of-the-Art Literature Review on Capacitance Resistance Models for Reservoir Characterization and Performance Forecasting. *Energies*. 2018;11(12):33–68. DOI: 10.3390/en11123368.
8. Sayarpour M. *Development and Application of Capacitance-Resistive Models to Water/CO₂ Floods*. Dissertation by Ph.D. USA: The Faculty of the Graduate School of the University of Texas at Austin; 2008. 236 p. DOI: 10.13140/RG.2.1.1798.3847.
9. Степанов С. В., Поспелова Т. А. Новая концепция математического моделирования для принятия решений по разработке месторождений. *Нефтяное хозяйство*. 2019;4:50–53. DOI: 10.24887/0028-2448-2019-4-50-53.
10. Olenchikov D., Posvyanskii D. Application of CRM-Like Models for Express Forecasting and Optimizing Field Development. *The SPE Russian Petroleum Technology Conference*. 2019. SPE-196893-MS. DOI: 10.2118/196893-MS.
11. Sayarpour M., Kabir C. S., Lake L. W. Field Applications of Capacitance-Resistive Models in Waterfloods. *SPE Annual Technical Conference and Exhibition*. 2008. DOI: 10.2118/114983-MS.
12. Sayarpour M., Zuluaga E., Kabir C. S. et al. The Use of Capacitance-Resistance Models for Rapid Estimation of Waterflood Performance and Optimization. *Journal of Petroleum Science and Engineering*. 2009;69:227–238. DOI: 10.1016/j.petrol.2009.09.006.
13. Sayarpour M., Kabir C. S., Sepehrnoori K. et al. Probabilistic History Matching with the Capacitance-Resistance Model in Waterfloods: A Precursor to Numerical Modeling. *The SPE Improved Oil Recovery Symposium*. 2010. DOI: 10.2118/129604-MS.
14. Басниев К. С., Кочина И. Н., Максимов В. М. *Подземная гидродинамика*. М.: Недра; 1993. 416 с.
15. Bolshakov M., Deeva T., Pustovskikh A. Fluid Flow Prediction with Development System Interwell Connectivity Influence. *IOP Conf. Series: Earth and Environmental Science*. 2016;33:012036. DOI: 10.1088/1755-1315/33/1/012036.
16. Soroush M., Kaviani D., Jensen J. L. Interwell Connectivity Evaluation in Cases of Changing Skin and Frequent Production Interruptions. *Journal of Petroleum Science and Engineering*. 2014;122:616–630. DOI: 10.1016/j.petrol.2014.09.001.
17. Наугольников М., Тепляков Н., Большаков М. С. Cost-БФА: инструмент стоимостной оптимизации заводнения. *Российская нефтегазовая техническая конференция SPE*. 2018. DOI: 10.2118/191580-18RPTC-RU. Режим доступа: https://www.researchgate.net/publication/328221524_Cost-Engineering_Waterflooding_Management_Methods_Russian.

К ЗАДАЧЕ ПОСТРОЕНИЯ УРАВНЕНИЯ ЭЙЛЕРА–ЛАГРАНЖА С ПОМОЩЬЮ ЛЕММЫ ПУАНКАРЕ

В. П. Кошечев

Московский авиационный институт (национальный исследовательский университет), филиал
«Стрела», г. Жуковский, Московская область, Российская Федерация
ORCID: <https://orcid.org/0000-0002-0724-9760>, koshcheev1@yandex.ru

Аннотация: показано, что уравнение Эйлера–Лагранжа для динамической системы с одной степенью свободы может быть построено с помощью леммы Пуанкаре. В приложении к статье доказана теорема существования и единственности свойства кососимметричности простейших (базовых) дифференциальных форм.

Ключевые слова: уравнение Эйлера–Лагранжа, динамическая система с одной степенью свободы, лемма Пуанкаре, простейшие (базовые) дифференциальные формы.

Для цитирования: Кошечев В. П. К задаче построения уравнения Эйлера–Лагранжа с помощью леммы Пуанкаре. *Успехи кибернетики*. 2025;6(2):40–43.

Поступила в редакцию: 06.12.2024.

В окончательном варианте: 14.01.2025.

CONSTRUCTING THE EULER–LAGRANGE EQUATION USING THE POINCARÉ LEMMA

V. P. Koshcheev

Moscow Aviation Institute (National Research University), Strela branch, Zhukovsky, Moscow Region,
Russian Federation
ORCID: <https://orcid.org/0000-0002-0724-9760>, koshcheev1@yandex.ru

Abstract: we demonstrate that the Euler–Lagrange equation for a dynamical system with one degree of freedom can be derived using the Poincaré lemma. In the appendix, we prove the theorem on the existence and uniqueness of the skew-symmetry property of the simplest (basic) differential forms.

Keywords: Euler–Lagrange equation, 1-DOF dynamic system Poincaré lemma, simplest (basic) differential forms.

Cite this article: Koshcheev V. P. Constructing the Euler–Lagrange Equation Using the Poincaré Lemma. *Russian Journal of Cybernetics*. 2025;6(2):40–43.

Original article submitted: 06.12.2024.

Revision submitted: 14.01.2025.

В [1] с помощью леммы Пуанкаре построена первая пара уравнений Максвелла, которая не содержит плотности заряда и тока. Построение уравнения Эйлера–Лагранжа для динамической системы с одной степенью свободы с помощью леммы Пуанкаре составляет цель настоящей работы.

Классическое действие для динамической системы с одной степенью свободы запишем в виде определенного интеграла с переменным верхним пределом

$$S(t) = \int_{t_0}^t L\left(t', q(t'), \frac{dq(t')}{dt'}\right) dt'. \quad (1)$$

Тогда

$$dS = L\left(t', q(t'), \frac{dq(t')}{dt'}\right) dt' \Big|_{t'=t} = L(t, q, \dot{q}) dt, \quad (2)$$

где $\dot{q} = \frac{dq}{dt}$.

Построим дифференциал второго порядка

$$ddS = \left(\frac{\partial L}{\partial t} dt + \frac{\partial L}{\partial q} dq + \frac{\partial L}{\partial \dot{q}} d\dot{q} \right) \wedge dt. \quad (3)$$

Так как (см. приложение) $dt \wedge dt = 0$ и $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} dq \right) = \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) dq + \frac{\partial L}{\partial \dot{q}} d\dot{q}$, то

$$d \left(dS - \frac{\partial L}{\partial \dot{q}} dq \right) = \left(-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial q} \right) dq \wedge dt. \quad (4)$$

Покажем, что

$$dS - \frac{\partial L}{\partial \dot{q}} dq = -H dt \quad (5)$$

является 1 — формой, так как выполняется лемма Пуанкаре

$$d \left(dS - \frac{\partial L}{\partial \dot{q}} dq \right) = d(-H dt) = -\frac{dH}{dt} dt \wedge dt = 0. \quad (6)$$

Тогда из уравнения (4) получим уравнение Эйлера–Лагранжа

$$0 = -\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial q}, \quad (7)$$

так как $dq \wedge dt \neq 0$.

Приложение

Как физики [1], так и математики [2] вынуждены постулировать свойство кососимметричности простейших дифференциальных форм, чтобы не излагать достаточно громоздкую стандартную процедуру их введения [3]. Новый подход к задаче построения определителя Якоби был предложен в [4]. Было показано, что произведение дифференциалов как в одной, так и в другой системах координат являются кососимметричными, если они связаны между собой с помощью определителя Якоби. В рамках нового подхода [4] докажем теорему существования и единственности свойства кососимметричности простейших (базовых) дифференциальных форм.

Известно, что инвариантность формы первого дифференциала функции многих (нескольких) переменных в разных системах координат достигается с помощью дополнительного условия, согласно которому дифференциалы независимых переменных преобразуются как компоненты контравариантного вектора, а частные производные преобразуются как компоненты ковариантного вектора. Сперва рассмотрим функцию двух независимых переменных в старой системе координат

$$f = f(u, v), \quad (\text{П1})$$

которые зависят от независимых переменных в новой системе координат

$$\begin{cases} u = u(x, y) \\ v = v(x, y). \end{cases}$$

Тогда

$$\begin{cases} du = u_x dx + u_y dy, \\ dv = v_x dx + v_y dy, \end{cases} \quad (\text{П2})$$

где

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y}, \quad v_x = \frac{\partial v}{\partial x}, \quad v_y = \frac{\partial v}{\partial y}.$$

Запишем систему уравнений (П2) в виде

$$A \begin{pmatrix} du \\ dv \end{pmatrix} = A J A^{-1} A \begin{pmatrix} dx \\ dy \end{pmatrix}, \quad (\text{П3})$$

где

$$J = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \quad \text{— матрица Якоби,}$$

$A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$ — диагональная матрица; $A^{-1} = \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \beta^{-1} \end{pmatrix}$ — обратная матрица; $AA^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; α и β — произвольные числа.

Видно, что

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} = \begin{pmatrix} du & 0 \\ 0 & dv \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \quad (\text{П4})$$

В результате получим

$$\left[\begin{pmatrix} du & 0 \\ 0 & dv \end{pmatrix} - AJA^{-1} \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix} \right] \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0. \quad (\text{П5})$$

Известно, что тривиальное решение системы однородных линейных алгебраических уравнений существует, и оно единственно. Рассмотрим обратную задачу (П5), когда элементы матрицы столбца не равны нулю, то есть $\alpha \neq 0$ и $\beta \neq 0$, а элементы квадратной матрицы являются неизвестными величинами. Очевидно, что тривиальное решение уравнения (П5) существует, и оно единственное

$$\begin{pmatrix} du & 0 \\ 0 & dv \end{pmatrix} = AJA^{-1} \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix}. \quad (\text{П6})$$

Так как определитель произведения матриц одинаковых порядков в (П6) равен произведению определителей этих матриц, то

$$\det \begin{pmatrix} du & 0 \\ 0 & dv \end{pmatrix} = \det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \cdot \det \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix}, \quad (\text{П7})$$

где $\det A \cdot \det A^{-1} = 1$.

В результате получим

$$du \wedge dv = \det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} dx \wedge dy, \quad (\text{П8})$$

где $\det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$ — определитель Якоби.

Произведение дифференциалов независимых переменных будем обозначать символом $\wedge$, который в алгебре определяет операцию внешнего умножения кососимметричных форм [3]. Покажем, что простейшие дифференциальные формы являются кососимметричными, то есть $du \wedge du = dv \wedge dv = 0$ и $du \wedge dv = -dv \wedge du$. Так как

$$\begin{cases} dv = v_x dx + v_y dy, \\ du = u_x dx + u_y dy, \end{cases}$$

то

$$\det \begin{pmatrix} dv & 0 \\ 0 & du \end{pmatrix} = \det \begin{pmatrix} v_x & v_y \\ u_x & u_y \end{pmatrix} \cdot \det \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix} = -\det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \cdot \det \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix}.$$

Тогда с помощью (П7) и (П8) получим

$$dv \wedge du = \det \begin{pmatrix} v_x & v_y \\ u_x & u_y \end{pmatrix} dx \wedge dy = -\det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} dx \wedge dy = -du \wedge dv.$$

Если

$$\begin{cases} du = u_x dx + u_y dy, \\ du = u_x dx + u_y dy, \end{cases}$$

то

$$\det \begin{pmatrix} du & 0 \\ 0 & du \end{pmatrix} = \det \begin{pmatrix} u_x & u_y \\ u_x & u_y \end{pmatrix} \cdot \det \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix} = 0 \cdot \det \begin{pmatrix} dx & 0 \\ 0 & dy \end{pmatrix} = 0.$$

Тогда

$$du \wedge du = \det \begin{pmatrix} u_x & u_y \\ u_x & u_y \end{pmatrix} dx \wedge dy = 0 \cdot dx \wedge dy = 0.$$

Если определитель Якоби не равен нулю, то легко получить

$$dx \wedge dx = dy \wedge dy = 0 \quad \text{и} \quad dx \wedge dy = -dy \wedge dx.$$

Таким образом, произведение дифференциалов как в одной, так и в другой системах координат являются кососимметричными, если они связаны между собой с помощью определителя Якоби.

ЛИТЕРАТУРА

1. Райдер Л. *Квантовая теория поля*. ПЛАТОН; 1998.
2. Яковлев Г. Н. *Лекции по математическому анализу*. Ч. 2. М.: Физматлит; 2004.
3. Зорич В. А. *Математический анализ: в 2-х частях. Часть 2*. М.: Наука; 1984.
4. Кошечев В. П. К задаче построения определителя Якоби. *Успехи кибернетики*. 2021;2(4):87–89. DOI: 10.51790/2712-9942-2021-2-4-10.

КЛАССИЧЕСКИЕ И КВАНТОМЕХАНИЧЕСКИЕ ПОПРАВКИ К РЕШЕНИЮ УРАВНЕНИЯ ЭЙЛЕРА–ЛАГРАНЖА

В. П. Кошечев

*Московский авиационный институт (национальный исследовательский университет), филиал
«Стрела», г. Жуковский, Московская область, Российская Федерация
ORCID: <https://orcid.org/0000-0002-0724-9760>, koshcheev1@yandex.ru*

Аннотация: с помощью уравнения Якоби построена цепочка замкнутых систем обыкновенных дифференциальных уравнений первого порядка, описывающих эволюцию моментов, начальные условия к которым могут содержать постоянную Планка. Показано, что эволюция моментов разных порядков происходит независимо друг от друга. Показано, что источником динамического хаоса являются, в частности, ненулевые начальные условия к системе уравнений для вторых моментов, так как средние квадраты флуктуаций координаты и импульса, которые являются мерой «разбегания» траекторий, возрастают экспоненциально быстро в области отрицательной гауссовой кривизны потенциала. Так как решения уравнения Якоби должны являться малыми поправками к решению уравнения Эйлера–Лагранжа, то обратное влияние этих поправок на решение уравнения Эйлера–Лагранжа может быть учтено с помощью розыгрыша (метод Монте-Карло) динамических переменных, что, в свою очередь, приведет к появлению ненулевых начальных условий для моментов высших порядков.

Ключевые слова: уравнение Эйлера–Лагранжа, уравнение Якоби, цепочка замкнутых систем обыкновенных дифференциальных уравнений первого порядка, средних значений корреляторов оператора координаты и импульса, динамический хаос.

Для цитирования: Кошечев В. П. Классические и квантовомеханические поправки к решению уравнения Эйлера–Лагранжа. *Успехи кибернетики*. 2025;6(2):44–46.

Поступила в редакцию: 27.12.2024.

В окончательном варианте: 01.02.2025.

CLASSICAL AND QUANTUM-MECHANICAL CORRECTIONS TO THE SOLUTION OF THE EULER–LAGRANGE EQUATION

V. P. Koshcheev

*Moscow Aviation Institute (National Research University), Strela branch, Zhukovsky, Moscow Region,
Russian Federation
ORCID: <https://orcid.org/0000-0002-0724-9760>, koshcheev1@yandex.ru*

Abstract: using the Jacobi equation, we constructed a chain of closed systems of first-order ordinary differential equations that describe the evolution of moments, where the initial conditions can include Planck's constant. We showed that the evolution of moments of different orders occurs independently of each other. We demonstrated that the source of dynamic chaos is, in particular, nonzero initial conditions for the system of equations governing second moments. This is because the mean square fluctuations of position and momentum, which measure the “divergence” of trajectories, grow exponentially fast in the region of negative Gaussian curvature of the potential. Since the solutions of the Jacobi equation must be small corrections to the solution of the Euler–Lagrange equation, the reverse influence of these corrections on the Euler–Lagrange equation can be accounted for using the Monte Carlo method for dynamic variables. This, in turn, leads to the emergence of nonzero initial conditions for higher-order moments.

Keywords: Euler–Lagrange equation, Jacobi equation, chain of closed systems of first-order ordinary differential equations, average values of position and momentum operator correlators, dynamic chaos.

Cite this article: Koshcheev V. P. Classical and Quantum-Mechanical Corrections to the Solution of the Euler–Lagrange Equation. *Russian Journal of Cybernetics*. 2025;6(2):44–46.

Original article submitted: 27.12.2024.

Revision submitted: 01.02.2025.

В рамках нового подхода к построению квазиклассического приближения в квантовой механике в [1] была построена замкнутая система обыкновенных дифференциальных уравнений первого порядка для средних значений корреляторов оператора координаты и импульса, которая не содержит явно постоянной Планка, а начальные условия к этой системе уравнений удовлетворяют системе обобщенных соотношений неопределенностей Гейзенберга. В [1] была поставлена задача построения точной

механической аналогии этой системы уравнений. В настоящей работе представлена попытка решения этой задачи.

Пусть нам дано уравнение Эйлера–Лагранжа

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = 0, \quad (1)$$

где $L = L(t, x, \dot{x})$ — функция Лагранжа; $\dot{x} = \frac{dx}{dt}$.

Пусть $x = \bar{x} + \delta x$ и $\dot{x} = \dot{\bar{x}} + \delta \dot{x}$. Тогда

$$L(t, \bar{x} + \delta x, \dot{\bar{x}} + \delta \dot{x}) = L(t, \bar{x}, \dot{\bar{x}}) + \delta L + \dots, \quad (2)$$

где $\delta L = \frac{\partial L(t, \bar{x}, \dot{\bar{x}})}{\partial \bar{x}} \delta x + \frac{\partial L(t, \bar{x}, \dot{\bar{x}})}{\partial \dot{\bar{x}}} \delta \dot{x}$.

Подставим (1) в (2) и получим уравнение Якоби

$$\frac{\partial \delta L}{\partial x} - \frac{d}{dt} \frac{\partial \delta L}{\partial \dot{x}} = 0,$$

которое записывают в виде [2]

$$\left(L_{xx} - \frac{dL_{x\dot{x}}}{dt} \right) \delta x - \frac{d}{dt} (L_{\dot{x}\dot{x}} \delta \dot{x}) = 0, \quad (3)$$

где обозначим $x \equiv \bar{x}$ для удобства.

Если $L = \frac{m\dot{x}^2}{2} - U(x, t)$, как в [1], то уравнения Эйлера–Лагранжа и Якоби имеют вид, соответственно:

$$U_x + m\ddot{x} = 0; \quad (4)$$

$$U_{xx}\delta x + m\delta\ddot{x} = 0. \quad (5)$$

Видно, что с помощью уравнения Якоби (5) могут быть построены замкнутые системы дифференциальных уравнений первого порядка для трех новых динамических переменных:

$$\begin{cases} \frac{d(\delta x^2)}{dt} = 2(\delta x \delta \dot{x}) \\ \frac{d(\delta x \delta \dot{x})}{dt} = (\delta \dot{x}^2) - \frac{U_{xx}}{m}(\delta x^2) \\ \frac{d(\delta \dot{x}^2)}{dt} = -\frac{2U_{xx}}{m}(\delta x \delta \dot{x}), \end{cases} \quad (6)$$

четырёх новых динамических переменных:

$$\begin{cases} \frac{d(\delta x^3)}{dt} = 3(\delta x^2 \delta \dot{x}) \\ \frac{d(\delta x^2 \delta \dot{x})}{dt} = 2(\delta x \delta \dot{x}^2) - \frac{U_{xx}}{m}(\delta x^3) \\ \frac{d(\delta x \delta \dot{x}^2)}{dt} = (\delta \dot{x}^3) - \frac{2U_{xx}}{m}(\delta x^2 \delta \dot{x}) \\ \frac{d(\delta \dot{x}^3)}{dt} = -\frac{3U_{xx}}{m}(\delta \dot{x}^2 \delta x) \end{cases} \quad (7)$$

и так далее.

Видно, что замкнутые системы уравнений (6), (7) зависят от решения уравнения (4) и эволюционируют независимо друг от друга. Так как решения уравнения Якоби должны являться малыми поправками к решению уравнения Эйлера–Лагранжа, то обратное влияние этих поправок на решение уравнения Эйлера–Лагранжа может быть учтено с помощью розыгрыша (метод Монте-Карло) динамических переменных, что, в свою очередь, приведет к появлению ненулевых начальных условий для

моментов высших порядков. Если $m\delta\dot{x} = \delta p$, то система уравнений (6) с точностью до обозначений совпадает с системой уравнений [1], которую запишем в виде [3]:

$$\begin{cases} \frac{d}{dt}\overline{\delta\hat{x}^2} = \frac{2}{m}\overline{\delta\hat{x}\delta\hat{p}_x} \\ \frac{d}{dt}\overline{\delta\hat{x}\delta\hat{p}_x} = \frac{1}{m}\overline{\delta\hat{p}_x^2} - U_{xx}(x)\overline{\delta\hat{x}^2} \\ \frac{d}{dt}\overline{\delta\hat{p}_x^2} = -2U_{xx}(x)\overline{\delta\hat{x}\delta\hat{p}_x}. \end{cases} \quad (8)$$

С помощью волновых функций, минимизирующих соотношение неопределенностей Гейзенберга, вычислены начальные условия к системе дифференциальных уравнений (8):

$$\begin{aligned} \overline{\delta\hat{x}\delta\hat{p}_x}\Big|_{t=0} &= 0, \\ \overline{\delta\hat{x}^2}\Big|_{t=0} \cdot \overline{\delta\hat{p}_x^2}\Big|_{t=0} &= \frac{\hbar^2}{4}. \end{aligned} \quad (9)$$

Исследуем на устойчивость решения системы дифференциальных уравнений (8) на небольших временных интервалах $x(t + \Delta t) \approx x(t)$. Решения системы уравнений (8) будем искать в виде:

$$\begin{cases} \overline{\delta\hat{x}^2} = Ae^{\lambda t} \\ \overline{\delta\hat{x}\delta\hat{p}_x} = Be^{\lambda t} \\ \overline{\delta\hat{p}_x^2} = Ce^{\lambda t}, \end{cases} \quad (10)$$

где A, B, C — постоянные интегрирования, а λ — показатель Ляпунова.

Подставим (10) в (8) и получим:

$$\begin{cases} \lambda A = \frac{2}{m}B \\ \lambda B = \frac{1}{m}C - U_{xx}(x)A \\ \lambda C = -2U_{xx}(x)B. \end{cases} \quad (11)$$

Найдем три корня характеристического уравнения:

$$\begin{cases} \lambda_1 = 0 \\ \lambda_2 = 2\sqrt{\frac{-U_{xx}(x)}{m}} \\ \lambda_3 = -2\sqrt{\frac{-U_{xx}(x)}{m}}. \end{cases} \quad (12)$$

Видно, что источником динамического хаоса являются, в частности, ненулевые начальные условия к системе уравнений (8), так как средние квадраты флуктуаций координаты и импульса (10), которые являются мерой «разбегания» траекторий, возрастают экспоненциально быстро $\lambda_2 > 0$ в области отрицательной гауссовой кривизны потенциала $U_{xx}(x) < 0$.

ЛИТЕРАТУРА

1. Bagrov V. G., Belov V. V., Kondrat'eva M. F. The Semiclassical Approximation in Quantum Mechanics. A New Approach. *Theoretical and Mathematical Physics*. 1994;98(1):34–38.
2. Гельфанд И. М., Фомин С. В. *Вариационное исчисление*. М.: Физматгиз; 1961. 228 с.
3. Koshcheev V. P., Morgun D. A., Panina T. A., Shtanov Y. N. Influence of Quantum Fluctuations on the Stochastic Dynamics of the Channeling Effect of Relativistic Electrons and Positrons. *Journal of Surface Investigation. X-ray, Synchrotron and Neutron Techniques*. 2012;6(1):168–171.

СОВРЕМЕННЫЕ ПРОБЛЕМЫ И РЕШЕНИЯ В ОБЛАСТИ МАШИННОГО ПЕРЕВОДА

Я. Д. Бакуменко^а, А. Н. Таджибова^б*Сургутский государственный университет, г. Сургут, Российская Федерация*^а ORCID: <http://orcid.org/0009-0003-2372-5397>, ✉ bakumenko_yad@edu.surgu.ru^б ORCID: <http://orcid.org/0009-0006-3239-6902>, tadzhibova_an@surgu.ru

Аннотация: машинный перевод играет ключевую роль в современном обществе, обеспечивая доступ к информации на различных языках в условиях глобализации. В данной статье проводится комплексный анализ проблем и достижений в области машинного перевода, с акцентом на использование передовых технологий, таких как нейронные сети и трансформеры, для повышения точности перевода. Рассматриваются основные метрики качества перевода, включая BLEU, METEOR и POS-BLEU, и их применение к различным системам перевода, таким как DeepL, Яндекс и Google Translate. Основное внимание уделяется сравнению эффективности этих систем с точки зрения лексического соответствия, семантической точности и синтаксической структуры. Для более глубокой оценки перевода была разработана специализированная программа на Python, которая проводит автоматический анализ перевода на основе вышеуказанных метрик. Программа выявляет слабые места в переводах и генерирует отчеты для дальнейшей оптимизации. По данным исследования, DeepL и Яндекс демонстрируют высокие результаты по метрике BLEU, что указывает на хорошее лексическое соответствие и точность синтаксической структуры. Однако все три системы сталкиваются с трудностями в области синтаксиса, лексики и использования артиклей, что требует дальнейших улучшений. В работе подчеркивается важность многогранной оценки качества перевода и необходимости интеграции машинного анализа с экспертным контролем. Также рассматривается роль этих технологий в специализированных областях, где точность перевода имеет критическое значение. Авторы отмечают необходимость дальнейших исследований в области машинного перевода для улучшения качества перевода и его адаптации к различным языковым и культурным контекстам.

Ключевые слова: машинный перевод, нейронные сети, трансформеры, многоязычные модели, оценка качества, BLEU, METEOR, POS-BLEU.

Для цитирования: Бакуменко Я. Д., Таджибова А. Н. Современные проблемы и решения в области машинного перевода. *Успехи кибернетики*. 2025;6(2):47–59.

Поступила в редакцию: 14.04.2025.

В окончательном варианте: 09.06.2025.

CURRENT CHALLENGES AND APPROACHES IN MACHINE TRANSLATION

Ya. D. Bakumenko^а, A. N. Tadzhibova^б*Surgut State University, Surgut, Russian Federation*^а ORCID: <http://orcid.org/0009-0003-2372-5397>, ✉ bakumenko_yad@edu.surgu.ru^б ORCID: <http://orcid.org/0009-0006-3239-6902>, tadzhibova_an@surgu.ru

Abstract: we conducted a comprehensive analysis of the current challenges and recent advancements in machine translation (MT), which plays a vital role in enabling multilingual communication in today's globalized world. We focused on the application of advanced technologies – neural networks and transformer architectures – to improve translation accuracy. We evaluated leading MT systems (DeepL, Yandex Translate, and Google Translate) using standard quality metrics: BLEU, METEOR, and POS-BLEU. Our evaluation emphasized lexical matching, semantic fidelity, and syntactic structure. We developed a Python-based tool that automatically assessed translation quality using these metrics. The program identified weaknesses in translations and generated detailed reports to guide optimization. Our results showed that DeepL and Yandex achieved higher BLEU scores, indicating better performance in lexical and syntactic accuracy. However, all three systems exhibited issues with syntax, vocabulary choices, and article usage, highlighting areas that require further development.

We emphasize the importance of multi-dimensional evaluation – combining automated metrics with expert assessment. We also examined MT performance in specialized domains, where high translation precision is critical. Our findings underscore the need for continued research to further improve MT systems and adapt them to diverse linguistic and cultural contexts.

Keywords: machine translation, neural networks, transformers, multilingual models, quality evaluation, BLEU, METEOR, POS-BLEU.

Cite this article: Bakumenko Ya. D., Tadzhibova A. N. Current Challenges and Approaches in Machine Translation. *Russian Journal of Cybernetics*. 2025;6(2):47–59.

Original article submitted: 14.04.2025.

Revision submitted: 09.06.2025.

Введение

Машинный перевод (МП) является необходимым в условиях современной глобализации и цифровизации, когда потребность в доступе к информации на различных языках стала неотъемлемой частью нашей жизни.

Термин *машинный перевод* понимается, по крайней мере, в двух смыслах. В узком смысле — это процесс перевода текста с одного языка на другой с использованием компьютера, реализуемый полностью или почти полностью. Машинный перевод в широком смысле — это область научных исследований, находящаяся на стыке лингвистики, математики, кибернетики, и имеющая целью построение систем, реализующих машинный перевод в узком смысле [1, с. 3].

Машинный перевод является автоматизированным процессом перевода текста с одного языка на другой с использованием алгоритмов. Идея его создания появилась в 1947 году, а первая демонстрация системы состоялась в 1954 году в штаб-квартире IBM. Основная задача машинного перевода — воспроизвести содержание текста с учетом лингвистических и культурных особенностей. Однако, в отличие от человека, компьютерные системы не способны полноценно понимать смысл и контекст, а лишь следуют заложенным алгоритмам и словарям. Это создает ряд трудностей, особенно при передаче оттенков значений, интонаций и сложных грамматических конструкций. Для эффективной работы машинного перевода требуются точные языковые правила и универсальные модели, что делает его разработку междисциплинарной задачей. Немаловажным аспектом является адаптация лингвистической терминологии и формализация языковых единиц, особенно для языков с богатой морфологией и сложным синтаксисом.

Цель исследования

Целью данного исследования является анализ проблем и решений в области машинного перевода, включая использование нейронных сетей для повышения точности перевода, а также учет культурных реалий в процессе перевода. Особое внимание уделено сравнению популярных систем машинного перевода (DeepL, Яндекс, Google Translate) и нейронной сети GPT, с акцентом на использование метрик качества перевода, таких как BLEU, METEOR и POS-BLEU.

Материалы и методы исследования

В процессе исследования использовались эмпирические данные и теоретические исследования в области машинного перевода, включая работы, посвященные нейронным сетям, трансформерам, многоязычным моделям и методам оценки качества перевода, BLEU, METEOR и другим популярным метрикам. Также были проанализированы современные системы машинного перевода, такие как Google Translate, DeepL и Яндекс.

Для более глубокой оценки качества перевода нами была разработана программа на языке Python, которая проводит автоматический анализ с использованием вышеуказанных метрик. Программа оценивает каждое предложение по метрикам BLEU, METEOR и POS-BLEU, позволяет выявить проблемные зоны в переводах и генерирует отчет, который используется для дальнейшей оптимизации и улучшения качества перевода.

Перевод — это не просто преобразование слов и предложений на другом языке, но также и учет контекста, многозначности отдельных слов. Многозначные слова особенно затрудняют задачу: одно и то же слово может означать разные вещи в зависимости от ситуации. Например, слово «ключ» в русском языке может обозначать как предмет для открытия дверей, так и источник воды. Если система не распознает правильное значение с учетом контекста, перевод может оказаться ошибочным. Аналогичные сложности возникают и с анафорическими связями, такими, как местоимения. В предложении «Иван пришел в дом, и он был пуст» местоимение «он» должно относиться к «дому», а не к «Ивану», что требует от системы способности анализировать контекст и структуру текста. Для решения этих проблем разрабатываются правила, касающиеся фразеологического анализа и распознавания контекста.

Проблемы, связанные с контекстуальной неоднозначностью, требуют внедрения и более сложных методов. Одним из таких методов является обучение на обширных объемах данных, что позволяет системе «видеть» множество примеров использования слов в различных контекстах и правильно определять их значения в зависимости от ситуации. Для этого создаются и анализируются большие текстовые корпуса, охватывающие разные стили и темы, чтобы помочь системе точнее передавать смысл фраз. Трансформеры, такие как модели BERT и GPT, значительно улучшают понимание контекста и помогают сохранять точность смысловой передачи. Данные модели анализируют текст комплексно, одновременно обрабатывая все компоненты предложения, следовательно, это дает возможность более глубоко «понимать» значения слов в контексте всего текста. Модели обучаются на параллельных текстах и затем дообучаются для более точного перевода.

Недавние эмпирические улучшения, достигнутые благодаря трансферному обучению с использованием языковых моделей, продемонстрировали, что предварительное обучение языковой модели эффективно для улучшения многих задач обработки естественного языка. К ним относятся задачи на уровне предложений, такие как логический вывод на естественном языке и перефразирование, которые направлены на предсказание взаимосвязей между предложениями путем их целостного анализа, а также задачи на уровне токенов, например, распознавание именованных объектов и ответы на вопросы, где требуются модели для получения детализированной информации [2].

Для того чтобы адаптировать перевод для самых разнообразных текстов, создаются специализированные подкорпусы данных, предназначенные для конкретных областей, что позволяет более точно учитывать специфическую терминологию и стиль. Различия в синтаксисе и морфологии между языками также требуют крайне точной настройки систем; система должна учитывать порядок слов и правила изменения окончаний, которые особенно важны для языков с развитой грамматикой.

Значительного прогресса в областях обработки языка удалось добиться на более низких уровнях обработки речи: для письменной формы языка к ним относятся сегментация текста, лексический, морфосинтаксический и синтаксический анализ; для обработки устной речи — распознавание речи, преобразование текста в речь и распознавание говорящего. В результате многие приложения теперь вошли в наше повседневное использование. В том, что касается обработки письменной речи — это проверка орфографии и грамматики, одноязычные межъязыковые поисковые системы, машинные онлайн-переводчики, в области обработки устной речи стоит отметить голосовые системы GPS, системы диктовки, транскрипцию и автоматическую индексацию аудиовизуального контента. Данный список показывает, что многие из этих приложений объединяют устную и письменную формы языка (транскрипция речевого сообщения в текст, перевод текста в речь) [2, с. 155].

Помимо многозначности, одной из сложных проблем в области машинного перевода является адекватный перевод фразеологизмов. Фразеологизмы представляют собой устойчивые лексические сочетания, значение которых определяется целым выражением, а не отдельными словами. Так, английские выражения, такие как «it's high time» (давно пора), «take your time» (не торопись) или «help yourself» (угощайся), не имеют буквального эквивалента в других языках, но их смысл можно передать соответствующими выражениями. С точки зрения смысловой связанности компоненты фразеологизмов можно разделить на образные и необразные. Необразные фразеологизмы сохраняют значение составляющих слов, но их элементы нельзя менять, как, например, в словосочетаниях «играть роль» и «иметь значение». Такие фразеологизмы чаще встречаются в научной и деловой речи.

Образные фразеологизмы, напротив, имеют значение, которое не является суммой значений отдельных слов. В языке встречаются такие примеры, как «погнаться за двумя зайцами», «попасть как кур в ошип» или «лиха беда начало» на русском языке и «through thick and thin» (во что бы то ни стало), «tooth and nail» (не жалея сил, засучив рукава), «hit the nail on the head» (попасть в точку) на английском.

Сложность машинного перевода фразеологизмов заключается в том, что их компоненты теряют свой буквальный смысл и их необходимо переводить не дословно, а с учетом культурных и языковых особенностей. В зависимости от ситуации для перевода фразеологизмов переводчик может использовать различные методы: фразеологический эквивалент, фразеологический аналог, калькирование или описательный перевод.

Машинные системы часто переводят такие фразы неверно, что приводит к искажению смысла. Чтобы решить эту проблему, современные технологии, основанные на нейронных сетях, обучаются

на больших объемах данных. Модели анализируют примеры перевода идиом в различных контекстах, чтобы подбирать подходящие аналоги на другом языке. Алгоритмы BERT и GPT способны учитывать контекст и заменять, например, «a piece of cake» на «проще простого» в русском языке. Это помогает сохранить точное значение выражений.

Следует отметить, что на 95% языков говорит только 6% населения мира. Некоторые лингвисты считают, что 90% языков исчезнут в течение столетия. Таким образом, по наличию данных о языке и автоматических систем его обработки можно выделить языки с большим или меньшим объемом ресурсов или полным их отсутствием, а также устные языки без системы письма. Наличие данных имеет решающее значение для развития используемых систем, часто основанных на статистическом подходе. Системы перевода не способны поддерживать такие языки на уровне более популярных, таких как английский или китайский. Например, языки с развитой морфологией, такие как финский или венгерский, имеют сложную структуру слова, включающую многочисленные суффиксы, падежи и склонения. Похожая проблема возникает и с диалектами, где словарный состав и синтаксические особенности могут существенно отличаться даже в пределах одного языка.

В этой связи машинный перевод требует существования параллельных корпусов, однако их число ограничено. Поэтому мы стараемся восполнить данный пробел путем разработки методов, использующих «шумные» параллельные корпуса текстов, сопоставимые корпуса (тексты одной тематики на разных языках) или квазисопоставимые корпуса, которые являются более доступными, особенно благодаря распространению Интернета [2, с. 161].

Facebook AI создала модель M2M-100, которая обучается сразу на параллельных текстах многих языков. Благодаря этому подходу она умеет работать с редкими языками, используя общие лингвистические закономерности. Например, опыт перевода между испанским и португальским помогает точнее переводить на каталонский, поскольку эти языки похожи по структуре. В отличие от традиционных систем, ориентированных на переводы через английский, M2M-100 сразу переводит между 100 языками и поддерживает 9 900 направлений. Для этого она группирует языки по близости и применяет «случайную перенаправку» параметров внутри каждой группы. Модель обучалась на большом корпусе CommonCrawl, что позволило сохранить точность даже там, где не было пар с английским языком. В итоге прямой перевод между неанглоязычными парами показал более высокие оценки BLEU по сравнению с двуязычными моделями, использующими данные WMT систем и работающими через английский.

Помимо этого, в M2M-100 внедрен метод переноса обучения: сначала модель учится на одном, «большом» языке, а затем дообучается на небольшом наборе данных редкого языка. После обучения на немецком ей проще адаптироваться к люксембургскому. Также применяется активное обучение: модель выбирает для дообучения сложные фразы, которые затем переводят носители языка или эксперты. Эти проверки постепенно наращивают корпус, что повышает качество перевода для малораспространенных языков и диалектов, поддерживая языковое разнообразие.

Следует также подчеркнуть, что в секторе информационных технологий значительные усилия в сфере развития многоязычия предпринимают и другие крупные американские компании. Одной из ведущих компаний в этом направлении является Google, которая значительно расширила количество поддерживаемых языков и улучшила качество своих переводческих сервисов. По состоянию на июнь 2024 года Google Translate поддерживает 243 языка, что на 98 языков больше, чем было доступно ранее. Этот сервис продолжает активно развиваться, добавляя новые языки и улучшая качество перевода. В 2024 году было добавлено 110 новых языков, включая редкие и исчезающие языки — кантонский, манкс, н'ко, панджаби (шахмуки), тамазигт (амазиг) и ток-писин. Благодаря этим обновлениям Google расширяет охват более чем на 614 миллионов человек по всему миру. Google активно работает и над сохранением исчезающих языков, таких как диула, волоф, бауле и других. Кроме того, Google использует новейшие модели PaLM 2, которые позволяют эффективно обучать нейросети для перевода между тесно связанными языками. Модели PaLM 2 также улучшили работу с региональными языками — башкирским, чеченским, осетинским и крымскотатарским.

Компания Microsoft также активно развивает свои языковые технологии. На июнь 2025 года Microsoft Translator поддерживает 179 языков и диалектов, включая 21 вариант испанского языка и 16 вариантов арабского. Важно отметить, что компания продолжает улучшать свои инструменты, например, с Microsoft Translator Pro, который был выпущен в ноябре 2024 года. Этот сервис теперь включает

точный перевод речи в реальном времени с поддержкой транскрипции на одном языке, улучшенной многозадачности и шумоподавления, что значительно расширяет возможности перевода в условиях динамичных ситуаций. Microsoft также предлагает поддержку офлайн-переводов в своих приложениях для Android и iOS, улучшая доступность для пользователей с ограниченным Интернет-доступом. Помимо этого, Microsoft интегрировала свои языковые инструменты в популярные офисные программы Microsoft 365, где доступна проверка орфографии и грамматики на 126 языках, включая учет региональных вариантов на 233 и 61 языках для орфографии и грамматики соответственно. Это демонстрирует стремление компании сделать свои инструменты доступными для широкой аудитории пользователей по всему миру.

Технические ограничения также играют значительную роль в определении качества перевода. Сегодня существуют различные типы систем машинного перевода, включая полностью автоматизированные, машинно-ассистируемые переводы (МАНТ и НАМТ) и системы переводческой памяти (Translation Memory, TM). В зависимости от роли человека в процессе перевода системы могут по-разному справляться с задачей, однако в каждом случае на точность влияет ряд факторов. Например, системы TM, которые широко используются профессиональными переводчиками (DL Trados Studio, MemoQ, Wordfast, OmegaT, Across и др.), позволяют сохранять уже переведенные фразы или части текста, чтобы предложить их для повторного использования. Это снижает трудоемкость перевода и улучшает качество за счет использования проверенных переводов.

С переходом к нейронному машинному переводу (NMT) технологии стали учитывать более широкий контекст. Первая волна NMT базировалась на архитектуре «энкодер–декодер» с рекуррентными нейронными сетями, однако с развитием Transformer-модели в 2017 году качество перевода значительно возросло благодаря механизму внимания. В отличие от более старых систем машинного перевода, NMT устранил многие недостатки предыдущих подходов, а именно проблемы со связанностью текста. Однако, несмотря на значительный прогресс, полностью автоматизированные системы (например, Google Translate) и другие сервисы все еще сталкиваются с рядом ограничений: трудности при переводе специализированных текстов, дефицит параллельных данных и проблемы с доменным рассогласованием (ситуацией, при которой тематическая, жанровая или стилевая область “domain” текстов, на которых обучалась модель машинного перевода, не совпадает с областью входных или целевых текстов, которые системе приходится переводить на практике).

В процессе совершенствования системы машинного перевода компания Яндекс применяет несколько ключевых методов, направленных на повышение точности перевода, особенно с русского языка. Одним из таких методов является использование морфологических трансформаций. Основное преимущество этого подхода заключается в том, что он помогает преобразовать русские предложения в более удобные для перевода формы. Как известно, русский язык обладает богатой морфологией, и это часто является причиной трудностей в переводе. Применение морфологических трансформаций решает проблему изменяющихся окончаний и форм слов, что, в свою очередь, существенно улучшает точность перевода.

Кроме того, следует отметить, что Яндекс использует модели последовательностей операций (Operation Sequence Model). Этот подход интегрирует операции переноса и лексические переводы в последовательности минимальных единиц перевода. Такой метод упрощает процесс перевода на уровне синтаксиса и порядка слов, что особенно важно при переводе между языками с различной грамматической структурой. Обратим внимание на то, что этот подход позволяет решить одну из основных проблем машинного перевода, связанную с различиями в грамматических системах. В дополнение к вышеупомянутым методам Яндекс использует подход трансфем (transfemes), который преобразует слова в более структурированные единицы перевода. Это, безусловно, помогает снизить сложность перевода между языками с разными грамматическими и синтаксическими особенностями. Применение этих методов значительно улучшает точность перевода, что подтверждается результатами экспериментов, проведенных на тестах WMT13 и WMT14. Результаты этих тестов показали повышение показателей BLEU.

Не менее важным инструментом в области машинного перевода является переводчик DeepL от одноименной немецкой компании, специализирующейся на разработке технологий машинного перевода с использованием искусственного интеллекта. В исследовании Линлина Ли 2024 г. был проведен сравнительный анализ качества перевода с использованием DeepL, Google Translate и Microsoft

Translator. Оценка проводилась по таким параметрам, как точность, плавность и естественность, с применением стратегии «расчета семантического сходства».

Одним из важных выводов исследования стало то, что DeepL продемонстрировал 94% точности перевода, что значительно превосходит результаты Google Translate (86%). В исследовании использовались различные текстовые корпуса и точность перевода была оценена в общем контексте, без уточнения конкретной языковой пары. В части плавности перевода DeepL показал более высокий результат, достигнув 146% в изменении длины предложений. DeepL также оказался лидером среди трех сервисов, особенно в отношении культурной адаптации (92%) и использования идиоматических выражений (94%) [18, с. 715–716].

Процесс работы DeepL включает несколько стадий. На первом этапе текст разделяется на категории, такие как профессиональная терминология, общий словарный запас и простые предложения. Это позволяет точнее выбирать и обрабатывать слова, что, в свою очередь, улучшает качество перевода. На следующем этапе DeepL использует технологии тегирования частей речи и синтаксического анализа для грамматической корректности перевода. Как правило, ошибки на этом этапе могут привести к некорректному переводу, поэтому важно обеспечить правильную обработку синтаксиса и контекста. Для оценки качества перевода DeepL использует методику, основанную на расчете семантического сходства. Это позволяет объективно сравнить перевод с оригиналом и другими переводами. В результате система демонстрирует превосходство в точности перевода, особенно для более стилистически сложных текстов. Однако исследование также показало, что Google Translate имеет небольшое преимущество в адаптации к контексту и культурным особенностям. Это открывает возможности для дальнейших улучшений в области машинного перевода.

Перспективы развития машинного перевода тесно связаны с нейронными сетями, а именно с трансформерами seq2seq, BERT, T5 и GPT. Эти модели позволяют значительно точнее моделировать контекст, справляясь с трудными синтаксическими структурами. Языковые модели, такие как GPT, показывают высокий уровень генерации текста, но, несмотря на это, они все еще сталкиваются с проблемами точности при интерпретации сложных структур. Это, в свою очередь, требует дальнейших улучшений, таких как fine-tuning для специфических задач и Zero-Shot Learning, который позволит моделям обучаться на новых языках с минимальными данными.

Будущее машинного перевода заключается в развитии гибридных и интерактивных подходов, которые объединяют нейронные и статистические методы для повышения точности. Интерактивные системы, где пользователи могут вносить коррективы в перевод, позволяют адаптировать его под специфические задачи, особенно в таких областях, как юриспруденция или медицина, где требуется высокая точность. В дополнение к этому ожидается развитие мультимедийных моделей, которые объединяют текст и изображения, а также адаптируются под контекст пользователя. Это откроет новые возможности для профессиональных приложений, что, безусловно, повысит качество и доступность машинного перевода. Более того, для отслеживания прогресса в сфере машинного перевода используются различные системы оценки, такие как BLEU (Bilingual Evaluation Understudy) и METEOR (Metric for Evaluation of Translation with Explicit ORdering). Метрика BLEU, одна из первых автоматизированных метрик, измеряет совпадение между машинным переводом и эталонными переводами на основе вычисления n-грамм (последовательностей слов). Чем больше таких совпадений, тем выше оценка, что указывает на схожесть с человеческим переводом. METEOR предлагает более гибкий подход, принимая во внимание лексические варианты — синонимы и морфологические изменения слов. Она включает этапы выравнивания слов, подсчет совпадений по смыслу, а также штрафы за отклонение от порядка слов, что делает ее более чувствительной к качеству перевода.

Современные исследователи работают над усовершенствованием этих метрик, добавляя учет синтаксических и семантических аспектов, чтобы они могли оценивать точность и естественность перевода более комплексно. В дополнение к традиционным метрикам оценки качества машинного перевода, которые фокусируются на лексическом и частично семантическом соответствии между эталонным и машинным переводами, все большее внимание уделяется метрикам, основанным на синтаксическом анализе. Такие метрики позволяют оценить, насколько хорошо машинный перевод сохраняет грамматическую структуру исходного текста. Они анализируют соответствие синтаксических деревьев перевода и референтного текста, выявляя отклонения и ошибки. Метрика POS tagging анализирует последовательность частей речи (POS), используемых в машинном переводе, и сравнивает ее с референт-

ным текстом. Она оценивает, насколько корректно соблюдены грамматические правила языка перевода. Эти улучшения позволяют моделям машинного перевода ориентироваться на более высокий уровень адекватности и плавности перевода.

Результаты исследования и их обсуждение

В рамках данного исследования была разработана программа на языке Python. Программа предназначена для автоматической оценки качества машинного перевода с использованием трех популярных метрик — BLEU, METEOR и POS-BLEU. Она использует несколько библиотек Python, включая nltk, Stanza и docx, чтобы провести анализ перевода. Эта программа анализирует перевод с точки зрения лексической точности, семантического соответствия и синтаксической структуры, что делает ее полезной для оценки и улучшения качества машинного перевода.

Программа начинает с чтения текстов перевода, которые необходимо оценить. Эталонный перевод и машинный перевод загружаются из текстовых файлов с помощью функции `read_translation_file`. Эта функция открывает файл и читает его содержимое, затем с помощью токенизатора из библиотеки nltk разбивает текст на предложения:

```
def read_translation_file(file_path):
    with open(file_path, 'r', encoding='utf-8') as file:
        text = file.read().strip()
    tokenizer = nltk.data.load('tokenizers/punkt/russian.pickle')
    sentences = tokenizer.tokenize(text)
    return sentences
```

Библиотека nltk применяется для токенизации текста, что позволяет делить текст на отдельные предложения. Токенизатор используется для обработки русского языка.

После токенизации текста каждое предложение разбивается на слова с использованием стандартного метода `.split()` в Python. Это необходимо для последующего анализа на уровне слов и фраз, что позволяет метрикам BLEU и METEOR корректно вычислять оценки:

```
def tokenize_and_split(sentences):
    return [sentence.split() for sentence in sentences]
```

POS tagging (частеречная разметка)

Для анализа синтаксической структуры перевода используется библиотека Stanza, которая выполняет POS-тегирование. В процессе этого анализа каждому слову в тексте присваивается соответствующая часть речи (например, существительное, глагол, прилагательное):

```
import stanza
stanza.download('ru') # Загрузка модели для русского языка
nlp = stanza.Pipeline('ru', processors='tokenize,pos', tokenize_pretokenized=False)
def pos_tag_sentences(sentences, nlp):
    tagged_sentences = []
    for sentence in sentences:
        doc = nlp(sentence)
        tags = [word.upos for sent in doc.sentences for word in sent.words]
        tagged_sentences.append(tags)
    return tagged_sentences
```

Модель `stanza.Pipeline` используется для обработки текста на русском языке и возвращает последовательности тегов частей речи, что необходимо для метрики POS-BLEU. После подготовки данных программа оценивает каждый перевод с использованием трех метрик.

BLEU оценивает лексическое совпадение между эталонным переводом и машинным переводом на основе n-грамм. Метрика BLEU используется для оценки точности перевода:

```
bleu_score = corpus_bleu(reference_corpus, machine_translation)
```

Для каждого предложения также вычисляется BLEU-оценка с помощью функции `sentence_bleu` из библиотеки nltk:

```
sent_bleu = sentence_bleu([ref], mt)
```

METEOR — более гибкая метрика, которая учитывает синонимы и морфологические изменения слов, а также порядок слов. METEOR лучше подходит для оценки семантического соответствия перевода:

```
score = single_meteor_score(ref_tokens, mt_tokens)
```

POS-BLEU оценивает синтаксическую правильность перевода, сравнивая POS-теги эталонного перевода и машинного перевода:

```
pos_bleu = sentence_bleu([ref_tags], mt_tags)
```

Программа проверяет, соответствует ли BLEU, METEOR и POS-BLEU пороговым значениям:

```
if sent_bleu < low_bleu_threshold:
    print(f'Низкий BLEU для предложения {i+1}: {sent_bleu:.4f}')
    print('Эталон:', ' '.join(ref))
    print('Машинный перевод:', ' '.join(mt))
    print('---')
```

Если оценка для предложения ниже заданного порога (например, 0.5), программа добавляет это предложение в отчет и выделяет его красным в документе, используя библиотеку docx.

После вычисления всех метрик и выделения низких оценок программа создает отчет в формате .docx. В отчете содержится общая информация о результатах анализа и подробности для каждого предложения:

```
document = Document()
document.add_heading('Отчет по оценке машинного перевода', 0)
document.add_heading('Общая оценка', level=1)
document.add_paragraph(f'Корпусный BLEU: {bleu_score:.4f}')
document.add_paragraph(f'Порог BLEU: {low_bleu_threshold:.4f}')
document.add_paragraph(f'Средний METEOR: {avg_meteor_score:.4f}')
document.add_paragraph(f'Порог METEOR: {low_meteor_threshold:.4f}')
document.add_paragraph(f'Корпусный POS-BLEU: {pos_bleu_corpus_score:.4f}')
document.add_paragraph(f'Порог POS-BLEU: {low_pos_bleu_threshold:.4f}')
```

В отчете отображаются все вычисленные значения для каждой метрики, а также указываются предложения с низкими оценками, выделенные красным цветом.

Программа позволяет проводить комплексный анализ качества перевода, выявляя слабые места перевода по каждой из метрик. Примером работы программы может служить анализ перевода научной статьи, выполненной с использованием различных машинных переводческих сервисов. Результаты показывают, что, несмотря на высокие оценки по метрике BLEU, некоторые предложения могут иметь низкую оценку по METEOR или POS-BLEU, что позволяет глубже понять, где именно требуется улучшение.

Программа обеспечивает автоматизацию процесса анализа качества перевода, позволяя пользователю получить полный отчет, который может быть использован для улучшения перевода и оптимизации работы машинных переводческих сервисов. Благодаря тому, что программа выделяет низкие оценки, она способствует более эффективному выявлению ошибок и упрощает анализ качества перевода, предоставляя ясную и структурированную информацию для дальнейшего использования.

Для оценки качества перевода была выбрана научная статья «Использование машинного обучения и нейронных сетей для автоматической верификации заданий в текстовом и графическом представлении и помощи преподавателю» авторов Н. О. Бесшапошников, М. С. Дьяченко и др., переведенная с русского языка на английский. Оценка качества перевода проводилась с использованием Google Translate, Яндекс Переводчика, DeepL и ChatGPT, а именно, модели o4-mini.

Таблица

Результаты анализа

	Яндекс переводчик	Google Translate	DeepL	ChatGPT
BLEU	0.5209	0.5095	0.3248	0.3074
METEOR	0.7395	0.7197	0.4806	0.5901
POS tagging	0.8967	0.8971	0.7942	0.8677
Среднее значение	0.7190	0.7088	0.5332	0.5884

В ходе проведенного исследования было установлено, что различия между сервисами могут оказывать значительное влияние на качество перевода. Использование трех метрик позволило получить всестороннюю картину эффективности этих сервисов при переводах с русского языка на английский язык.

Исследование показало, что Яндекс и Google демонстрируют лучшие результаты по метрике BLEU, что свидетельствует о высоком уровне лексического соответствия и синтаксической структуры перевода. Однако разница между ними минимальна. GPT показал более низкие значения по BLEU, что говорит о большем количестве перефразирований и изменении структуры предложения. Важно отметить, что Яндекс и Google также лидируют и по METEOR, что подтверждает более высокое сохранение смысла перевода в сравнении с DeepL.

Яндекс продемонстрировал отличные результаты по переводу научной статьи, обеспечивая высокий уровень точности, особенно при передаче контекста и синтаксической структуры на уровне предложения. В то же время Google показал более высокую точность в переводе специфических эквивалентов, где точность транслитерации и культурная адаптация имеют ключевое значение. Это указывает на то, что Яндекс лучше справляется с передачей общего контекста и структуры, тогда как Google показывает более высокую эффективность в переводе узкоспециализированных терминов.

Различия в производительности этих сервисов могут быть связаны с различиями в их архитектурах. DeepL использует исключительно нейронные сети, что позволяет модели анализировать весь контекст предложения комплексно. В свою очередь, Яндекс применяет гибридный подход, комбинируя нейронные сети и статистические методы, что делает его особенно эффективным при переводах, где важно правильно обработать редкие слова и специфические выражения. Google, как и Яндекс, использует гибридный подход, но его результаты по метрике BLEU и другим метрикам слегка уступают, что также указывает на возможные различия в подходах к обучению и архитектуре. Более того, мы можем предположить, что различия в результатах этих систем могут также быть обусловлены различиями в объемах и характере обучающих данных. Например, Яндекс, ориентированный на русскоязычные тексты, имеет более глубокое понимание структуры и специфики русского языка, что способствует большей точности перевода в данной языковой паре. В то время как DeepL, несмотря на высокие результаты в переводах между западноевропейскими языками, имеет меньший фокус на русском языке, что ограничивает его эффективность в переводах с и на русский. Что касается Google, его результаты несколько уступают Яндексу, вероятно, из-за меньшего объема выборки для русского языка в обучении системы, а также особенностей алгоритмов обучения, ориентированных на более широкий спектр языков. Это может объяснять, почему Google менее эффективен при переводах с русского языка.

Результаты ChatGPT по всем метрикам заметно хуже, чем у специализированных сервисов перевода. Это, вероятно, связано с тем, что ChatGPT не является специализированным переводчиком и не был обучен с учетом fine-tuning для задач машинного перевода. Хотя ChatGPT может генерировать текст на различных языках, его основная цель — это диалоговое взаимодействие и генерация осмысленных текстов, а не перевод. Отсутствие адаптации под конкретные задачи перевода, такие как работа с техническими текстами или специализированной лексикой, ограничивает точность его переводов.

Несмотря на высокие результаты по BLEU и METEOR, все системы сталкиваются с определенными проблемами в области синтаксиса, лексики и использования артиклей. Например, машинный перевод продолжает испытывать трудности с точной передачей сложных грамматических структур, что иногда приводит к потерям в точности перевода. Ошибки на уровне синтаксиса и лексики делают переводы менее естественными. В то же время стоит отметить, что системы машинного перевода

не всегда соблюдают тема-рематическое членение, что также может привести к искажению смысла текста. Нарушение правильного распределения информации по теме и реме может менять акценты в предложении, создавая путаницу или теряя важные детали контекста. Для исправления таких ошибок необходима дополнительная экспертная доработка, поскольку они требуют глубокого понимания текста и его структуры, что машины пока не могут обеспечить на должном уровне.

В ходе проведенного исследования также следует обратить внимание, что метрики оценки качества машинного перевода могут сильно зависеть от эталонного (экспертного) перевода и часто занижают оценки, если допущен сильный стилистический или лексический отход от оригинала. Это особенно актуально для метрики BLEU, которая ориентирована на точное совпадение слов и фраз между машинным переводом и эталонным переводом, не учитывая синонимы и перефразирование. Такая зависимость от эталонного перевода снижает объективность данной метрики, особенно при ее использовании в одиночку.

Примером может служить следующий случай:

Оригинал: «Существенной помощью может стать использование интеллектуальных помощников преподавателя и автоматизированных систем проверки, построенных методами машинного обучения и технологии нейронных сетей». *Экспертный перевод:* «Smart educator assistants and automated assessment tools based on machine learning and neural networks can significantly alleviate the problem». *Перевод Google:* «The use of intelligent teacher assistants and automated checking systems built using machine learning methods and neural network technology can be of significant help». Итоговая оценка BLEU: 0.0000 (но METEOR = 0.4808, POS-BLEU = 0.7440).

Метрика BLEU присудила переводу нулевую оценку, но она не отражает глубины ошибок, связанных с неправильным выбором лексических эквивалентов. Грамматическая метрика POS-BLEU показала, что грамматическая составляющая перевода сохранена, однако это не свидетельствует о полном соответствии перевода. В данном случае именно METEOR, учитывающий синонимы и перефразирование, оказался более информативным. Несмотря на то, что перевод Google в целом сохраняет адекватность передачи общего смысла текста, в нем не была обеспечена эквивалентность лексических единиц. Например, словосочетание «системы проверки» была переведено как «checking systems», что является ошибкой. Вместо этого следовало использовать «assessment tools», что точно соответствует контексту и смыслу оригинала.

В рамках нашего исследования дополнительно был взят текст рецензируемой статьи из области медицины, а также статья с экспертным переводом, который ранее не был опубликован в сети Интернет. Эти тексты были использованы для того, чтобы оценить эффективность машинного перевода в различных контекстах и с различной сложностью лексики. Все оценки, полученные по трем метрикам, показали высокие результаты, со средним значением (0.6896), что свидетельствует о хорошем качестве перевода. При этом важно отметить, что корреляции между переводами статей, уже имеющих экспертные переводы в сети, и текстами, ранее не опубликованными, не было обнаружено. Гипотеза о том, что сервисы машинного перевода используют уже готовые переводы из сети, не была подтверждена результатами нашего анализа. Напротив, оказалось, что системы машинного перевода способны генерировать уникальные переводы даже для статей, не имеющих предварительных переводов в сети.

Системы машинного перевода показали достаточно высокую эффективность, что, вероятно, связано с функциональным стилем выбранных нами текстов. Научные статьи имеют строгую структуру, фиксированную грамматику и часто используют интернациональные термины, что делает их легче для перевода. Кроме того, в научных текстах отсутствует имплицитность и слова не имеют двойственного значения. Однако перевод медицинских диагнозов, заключений и специфической терминологии остается более сложной задачей. Несмотря на это, современные сервисы машинного перевода достаточно адекватно справляются с научными текстами. Важно отметить, что для языковой пары русский-английский лучшим сервисом по результатам нашего исследования оказался Яндекс, что подтверждается его высокими показателями по метрикам качества перевода. Развитие технологий машинного перевода открыло новые возможности для глобальных взаимодействий, обмена опытом и научными открытиями, но тем не менее необходимо проводить постредактуру переводов экспертами, способными выявить и исправить ошибки, которые машина не в силах корректно интерпретировать.

Автоматическая оценка в сочетании с экспертным анализом позволяет повысить точность проверки переводов. Автоматизированные инструменты помогают быстро выявить ключевые ошибки и

несоответствия, освобождая переводчиков и редакторов от рутинных задач и позволяя им сосредоточиться на исправлении значимых недостатков. Тем самым сочетание машинного анализа и экспертного контроля делает процесс перевода более эффективным и надежным.

Заключение

Машинный перевод, бесспорно, занимает важное место в эпоху глобализации, облегчая доступ к информации. Современные системы МП, основанные на нейронных сетях и трансформерах, демонстрируют значительные успехи в воспроизведении сложных языковых структур, но также сталкиваются с множеством вызовов, таких как передача культурных нюансов, идиоматических выражений и перевод для малораспространенных языков. Разработки в области многоязычных моделей, трансферного обучения и активного обучения открывают новые возможности для улучшения качества перевода редких языков и диалектов. Будущее машинного перевода связано с развитием гибридных и интерактивных систем, которые учитывают пользовательский контекст и адаптируются к специфике текста. Метрики автоматической оценки продолжают совершенствоваться, позволяя более точно оценивать качество перевода. Включение синтаксических и семантических аспектов в эти метрики помогает лучше отслеживать прогресс, делая машинный перевод адекватным.

Для повышения качества машинного перевода необходимы дальнейшие исследования и усовершенствования в области лексического анализа, морфологии и синтаксиса, а также разработка методов, способных лучше справляться с переводом редких и сложных слов, особенно в специализированных областях. Оценка машинного перевода должна быть многогранной, и использование нескольких метрик, таких как BLEU, METEOR и POS-BLEU, является важным инструментом для повышения точности и адекватности перевода. Постоянное совершенствование этих технологий требует интеграции автоматической оценки с экспертным анализом, что поможет улучшить качество переводов и снизить количество ошибок.

Машинный перевод продолжает развиваться и демонстрирует значительные успехи в области обработки сложных языковых структур. Однако важно продолжать исследования для улучшения адаптации к редким языкам и сохранения точности перевода, особенно в контексте культурных различий.

ЛИТЕРАТУРА

1. Воронович В. В. *Машинный перевод: конспект лекций по специальности «Компьютерная лингвистика, Компьютерное обучение языкам»*. Минск: Белорусский государственный университет; 2013. 39 с.
2. *Net.lang: на пути к многоязычному киберпространству*: пер. с англ. / ред. перевода Е. И. Кузьмин, А. В. Паршакова. М.: Межрегиональный центр библиотечного сотрудничества; 2014. 464 с.
3. Jawahar G., Sagot B., Seddah D. What Does BERT Learn about the Structure of Language? *ACL 2019-57th Annual Meeting of the Association for Computational Linguistics*. 2019.
4. Devlin J. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805. 2018.
5. Rogers A., Kovaleva O., Rumshisky A. A Primer in BERTology: What We Know about How BERT Works. *Transactions of the Association for Computational Linguistics*. 2021;8:842–866.
6. Shukla A. et al. An Evaluation of Google Translate for Sanskrit to English Translation via Sentiment and Semantic Analysis. *Natural Language Processing Journal*. 2023;4:100025.
7. Caswell I. 110 New Languages are Coming to Google Translate. *Google Translate Blog*. 2024.
8. Fan A. et al. Beyond English-Centric Multilingual Machine Translation. *Journal of Machine Learning Research*. 2021;22:1–48.
9. Koehn P., Knowles R. *Six Challenges for Neural Machine Translation*. arXiv:1706.03872. 2017.
10. Maruf S., Saleh F., Haffari G. A Survey on Document-Level Neural Machine Translation: Methods and Evaluation. *ACM Computing Surveys (CSUR)*. 2021;54(2):1–36.
11. Stahlberg F. *Neural Machine Translation: A Review and Survey*. arXiv:1912.02047. 2019.
12. Fitria T. N. Performance of Google Translate, Microsoft Translator, and DeepL Translator: Error Analysis of Translation Result. *Al-Lisan: Jurnal Bahasa*. 2023;8(2):115–138.
13. Varela-Salinas M. J. et al. Google Translate and DeepL: Breaking Taboos in Translator Training. Observational Study and Analysis. *Ibérica*. 2023;45:243–266.

14. Borisov A., Galinskaya I. Y. Yandex School of Data Analysis Russian-English Machine Translation System for WMT14. *Proceedings of the Ninth Workshop on Statistical Machine Translation*. 2014:66–70.
15. Cambedda G., Di Nunzio G. M., Nosilia V. A Study on Automatic Machine Translation Tools: A Comparative Error Analysis Between DeepL and Yandex for Russian-Italian Medical Translation. *Umanistica Digitale*. 2021;10:139–163.
16. Kulikov V., Kulikova V., Yerkebulan G. Google/Yandex Translation Detection in the Patterns Identifying System of Multilingual Texts. *International Journal of Computers*. 2021;20(1):72–77.
17. Kamaluddin M. I. et al. Accuracy Analysis of DeepL: Breakthroughs in Machine Translation Technology. *Journal of English Education Forum (JEEF)*. 2024;4(2):122–126.
18. Linlin L. Artificial Intelligence Translator DeepL Translation Quality Control. *Procedia Computer Science*. 2024;247:710–717.
19. Sriram A. et al. *Cold Fusion: Training seq2seq Models Together with Language Models*. arXiv:1708.06426. 2017.
20. Egonmwan E., Chali Y. Transformer and seq2seq Model for Paraphrase Generation. *Proceedings of the 3rd Workshop on Neural Generation and Translation*. 2019:249–255.
21. Li Z. et al. Seq2seq Dependency Parsing. *Proceedings of the 27th International Conference on Computational Linguistics*. 2018:3203–3214.
22. Sutskever I. *Sequence to Sequence Learning with Neural Networks*. arXiv:1409.3215. 2014.
23. Ni J. et al. *Sentence-T5: Scalable Sentence Encoders from Pre-trained Text-to-Text Models*. arXiv:2108.08877. 2021.
24. Chen G. et al. Towards Making the Most of Cross-Lingual Transfer for Zero-Shot Neural Machine Translation. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. 2022;1:142–157.
25. Rothman D. *Transformers for Natural Language Processing: Build, Train, and Fine-Tune Deep Neural Network Architectures for NLP with Python, Hugging Face, and OpenAI's GPT-3, ChatGPT, and GPT-4*. Birmingham: Packt Publishing Ltd; 2022.
26. Maltas S. I. et al. *Efficient Finetuning Strategies for Multilingual Neural Machine Translation*. Universitat Politècnica de Catalunya; 2024.
27. Bengesi S. et al. Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers. *IEEE Access*. 2024;PP(99):1–1.
28. Achiam J. et al. *GPT-4 Technical Report*. arXiv:2303.08774. 2023.
29. Stap D., Araabi A. ChatGPT is Not a Good Indigenous Translator. *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP)*. 2023:163–167.
30. Agarwal A., Lavie A. Meteor, M-BLEU and M-TER: Evaluation Metrics for High-Correlation with Human Rankings of Machine Translation Output. *Proceedings of the Third Workshop on Statistical Machine Translation*. 2008:115–118.
31. Babych B., Hartley T. Extending the BLEU MT Evaluation Method with Frequency Weightings. *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*. 2004:621–628.
32. Chen B., Cherry C. A Systematic Comparison of Smoothing Techniques for Sentence-Level BLEU. *Proceedings of the Ninth Workshop on Statistical Machine Translation*. 2014:362–367.
33. Papineni K. et al. BLEU: a Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. 2002:311–318.
34. Pérez D., Alfonseca E. Application of the BLEU Algorithm for Recognising Textual Entailments. *Proceedings of the First Challenge Workshop Recognising Textual Entailment*. 2005:9–12.
35. Banerjee S., Lavie A. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*. 2005:65–72.
36. Denkowski M., Lavie A. Extending the METEOR Machine Translation Evaluation Metric to the Phrase Level. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. 2010:250–253.
37. Denkowski M., Lavie A. Meteor Universal: Language-Specific Translation Evaluation for Any Target

- Language. *Proceedings of the Ninth Workshop on Statistical Machine Translation*. 2014:376–380.
38. Бешапошников Н. О., Дьяченко М. С., Леонов А. Г., Матюшин М. А., Орловский А. Е. Использование машинного обучения и нейронных сетей для автоматической верификации заданий в текстовом и графическом представлении и помощи преподавателю. *Успехи кибернетики*. 2020;1(2):35–41. DOI: 10.51790/2712-9942-2020-1-2-4.
39. Wang F. Methods for Evaluating the Translation Quality of Artificial Intelligence Translator DEEPL Based on Multi-Translation Parallel Corpus. *2024 First International Conference on Software, Systems and Information Technology (SSITCON)*. 2024:1–5.
40. Мыльников А. Г., Климов А. Е., Курбанниёзов Т. Ш., Буйместру Н. В., Черняева А. А., Гусарова Т. А. Экстренная тотальная дуоденопанкреатэктомия по поводу профузного желудочно-кишечного кровотечения, вызванного метастазами рака почки в поджелудочную железу. *Вестник Российского университета дружбы народов. Серия: Медицина*. 2023;27(2):246–253. DOI: 10.22363/2313-0245-2023-27-2-246-253.

ЛИНЕЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ И СТРУКТУРЫ ФИБОНАЧЧИ

Г. Е. Деев^{1,a}, С. В. Ермаков^{1,b}, S. Starkloff^{2,c}¹ Обнинский институт атомной энергетики, Национальный исследовательский ядерный университет «МИФИ», г. Обнинск, Российская Федерация² Базельский университет, г. Базель, Швейцария^a  georgdeo@mail.ru, ^b ermakov@iate.obninsk.ru ^c s.starkloff@stud.unibas.ch,

Аннотация: показано, что решение линейного разностного уравнения любого порядка может быть записано с помощью структуры Фибоначчи, определение которой дается по ходу изложения. Структура Фибоначчи представляет интерес сама по себе. Она состоит из последовательностей чисел, родственных хорошо известным числам Фибоначчи. С каждой такой последовательностью связано число Фидия. В статье приведены некоторые результаты относительно свойств структуры Фибоначчи и чисел Фидия.

Ключевые слова: разностные уравнения, числа Фибоначчи, числа Фидия.

Для цитирования: Деев Г. Е., Ермаков С. В., Starkloff S. Линейные разностные уравнения и структуры Фибоначчи. *Успехи кибернетики*. 2025;6(2):60–66.

Поступила в редакцию: 16.05.2025.

В окончательном варианте: 02.06.2025.

LINEAR DIFFERENCE EQUATIONS AND FIBONACCI STRUCTURES

G. E. Deev^{1,a}, S. V. Ermakov^{1,b}, S. Starkloff^{2,c}¹ Obninsk Institute for Nuclear Power Engineering, National Research Nuclear University MEPhI, Obninsk, Russian Federation² University of Basel, Basel, Switzerland^a  georgdeo@mail.ru, ^b ermakov@iate.obninsk.ru ^c s.starkloff@stud.unibas.ch

Abstract: we studied the solution of linear difference equations of arbitrary order and demonstrated that they can be expressed using the Fibonacci structure, which we defined in the course of our presentation. This structure, which consists of number sequences related to the classical Fibonacci numbers, is of independent interest. Each sequence in the structure has an associated Phidias number. In this paper, we present several results on the properties of the Fibonacci structure and Phidias numbers.

Keywords: difference equations, Fibonacci numbers, Phidias numbers.

Cite this article: Deev G. E., Ermakov S. V., Starkloff S. Linear Difference Equations and Fibonacci Structures. *Russian Journal of Cybernetics*. 2025;6(2):60–66.

Original article submitted: 16.05.2025.

Revision submitted: 02.06.2025.

Постановка задачи

Сеточная функция $y_i = y(x_i)$, $i \in Z = \{\dots - 2, -1, 0, 1, 2, \dots\}$, определенная на сетке $\omega = \{x | x = x_i, x_i = x_0 + i, i \in Z\}$, удовлетворяет уравнению

$$y_{n+k} = a_{k-1}y_{n+k-1} + \dots + a_1y_{n+1} + a_0y_n + f_n, \quad n \in Z \quad (1)$$

и начальным условиям

$$y_0 = \bar{y}_0, \quad y_1 = \bar{y}_1, \dots, y_{k-1} = \bar{y}_{k-1}, \quad (2)$$

где $\bar{y}_0, \bar{y}_1, \dots, \bar{y}_{k-1}$ — известные числа; f_n — известная функция аргумента n .

Задача состоит в нахождении y_n для всех n [1–4].

Решение задачи

Используя известный прием [4, 6], запишем задачу в векторно-матричном виде:

$$\begin{cases} \vec{y}_{n+1} = A\vec{y}_n + \vec{f}_n, & (n \in Z) \end{cases} \quad (1)$$

$$\begin{cases} \vec{y}_0 = \begin{pmatrix} \bar{y}_{k-1} \\ \vdots \\ \bar{y}_1 \\ \bar{y}_0 \end{pmatrix}, \end{cases} \quad (2)$$

где

$$\vec{y}_n = \begin{pmatrix} y_{n+k-1} \\ \dots \\ y_{n+1} \\ y_n \end{pmatrix}_{k \times 1}, \quad \vec{f}_n = \begin{pmatrix} f_n \\ 0 \\ \dots \\ 0 \end{pmatrix}_{k \times 1}, \quad A = A(k) = \begin{pmatrix} a_{k-1} & a_{k-2} & \dots & a_1 & a_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{k \times k}. \quad (3)$$

Из равенства (1) вытекает

$$\vec{y}_n = A^n \cdot \vec{y}_0 + \sum_{i=0}^{n-1} A^{n-1-i} \cdot \vec{f}_i, \quad (4)$$

откуда видно, что для нахождения решения надо знать степени матрицы A .

Параметр k определяет порядок разностного уравнения. Как устроены степени матрицы A для различных значений k , мы понимаем после разбора случаев $k = 2$ и $k = 3$. Поэтому опишем устройство матрицы $A^n(k)$ для произвольных k и запишем решение задачи (1)–(3).

Степени матрицы $A(k)$

То, как устроены степени матрицы $A^n(k)$ ($n = 0, 1, 2, \dots$) [5], становится понятно после разбора простых случаев, например, $k = 2$ и $k = 3$. Наблюдается интересная особенность, состоящая в том, что для описания степеней матрицы в полном объеме (т.е. чтобы знать все ее элементы) достаточно знать только верхние угловые элементы b_{11} как самой матрицы $A^n(k)$, так и некоторых ее предыдущих степеней. Причем все эти элементы устроены по принципу подобия и, следовательно, достаточно иметь представление хотя бы об одном из них, чтобы написать все остальные. Но мы имеем представление об элементе $b_{11}^n(k)$, имеем точную формулу, его описывающую. Для случаев $k = 2$ и $k = 3$ эти формулы таковы.

Случай $k = 2$:

$$b_{11}^n = b_{11}^n(2) = \sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0 \\ \nu_1 + 2\nu_2 = n}} \frac{(\nu_1 + \nu_2)!}{\nu_1! \cdot \nu_2!} \cdot a_1^{\nu_1} \cdot a_0^{\nu_2}. \quad (1)$$

Случай $k = 3$:

$$b_{11}^n = b_{11}^n(3) = \sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0, \nu_3 \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 = n}} \frac{(\nu_1 + \nu_2 + \nu_3)!}{\nu_1! \cdot \nu_2! \cdot \nu_3!} \cdot a_2^{\nu_1} \cdot a_1^{\nu_2} \cdot a_0^{\nu_3}. \quad (2)$$

На случай произвольного k эти формулы обобщаются:

$$b_{11}^n = b_{11}^n(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k}. \quad (3)$$

В этой формуле n принимает все целые значения, начиная с нуля: $n = 0, 1, 2, \dots$, а $k = 2, 3, \dots$. В дальнейшем эти значения будут дополнены. Доказательство формулы (4) проводится по индукции. Далее выясняется, что по элементам типа b_{11} могут быть расписаны все элементы первой строки матрицы $A^n(k)$ по приведенной далее схеме (5).

Используя эту схему, пишем первую строку матрицы (для лучшего зрительного восприятия элементы строки заключены в скобки):

$$\left\{ b_{11}^n = b_{11}^{n-1} \cdot a_{k-1} + b_{11}^{n-2} \cdot a_{k-2} + \dots + b_{11}^{n-(k-1)} \cdot a_1 + b_{11}^{n-k} \cdot a_0 \right\} \dots \left\{ b_{1,k-1}^n = b_{11}^{n-1} \cdot a_1 + b_{11}^{n-2} \cdot a_0 \right\} \left\{ b_{1,k}^n = b_{11}^{n-1} \cdot a_0 \right\}. \quad (4)$$

Фигурирующие в (4) элементы типа b_{11} вычисляются по формулам (3). Таким образом, первая строка матрицы $A^n(k)$ становится известной полностью. В принципе, ее достаточно, чтобы написать решение исходной задачи (1)–(3). Но мы можем выписать всю матрицу, используя закономерности, присущие степеням матрицы $A^n(k)$, относящиеся к связям между строками матриц различных степеней. Однако

в этом нет необходимости, если сосредоточиться на поиске решения задачи. Первая строка матричного равенства (4) из раздела «Решение задачи» дает искомое решение. Предварительно распишем элементы первой строки (5) в явном виде наподобие (4). Вычислительные подробности опускаем. Поскольку каждый элемент занимает всю ширину страницы, элементы строки будем записывать столбцом.

$${}^n b_{11} = {}^n b_{11}(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \quad (5)$$

$${}^n b_{12} = {}^n b_{12}(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n+1}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_2 + \nu_3 + \dots + \nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \quad (6)$$

$${}^n b_{13} = {}^n b_{13}(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n+2}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_3 + \dots + \nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \quad (7)$$

$${}^n b_{1k} = {}^n b_{1k}(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n+k-1}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k}, \quad (8)$$

$$(k = 2, 3, \dots)$$

После освобождения от векторной записи решение задачи (1)–(2) (раздел «Решение задачи») может быть записано в скалярном виде:

$$y_{n+k} = {}^n b_{11} \cdot \bar{y}_{k-1} + {}^n b_{12} \cdot \bar{y}_{k-2} + \dots + {}^n b_{1k} \cdot \bar{y}_0 + \sum_{i=0}^{n-k} {}^n b_{11} \cdot \bar{f}_i. \quad (9)$$

С помощью формул (5)–(8) его можно записать явно с использованием только входных данных, т.е. коэффициентов уравнения $a_{k-1}, \dots, a_1, a_0$, правой части $\bar{f}_n$, ($n = 0, 1, 2, \dots$) и начальных условий $y_0 = \bar{y}_0, y_1 = \bar{y}_1, \dots, y_{k-1} = \bar{y}_{k-1}$. Поскольку решение – это итоговый результат, выпишем его здесь:

$$\begin{aligned} y_{n+k} = & \left(\sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n+1}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \right) \cdot \bar{y}_{k-1} + \\ & + \left(\sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n+2}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_2 + \nu_3 + \dots + \nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \right) \cdot \bar{y}_{k-2} + \\ & \dots \\ & + \left(\sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n+k}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot \frac{\nu_k}{\nu_1 + \nu_2 + \nu_3 + \dots + \nu_k} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \right) \cdot \bar{y}_0 + \\ & + \sum_{i=0}^{n-k} \left(\sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n-i}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \right) \cdot \bar{f}_i \quad (10) \end{aligned}$$

Подтвердим правильность формул (10), сравнив следствие из (10) с частным случаем для ($k = 2$), $a_1 = 3$, $a_0 = -2$.

Для этих параметров следствие имеет вид:

$$y_{n+2} = \left(\sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0 \\ \nu_1 + 2\nu_2 = n+1}} \frac{(\nu_1 + \nu_2)!}{\nu_1! \cdot \nu_2!} \cdot \frac{\nu_1 + \nu_2}{\nu_1 + \nu_2} \cdot a_1^{\nu_1} \cdot a_0^{\nu_2} \right) \cdot \bar{y}_1 + \\ + \left(\sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0 \\ \nu_1 + 2\nu_2 = n+2}} \frac{(\nu_1 + \nu_2)!}{\nu_1! \cdot \nu_2!} \cdot \frac{\nu_2}{\nu_1 + \nu_2} \cdot a_1^{\nu_1} \cdot a_0^{\nu_2} \right) \cdot \bar{y}_0$$

При $n = 3$ находим интересующее нас значение, например, y_5 :

$$y_5 = \left(\sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0 \\ \nu_1 + 2\nu_2 = 4}} \frac{(\nu_1 + \nu_2)!}{\nu_1! \cdot \nu_2!} \cdot 3^{\nu_1} \cdot (-2)^{\nu_2} \right) \cdot \bar{y}_1 + \left(\sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0 \\ \nu_1 + 2\nu_2 = 5}} \frac{(\nu_1 + \nu_2)!}{\nu_1! \cdot \nu_2!} \cdot \frac{\nu_2}{\nu_1 + \nu_2} \cdot 3^{\nu_1} \cdot (-2)^{\nu_2} \right) \cdot \bar{y}_0 = 31\bar{y}_1 + (-30)\bar{y}_0.$$

Результат подтверждает правильность формулы (10).

Формула (10) дает явный вид решения любого линейного разностного уравнения с любой неоднородностью и любыми начальными данными.

На вид формула (10) громоздка. Но в действительности она легко применима на практике. При небольших n и k вычисления легко проводятся вручную, для больших n и k , ввиду ее очевидной алгоритмичности, она легко программируется. Она применима при любых n и k без каких-либо принципиальных осложнений. В этом отношении она заметно выигрывает при сравнении ее с обычной процедурой нахождения решения разностного уравнения, связанной с решением характеристического уравнения. При больших порядках характеристического уравнения ($k \geq 5$) оно принципиально неразрешимо и приходится прибегать к приближенным методам его решения, что, как известно, довольно тягостно. Формула (10) избавляет нас от необходимости решать характеристическое уравнение. Дело сводится к подстановке данных задачи, выполнению сложения, умножения и сокращения (что проще деления) и последующего остаточного деления в минимальном объеме.

Структура Фибоначчи

Ценность формулы (10) из раздела «Степени матрицы» состоит не только в ее привлекательных вычислительных особенностях, но и в том, что по ходу ее получения обнаружилась важная роль чисел Фибоначчи [7, 8] при решении разностных уравнений. Выяснилось, что верхние угловые элементы

$$b_{11}^n = b_{11}^n(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \cdot a_{k-1}^{\nu_1} \cdot a_{k-2}^{\nu_2} \cdot \dots \cdot a_1^{\nu_{k-1}} \cdot a_0^{\nu_k} \quad (1)$$

степеней $A^n(k)$ матрицы $A(k)$ имеют самое прямое отношение к числам Фибоначчи. А именно: при $k = 2$ и $a_1 = a_0 = 1$ элемент

$$b_{11}^n = b_{11}^n(2) = \sum_{\substack{\nu_1 \geq 0, \nu_2 \geq 0 \\ \nu_1 + 2\nu_2 = n}} \frac{(\nu_1 + \nu_2)!}{\nu_1! \cdot \nu_2!} \quad (2)$$

в точности дает нам знакомые числа Фибоначчи. Действительно,

$$b_{11}^0(2) = 1, \quad b_{11}^1(2) = 1, \quad b_{11}^2(2) = 2, \quad b_{11}^3(2) = 3, \quad b_{11}^4(2) = 5, \dots \quad (3)$$

Но самое замечательное состоит в том, что в формуле (1) содержатся не только известные числа Фибоначчи, но и целое семейство последовательностей чисел, родственных числам Фибоначчи. Это родство непосредственно ясно из того факта, что все они содержатся в формуле (1). В еще более явной форме оно проявляется тогда, когда мы словесно сформулируем их характерное свойство. Характерное свойство чисел Фибоначчи (4) известно как *правило* порождения чисел этой последовательности. Это правило формулируется так: *каждое последующее число последовательности равно сумме двух предыдущих*. Это правило назовем *Правилом-2*. Появившаяся здесь двойка связана с тем, что последовательность (4) соответствует значению $k = 2$. И саму последовательность чисел Фибоначчи (4) будем называть последовательностью Фибоначчи-2 или, кратко, последовательностью Ф-2. Другим значениям k соответствуют последовательности чисел, которые будем обозначать символом Ф- k . Последовательность чисел Ф- k , как и последовательность чисел Ф-2, получается из материнской формулы (1) путем взятия чисел « a » равными единице: $a_{k-1} = a_{k-2} = \dots = a_1 = a_0 = 1$. Тогда получается последовательность чисел, описываемая формулой

$$b_{11}^n(k) = \sum_{\substack{\nu_1 \geq 0, \dots, \nu_k \geq 0 \\ \nu_1 + 2\nu_2 + 3\nu_3 + \dots + k\nu_k = n}} \frac{(\nu_1 + \nu_2 + \dots + \nu_k)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_k!} \quad (4)$$

и состоящая при каждом k из чисел

$$b_{11}^0(k), \quad b_{11}^1(k), \quad b_{11}^2(k), \quad b_{11}^3(k), \quad \dots \quad (5)$$

Это есть последовательность Ф- k .

Возьмем в (5) конкретные значения k .

Пусть $k = 3$. Тогда

$$b_{11}^0(3) = 1, \quad b_{11}^1(3) = 1, \quad b_{11}^2(3) = 2, \quad b_{11}^3(3) = 4, \quad b_{11}^4(3) = 7, \quad b_{11}^5(3) = 13, \quad \dots$$

Для этой последовательности работает порождающее *Правило-3*: *каждое последующее число равно сумме **трех** предыдущих чисел*.

Пусть $k = 4$. Тогда

$$b_{11}^0(4) = 1, \quad b_{11}^1(4) = 1, \quad b_{11}^2(4) = 2, \quad b_{11}^3(4) = 4, \quad b_{11}^4(4) = 8, \quad b_{11}^5(4) = 15, \quad \dots$$

Для этой последовательности работает порождающее *Правило-4*: *каждое последующее число равно сумме **четырёх** предыдущих чисел*.

И так далее. Эти вычисления можно продолжать сколь угодно долго. В результате для последовательности (5) будет сформулировано общее порождающее *Правило- k* : *каждое последующее число равно сумме **k** предыдущих чисел*.

Надо отметить, что формула (4) работает и при $k = 1$. В этом случае последовательность (5) принимает вид:

$$b_{11}^0(1) = 1, \quad b_{11}^1(1) = 1, \quad b_{11}^2(1) = 1, \quad b_{11}^3(1) = 1, \quad b_{11}^4(1) = 1, \quad b_{11}^5(1) = 1, \quad \dots$$

Если в общее порождающее правило подставить $k = 1$, то, хотя понятие суммы для одного слагаемого не определено, можно с натяжкой придать этому правилу смысл и при $k = 1$.

Таким образом, семейство последовательностей Ф- k , ($k = 1, 2, 3, \dots$) образует структуру, которую можно представить таблицей.

Любое число в этой таблице может быть вычислено по формуле (4) за счет выбора k и n .

Для согласования обозначений в таблице и в формуле (4) надо еще отметить, что

$$b_{11}^n(k) = y_n(k).$$

Числа Фидия

Рассмотрим вопрос о фигурирующих в таблице числах Фидия. Одно из них, число 1,618..., хорошо известно, с ним связано множество интересных фактов, и его даже называют «числом Бога». Все остальные числа, стоящие в столбце «число Фидия», оказываются в этом отношении обделенными, хотя с математической точки зрения все числа этого столбца равноправны.

Таблица

Структура Фибоначчи

Номер числа n	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	...	Число Фидия
Посл-ть Ф-1: $y_n(1)$		1	1	1	1	1	1	1	1	1	1	1	1	1	...	1
Посл-ть Ф-2: $y_n(2)$	0	1	1	2	3	5	8	13	21	34	55	89	144	233	...	1,618...
Посл-ть Ф-3: $y_n(3)$	0	1	1	2	4	7	13	24	44	81	149	274	504	927	...	1,839...
Посл-ть Ф-4: $y_n(4)$	0	1	1	2	4	8	15	29	56	108	208	401	773	1490	...	1,927...
Посл-ть Ф-5: $y_n(5)$	0	1	1	2	4	8	16	31	61	120	236	464	912	1793	...	1,966...
Посл-ть Ф-6: $y_n(6)$	0	1	1	2	4	8	16	32	63	125	248	492	976	1936	...	1,983...

Число Фидия для последовательности Фибоначчи- k определяется равенством

$$\text{Fid}-k = \lim_{n \rightarrow \infty} \frac{y_{n+1}(k)}{y_n(k)}, \quad (k = 1, 2, 3, \dots), \quad (1)$$

где $y_n(k)$ — число Фибоначчи с номером n в последовательности Фибоначчи- k .

Судя по экспериментальному материалу, содержащемуся в таблице, числа Фидия растут с ростом k , но не превосходят числа 2, медленно приближаясь к 2-ке. И это не случайно, т.к. справедлива **теорема**:

$$\forall k \ [k \geq 1] \ [\text{Fid}-k \leq 2]. \quad (2)$$

Доказательство вытекает из выкладок:

$$\begin{aligned} \frac{y_{n+1}(k)}{y_n(k)} &= \frac{y_n + [y_{n-1} + \dots + y_{n-k+1}]}{[y_{n-1} + \dots + y_{n-k+1}] + y_{n-k}} = \frac{\frac{y_n}{[\dots]} + 1}{1 + \frac{y_{n-k}}{[\dots]}} = \frac{\frac{[\dots] + y_{n-k}}{[\dots]} + 1}{1 + \frac{y_{n-k}}{[\dots]}} = \\ &= \frac{2 + \frac{y_{n-k}}{[\dots]}}{1 + \frac{y_{n-k}}{[\dots]}} = \frac{2 + \eta}{1 + \eta} = (2 + \eta) \cdot (1 - \eta + \eta^2 - \eta^3 + \dots) = 2 - \eta + \eta^2 - \dots = \\ &= 2 - \eta \cdot (1 - \eta + \eta^2 - \dots) = 2 - \frac{\eta}{1 + \eta} < 2, \quad \eta = \frac{y_{n-k}}{[\dots]}. \end{aligned}$$

Возбуждение, связанное с числом Фидия-2, равным 1,618... («золотое сечение»), которое царит вот уже почти два тысячелетия, затеняет остальные числа Фидия- k ($k = 1, 3, 4, \dots$), которые ничем не хуже. Однако обнаруживается, что над всеми ними возвышается двойка, роль которой, как выясняется, становится не менее загадочной, чем роль самих чисел Фидия.

Числа Фибоначчи- ∞ . Число Фидия- ∞ для чисел Фибоначчи- ∞

Снова обратимся к таблице. Каждая последовательность чисел Фибоначчи- k (в таблице изображены последовательности Фибоначчи для $k = 1, 2, 3, 4, 5, 6$) начинается с заштрихованного участка. Так, последовательность Фибоначчи-5 начинается с участка, состоящего из чисел 1, 1, 2, 4, 8. Все последующие числа последовательности Фибоначчи-5 порождаются этой пятеркой чисел по правилу: каждое последующее число равно сумме **пяти** предыдущих. При любом k последовательность Фибоначчи- k начинается с k начальных чисел. Интересен закон образования этих k начальных чисел. Закон образования k начальных чисел таков: k начальных чисел последовательности Фибоначчи- k являются первыми k числами предыдущей последовательности Фибоначчи- $(k - 1)$. Этот закон не произволен. Он индуцируется матрицей Фибоначчи $A(k)$,

$$A = A(k) = \begin{pmatrix} a_{k-1} & a_{k-2} & \dots & a_1 & a_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{k \times k},$$

порождаемой исходным разностным уравнением k -го порядка. В таблице по мере роста k выделенная жирным шрифтом область, имеющая вид треугольника, постепенно расширяется за счет увеличивающегося основания. Примечательно, что по вертикалям числа треугольника не меняются, имеет место

стабилизация, начинающаяся с самого начала. Так, если восьмерка (число 8) попадает в выделенную жирным шрифтом область, то она остается в своем столбце до самого конца, пока не провалится в бесконечность. В бесконечности основание выделенного жирным шрифтом треугольника по длине тоже становится бесконечным, и числа, в нем оказавшиеся, тоже образуют последовательность, которую мы обозначим $\Phi-\infty$ и которую будем называть последовательностью «Фибоначчи- ∞ ». Она состоит из чисел:

$$\Phi-\infty : 1, 1, 2, 4, 8, 16, 32, \dots, \quad (1)$$

правило образования которых является предельным выражением всех предыдущих правил образования чисел Фибоначчи. Назовем его Правилем- ∞ . *Правило- ∞ таково: каждое последующее число равно сумме всех предыдущих чисел.* Для формирования этой последовательности достаточно иметь только одно начальное число — единицу (можно любое количество чисел вида (1) в качестве начального множества, но достаточно только одной единицы). Разностное уравнение, определяющее эти числа, имеет вид:

$$\forall_{n \in N_0} [y_{n+1} = y_n + y_{n-1} + \dots + y_1 + y_0], \quad (2)$$

где $N_0 = \{0, 1, 2, \dots\}$ — множество натуральных чисел, начинающееся с 0.

Разностное уравнение, очевидно, соответствует Правилу- ∞ .

Это уравнение особого типа. Оно не имеет конечного порядка. В таком случае будем говорить, что оно имеет порядок, равный ∞ . Понятно, что в общем виде такого сорта уравнения записываются так:

$$\forall_{n \in N_0} [y_{n+1} = a_n y_n + a_{n-1} y_{n-1} + \dots + a_1 y_1 + a_0 y_0], \quad (3)$$

где $N_0 = \{0, 1, 2, \dots\}$ — множество натуральных чисел, начинающееся с 0.

Как решаются такие уравнения, пока неясно. В отличие от этого, уравнение (2) имеет очевидное **решение**:

$$y_n = 2^{n-1}, \quad (n = 1, 2, 3, \dots); \quad y_0 = 1. \quad (4)$$

Число Фидия для этой последовательности, как и полагается, равно 2:

$$\text{Fid}-\infty = \lim_{n \rightarrow \infty} \frac{y_{n+1}(\infty)}{y_n(\infty)} = \lim_{n \rightarrow \infty} \frac{y_{n+1}}{y_n} = \lim_{n \rightarrow \infty} \frac{2^n}{2^{n-1}} = 2. \quad (5)$$

Это решение лишний раз показывает родство чисел Фибоначчи с биномом Ньютона.

ЛИТЕРАТУРА

1. Гельфонд А. О. *Исчисление конечных разностей*. М.: Наука; 1967.
2. Gelfond A. O. *Differenzenrechnung*. Berlin: Deutsch. Verl. Wiss.; 1958.
3. Самарский А. А. *Введение в теорию разностных схем*. М.: Наука; 1971.
4. Кострикин А. И. *Введение в алгебру*. М.: Наука; 1994.
5. Воеводин В. В. *Численные методы алгебры*. М.: Наука; 1976.
6. Успенский В. А. *Треугольник Паскаля*. М.: Наука; 1979.
7. Воробьев Н. Н. *Числа Фибоначчи*. М.: Наука; 1978.
8. Hackbusch W. *Tensor Spaces and Numerical Tensor Calculus*. Berlin: Springer; 2012. DOI: 10.1007/978-3-642-28027-6.

ПРОБЛЕМЫ И ОБЗОР АЛГОРИТМИЧЕСКИХ РЕШЕНИЙ В ТЕХНОЛОГИИ FFF 3D-ПЕЧАТИ

Э. А. Юнусов^{1,2,a}, Д. В. Горбунов^{2,3,b}

¹ ПАО «Сургутнефтегаз», ПУ СургутАСУнефть, Сургут, Российская Федерация

² Сургутский государственный университет, г. Сургут, Российская Федерация

³ Сургутский филиал федерального государственного автономного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Сургут, Российская Федерация

^a ✉ yunusov_ea@mail.ru, ^b Gorbunov_dv@surgu.ru

Аннотация: в статье рассматриваются существующие алгоритмы подготовки модели к 3D-печати. Проведен обзор основных программ для подготовки 3D-моделей к печати, а также рассмотрена возможность печати экзотическими материалами, такими как шоколад. На примере одной из программ изучены и разделены на категории алгоритмы подготовки модели к печати. Также представлен обзор общепринятых практик и рассмотрены перспективные разработки и решения на различных стадиях. В статье предложены и описаны варианты тестирования напечатанных моделей для оценки работоспособности перспективных алгоритмов. Кроме того, была проведена оценка скорости разработки новых алгоритмов. Приведены факторы, которые существенно ограничивают развитие технологии 3D-печати, например, такие, как коммерциализация и закрытая разработка, при которой результаты исследований не выкладываются в публичное пространство. На основе проведенного обзора проблем и алгоритмов в аддитивных технологиях можно сказать, что развитие алгоритмов подготовки моделей к 3D-печати может существенно улучшить характеристики готовых изделий без существенного повышения затрат на печать.

Ключевые слова: 3D-печать, алгоритмы печати, алгоритм заполнения, слайсер.

Для цитирования: Юнусов Э. А., Горбунов Д. В. Проблемы и обзор алгоритмических решений в технологии FFF 3D-печати. *Успехи кибернетики*. 2025;6(2):67–74.

Поступила в редакцию: 29.05.2025.

В окончательном варианте: 14.06.2025.

FFF 3D PRINTING TECHNOLOGY ALGORITHMS: PROBLEMS AND REVIEW

E. A. Yunusov^{1,2,a}, D. V. Gorbunov^{2,3,b}

¹ PJSC Surgutneftegaz, PM SurgutASUneft, Surgut, Russian Federation

² Surgut State University, Surgut, Russian Federation

³ Surgut Branch of Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, Surgut, Russian Federation

^a ✉ yunusov_ea@mail.ru, ^b gorbunov_dv@surgu.ru

Abstract: we examined existing algorithms for preparing models for 3D printing and provided an overview of the key software tools used in this process. We also explored the potential of printing with unconventional materials, such as chocolate. Using one of the programs as a case study, we analysed model preparation algorithms and categorised them based on their functions. In addition, we reviewed common industry practices and discussed promising developments and solutions at various stages of the workflow. We proposed and described methods for testing printed models to evaluate the performance of new algorithms. Furthermore, we assessed the pace of algorithm development and identified key barriers to progress in 3D printing, including commercial restrictions and the lack of publicly available research due to closed-source development. Based on our review, we conclude that improving algorithms for model preparation can enhance the quality of printed products without significantly increasing production costs.

Keywords: 3D printing, printing algorithms, filling algorithm, slicer.

Cite this article: Yunusov E. A., Gorbunov D. V. FFF 3D Printing Technology Algorithms: Problems and Review. *Russian Journal of Cybernetics*. 2025;6(2):67–74.

Original article submitted: 29.05.2025.

Revision submitted: 14.06.2025.

Введение

Современная технология аддитивной 3D-печати активно развивается и применяется на практике за счет аддитивности (вместо обычного процесса убирания лишнего материала здесь происходит добавление материала) и высокой скорости разработки прототипов. Данная технология позволила множеству людей использовать огромное количество разнообразных моделей (в том числе доступных в сети Интернет для некоммерческого использования), но в подавляющем большинстве случаев 3D-печать остается на уровне хобби. Несмотря на это сейчас некоторые производители начали выкладывать модели деталей для своей продукции в открытом доступе (например, Philips для триммера предоставляют в открытом доступе модели насадок), что повышает ремонтпригодность и способствует привлечению новых клиентов.

В рамках статьи рассматриваются проблемы аддитивных технологий и возможные их решения, а также существующие алгоритмы подготовки модели к 3D-печати, возможные пути совершенствования или новые алгоритмы, которые могут улучшить данную технологию. Новые алгоритмы печати способны повысить прочность готовых изделий и при этом снизить время, требуемое для печати.

Материалы и методы

В рамках статьи проведен обзор различного ПО (далее — слайсер) для подготовки 3D-модели в специальный код (G-Code) для 3D-принтера и представлены особенности слайсеров. Также рассмотрены существующие алгоритмы, встроенные в ПО. Как и большинство систем, слайсеры имеют и проблемы. Для решения этих проблем существует множество неоднозначных решений. Стоит отметить, что есть алгоритмы печати, которые еще не внедрены в слайсеры, но эти решения активно используются. Следует уделить внимание способам и алгоритмам тестирования готовых изделий, которые могут адекватно подтвердить корректность работы и зафиксировать улучшенные характеристики деталей, полученных на основе новых алгоритмов подготовки модели к 3D-печати.

На данный момент существует большое количество ПО, которое способно подготовить 3D-модели и преобразовать их в G-code для 3D-принтера [1]. В статье рассмотрены наиболее популярные слайсеры:

- Cura — открытый и свободный слайсер [2]. Создан компанией UltiMaker;
- PrusaSlicer — открытый, многофункциональный и часто обновляемый слайсер [3]. Был создан компанией Prusa Research;
- OrcaSlicer — открытый слайсер на базе Bambu Studio, созданный компанией SoftFever [4].

Все представленные слайсеры объединяет открытость исходного кода, популярность в сообществе 3D-печати и частичное использование одного и того же исходного кода (OrcaSlicer основан на Bambu Studio, который, в свою очередь, основан на исходном коде PrusaSlicer, который написан на базе исходного кода оригинального слайсера Slic3r). Каждый из этих слайсеров может быть использован для подготовки модели к любому принтеру (в том числе спроектированному самостоятельно) за счет возможности гибкой настройки конфигурации принтеров, даже для печати шоколадом [5], например, Cocoa Press (рис. 1).

Каждый из слайсеров имеет свои уникальные алгоритмы. Следует отметить, что эти алгоритмы со временем могут быть внедрены и в другие слайсеры или же на их основе разрабатываются более совершенные алгоритмы подготовки 3D-моделей к печати.

Большинство алгоритмов будет рассмотрено на базе слайсера OrcaSlicer (большая часть из которых реализована в других). В данном слайсере модель разделяется на отдельные слои и обрабатывается шаг за шагом (рис. 2). Далее происходит обработка каждого слоя алгоритмами генерации периметров, заполнения, мостов, крышек, дна.

В первую очередь, следует рассмотреть алгоритмы генерации периметров модели. Данные алгоритмы формируют внешнюю оболочку модели и определяют внешний вид и прочностные характеристики готового изделия. На текущий момент в OrcaSlicer периметры генерируются только стандартным вариантом наложения слоев друг на друга (рис. 3, модель слева), что приводит к образованию больших промежутков между полосами материала. Пустоты в готовом изделии приводят к уменьшению адгезии слоев, что, в свою очередь, уменьшает прочностные характеристики модели.

Для того чтобы минимизировать пустоты в готовом изделии, а также для повышения адгезионных характеристик слоев при печати разработаны более совершенные алгоритмы подготовки модели к печати. Следует отметить, что эти алгоритмы пока не внедрены в слайсеры. Ниже представлены два



Рис. 1. Пример печати шоколадом

алгоритма:

- использование так называемой «кирпичной кладки» (рис. 3, модель справа) [6], что подразумевает сдвиг по высоте у соседних слоев периметров. Таким образом достигается минимизация пустот между слоями;
- другой вариант улучшения заполнения — создать пустоту между двумя периметрами и заполнить их сверху жидким пластиком, выполнив, по своей сути, операцию литья под давлением, только без самого давления [7].

Для тестирования готовых изделий, полученных на основе представленных алгоритмов, можно

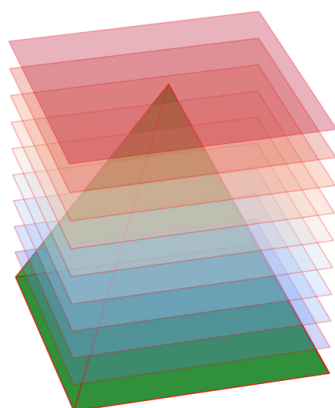


Рис. 2. Пример нарезки пирамиды на слои шаг за шагом для последующей обработки

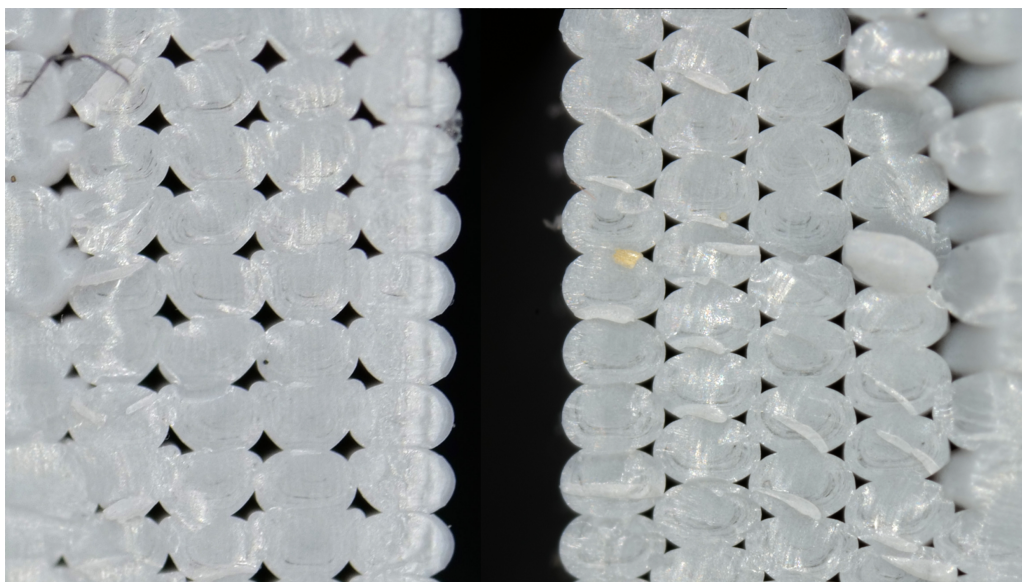


Рис. 3. Стандартные периметры и «кирпичные» периметры

применить классическую схему исследования с помощью метода трехточечного изгиба [8]. Пример работы машины представлен на рис. 4. При улучшении адгезии слоев и уменьшении пустот между слоями характеристики модели при изгибе должны увеличиться.

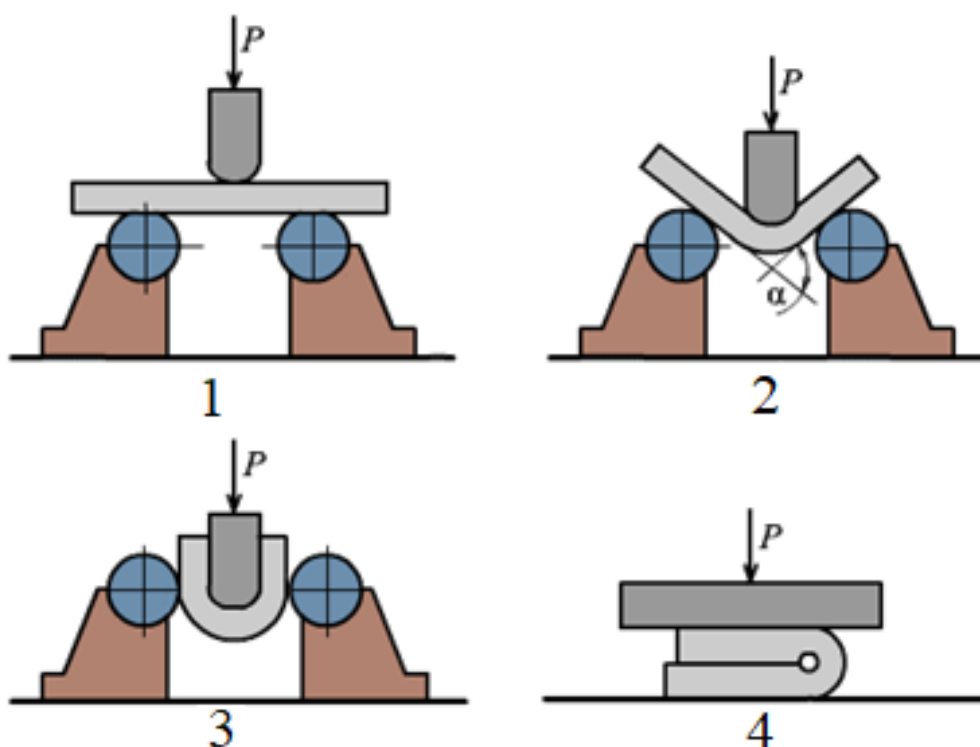


Рис. 4. Пример работы машины трехточечного изгиба

Следующий блок алгоритмов — генерация внутреннего заполнения. Эти алгоритмы необходимы для внутреннего заполнения модели (объем заполнения обычно указывается в %). Наиболее распространенные алгоритмы представлены на рис. 5.

Каждый из вариантов заполнения имеет свои достоинства и недостатки. В этой связи оператор 3D-принтера должен сам определить тип и плотность заполнения исходя из требований и предназначения детали. Следует отметить, что заполнение влияет на прочностные характеристики готового изделия, но эта зависимость между прочностью и заполнением нелинейная.

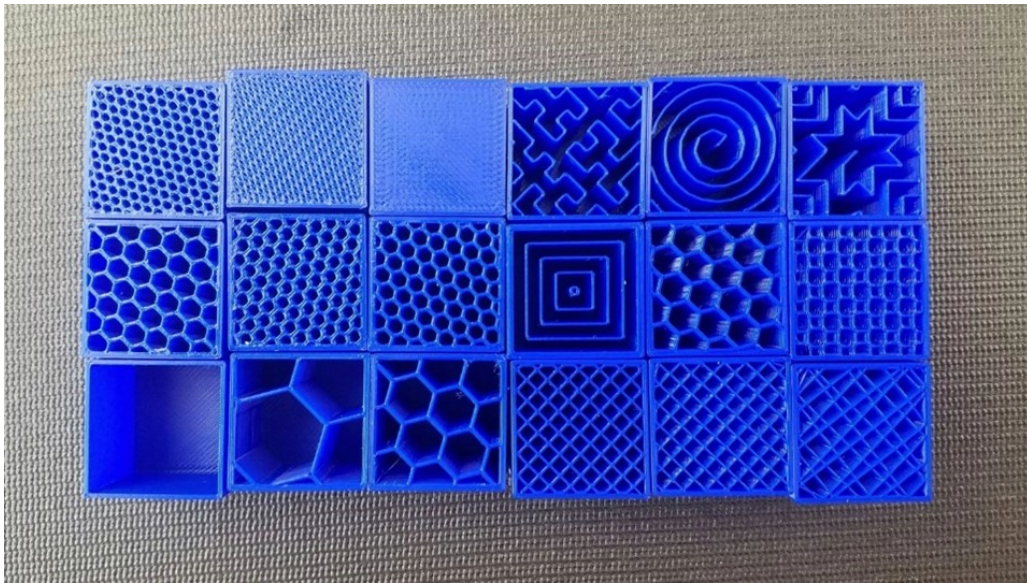


Рис. 5. Примеры различных вариантов заполнения

На данный момент существует несколько вариантов совершенствования алгоритмов заполнения, которые можно реализовать в слайсерах для повышения прочности готового изделия:

- применение нескольких вариантов заполнения для одной модели. В рамках исследования [9] были использованы сочетания паттернов с поворотом на 180 градусов, что привело к увеличению прочности при растяжении и сжатии;
- применение мультиматериальных заполнений и структур, которые сочетают различные характеристики. Примером такого сочетания является полиэтилентерефталатгликоль (PETG) и термопластичный полиуретан (TPU). В такой комбинации можно создать детали с достаточной прочностью и возможностью зафиксировать TPU, который слишком мягок для прочного крепления;
- создание нового алгоритма заполнения модели, который позволил бы заполнить печатную модель различными видами утяжелителей без необходимости на этапе моделирования создавать полости. При разработке такого алгоритма необходимо учесть, что заполнение утяжелителем (например, свинцовыми шариками) не должно приводить к появлению эффекта погремушки. На данный момент для того, чтобы увеличить массу готового изделия, часто используют остановку печати на указанном слое и загрузку утяжелителей, но этот вариант требует постоянного контроля печати, а также в большинстве случаев определенного размера утяжелителя для конкретной модели.

Для тестирования разных типов заполнения можно использовать трехточечный изгиб, а также специализированные виды тестов (по типу проверки работоспособности заполнения и утяжеления).

Далее необходимо рассмотреть алгоритмы, которые можно изменить, — блок алгоритмов генерации поддерживающего материала нависающих частей модели. Из-за специфики процесса 3D-печати сложные криволинейные формы не всегда возможно напечатать без дополнительных структур, в таком случае необходимо использовать дополнительный поддерживающий материал, поверх которого будет происходить печать основной фигуры (рис. 6).

На данный момент известны следующие пути, которые можно реализовать, чтобы повысить качество готового изделия:

- разработка новых алгоритмов генерации поддержек для специфичных задач. Таким примером может служить алгоритм генерации поддержек для различных плоских корпусов. Сейчас в большинстве случаев модели располагаются диагонально над столом печати. В таком случае достаточно небольшой опоры и нескольких единичных точек соприкосновения, чтобы модель могла быть отпечатана, при этом минимизируются дефекты на модели и сокращается расходуемый материал;
- применение нескольких алгоритмов генерации поддержек для одной модели, которые используются точно для разных моделей. Так как одни виды поддержек дают поверхность с меньшим количеством дефектов, а другие расходуют в целом гораздо меньше материала, можно использовать одновременно оба алгоритма на разных частях модели, чтобы получить более оптимальный расход

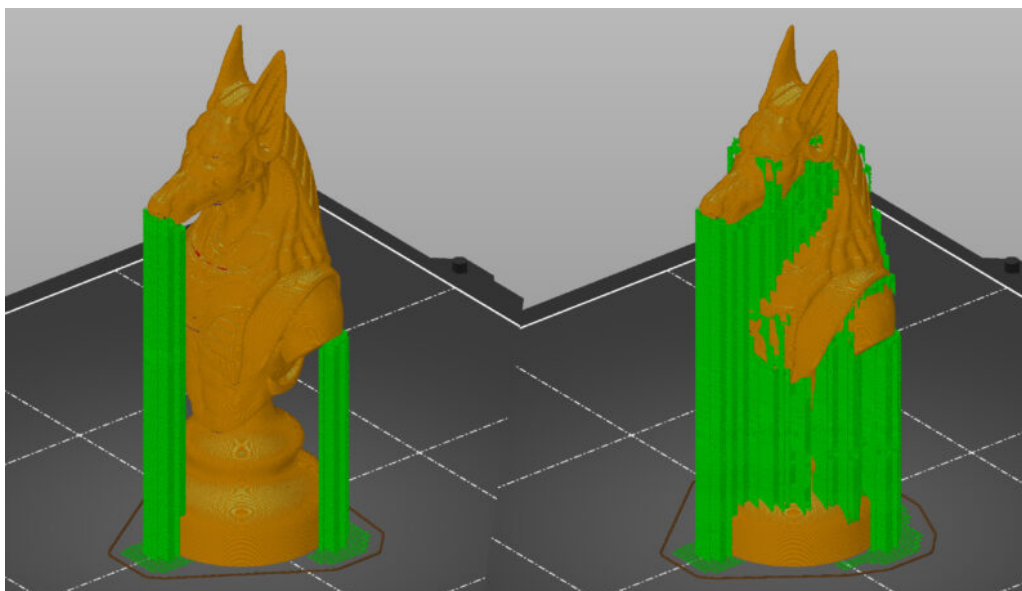


Рис. 6. Пример двух видов поддержек на одной и той же модели

материала и уменьшение дефектов внешней части модели, которая соприкасается с поддержкой.

Соответственно, тестирование обновленных алгоритмов поддержки будет представлять из себя печать одной и той же модели, только с различными вариациями поддержки, а оценка должна производиться по следующим параметрам:

- расход материала — чем меньше было израсходовано материала, тем лучше, но должно сохраняться такое же или меньшее количество дефектов на модели. Следует учесть, что печатать модель можно без поддержки, но результат такой печати будет в виде испорченного готового изделия;
- количество дефектов на модели: чем меньше поддержек соприкасается с моделью, тем меньше будет дефектов от поддержек на готовом изделии, что может существенно уменьшать расход на постобработку детали.

Следующий и последний блок алгоритмов — это блок дополнительных алгоритмов компенсации различных дефектов или особенностей 3D-печати. Таких алгоритмов достаточно много, так как есть большое количество типовых дефектов при 3D-печати, которые можно предусмотреть на этапе подготовки модели к печати. Примеры таких алгоритмов:

- алгоритм PA\LA — алгоритм компенсации подачи филамента [10]. Проблема возникает из-за особенностей филамента и несовершенства механики экструдера (механизма подачи филамента), например, резиноподобный TPU сильно растягивается и сжимается при печати, что приводит к резкому изменению ускорения при печати. Работа алгоритма PA\LA основана на изменении скорости подачи филамента соответственно необходимой экструзии (если ее нужно снизить, подача будет заранее уменьшена для того, чтобы избежать появления дефектов, и наоборот);

- алгоритмы специальной нарезки модели на слои, которые могут напечатать деталь единой нитью по спирали, что актуально для различных тонкостенных деталей и позволяет печатать более крепкую деталь относительно стандартных настроек;

- алгоритм компенсации эффекта «слоновой ноги» — алгоритм позволяет исправить дефект, связанный с неправильной настройкой принтера. Проявление такого дефекта выражается в утолщении нижних слоев печатающейся детали. Такой дефект в подавляющем большинстве случаев появляется вследствие неправильной настройки принтера (например, неправильной калибровки осей). На практике часто проще решить такую проблему программным способом без изменения настроек принтера;

- алгоритмы сброса материала. В случае многоцветной печати на основе систем, которые подают филамент через один и тот же хотенд, этот алгоритм позволяет вместо расхода времени и материала на очистку старого филамента в специально выделенной зоне эту же очистку проводить в заполнении поддержек и/или заполнении основной модели, ускоряя процесс печати и при этом уменьшая расход материала;

- алгоритмы уменьшения лишних движений. Вследствие неоптимальной нарезки может быть

огромное количество лишних движений экструдера, что увеличивает количество времени на печать, а при неисправностях увеличивается и количество дефектов на модели. Данный алгоритм переопределяет все холостые передвижения так, чтобы добиться наименьшего количества времени, затрачиваемого на них.

Соответственно, для каждого алгоритма из вышеперечисленных необходимы свои варианты тестирования, так как все они уникальные и изменяют один из критериев печати. Комбинация из всех вышепредставленных способов и алгоритмов может существенно осложнить процесс подготовки модели, но это приведет к повышению качества печати и уменьшению времени печати и используемого материала.

Результаты и их обсуждение

Подводя итог, можно сказать, что современные алгоритмы 3D-печати активно развиваются и изменяются, хотя темпы развития могли бы быть и быстрее. На данный момент очень активно развиваются следующие алгоритмы, которые впоследствии могут стать алгоритмами по умолчанию в программах-слайсерах:

- новые варианты алгоритмов генерации стенок, такие как кирпичные стенки и их модификации, позволяющие увеличивать прочность готовых изделий за счет повышения адгезии слоев;
- новые варианты и комбинирование существующих алгоритмов генерации заполнения моделей, которые также могут увеличить прочность только за счет использования комбинированных видов заполнения в одной модели;
- новые варианты и комбинирование алгоритмов генерации поддержек, уменьшающих расходы материалов и количество дефектов модели.

К сожалению, данная область 3D-печати (а именно новые алгоритмы) развивается не так быстро, как могла бы, из-за следующих факторов:

- проблемы с распространением нелегальных копий моделей [11]. В рамках исследования было отмечено, что на данный момент в сфере 3D-моделирования и печати наблюдается дефицит доступа к 3D-моделям. На существующем рынке создание моделей для последующей продажи является убыточным, так как любой покупатель получает для себя полную копию файлов для печати без каких-либо серьезных средств защиты, что способствует созданию целых сообществ людей, которые, единожды купив одну модель, выкладывают ее в свободный доступ, из-за чего официальную копию будут покупать реже. Для решения данных проблем возможно внедрение новых технологий на базе блокчейн, рассмотренных в статье [11].
- проблемы со свободным распространением алгоритмов. Множество технологий разрабатывается внутри коммерческих компаний, что замедляет развитие публичных технологий;
- проблема отсутствия единых стандартов тестов. Существует большое количество калибровочных моделей, которые на один и тот же параметр могут давать разные результаты, противоречащие друг другу.

Заключение

Проведенное исследование показало, что разработка и использование новых алгоритмов позволят добиться лучших результатов в 3D-печати без существенного повышения затрат на печать моделей, а также без необходимости специального обучения, только за счет программных изменений, хотя это и требует достаточно большого труда проектировщиков и программистов. Однако даже это развитие является достаточно медленным, так как значительное количество технологий и алгоритмов разрабатывается закрытыми коммерческими компаниями. Развитие алгоритмов печати способно упростить сложность подготовки моделей, что приведет к более массовому распространению 3D-принтеров в быту, что, в свою очередь, приведет к увеличению скорости развития аддитивных технологий.

ЛИТЕРАТУРА

1. Лисяк В. В. Обзор слайсеров для 3D-печати. *Известия ЮФУ. Технические науки*. 2022;3:60-74.
2. Смирнов М. А., Рыбкин Н. О., Ксенофонтова О. Л. FDM-технология: особенности применения, преимущества, недостатки. *Проблемы экономики, финансов и управления производством: сборник научных трудов вузов России*. 2021;48:115-122.

3. Kohout K. *PrusaSlicer 2.9.1: Smarter Sequential Printing and Stronger Multi-Material Interlocking*. Режим доступа: https://blog.prusa3d.com/prusaslicer-2-9-1-smarter-sequential-printing-stronger-multi-material-interlocking_111458.
4. Benedict O'Neill. *Orca Slicer vs. Cura: A Technical and Practical Comparison*. Режим доступа: <https://www.wevolver.com/article/orca-slicer-vs-cura-a-technical-and-practical-comparison>.
5. Athira V. A., Udayarajan C. T., Goksen G., Brennan C. S., Nisha P. A Brief Review on 3D Printing of Chocolate. *International Journal of Food Science and Technology*. 2023;58(6):2811-2828. DOI: 10.1111/ijfs.16415.2023.
6. Suiker A. S. J. Mechanical Performance of Wall Structures in 3D Printing Processes: Theory, Design Tools and Experiments. *International Journal of Mechanical Sciences*. 2018;137:145-170. DOI: <https://doi.org/10.1016/j.ijmecsci.2018.01.010>.
7. David O. K. Austin Colon Injection Printing: Additive Molding Via Shell Material Extrusion and Filling. *Additive Manufacturing*. 2020;36:101469. DOI: <https://doi.org/10.1016/j.addma.2020.101469>.
8. Балашов А. В., Черданцев А. О., Новиковский Е. А., Ананьин С. В., Белоплотов С. В. Исследование прочности изделий, полученных методом 3D-печати. *Ползуновский вестник*. 2016;2:61-64.
9. Aboelella M. G., Ebeid S. J., Sayed M. M. Layer Combination of Similar Infill Patterns on the Tensile and Compression Behavior of 3D Printed PLA. *Scientific Reports*. 2025;15:11759. DOI: <https://doi.org/10.1038/s41598-025-94446-8>.
10. Tronvoll S. A., Popp S., Elverum C. W., Welo T. Investigating Pressure Advance Algorithms for Filament-Based Melt Extrusion Additive Manufacturing: Theory, Practice and Simulations. *Rapid Prototyping Journal*. 2019;25(5):830-839. DOI: <https://doi.org/10.1108/RPJ-10-2018-0275>.
11. Гавриленко Т. В., Горбунов Д. В., Востров В. А. Состояние и основные проблемы развития аддитивных технологий. *Успехи кибернетики*. 2024;5(2):75-83. DOI: 10.51790/2712-9942-2024-5-2-08.

ПРОБЛЕМА УСТОЙЧИВОСТИ КРИТИЧЕСКОЙ ИНФРАСТРУКТУРЫ В УСЛОВИЯХ ВЫСОКОГО УРОВНЯ АВТОМАТИЗАЦИИ

Э. М. Исмаилов^{1,a}, Т. В. Гавриленко^{1,2,b}

¹ Сургутский филиал федерального государственного автономного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Сургут, Российская Федерация

² Сургутский государственный университет, г. Сургут, Российская Федерация

^a ORCID: <https://orcid.org/0009-0003-9913-3283>, trol.8@inbox.ru

^b ORCID: <https://orcid.org/0000-0002-3243-2751>, taras.gavrilenko@gmail.com

Аннотация: вследствие быстрого скачкообразного развития научно-технического прогресса и меняющейся геополитики возникает острая необходимость изучения и активного внедрения новых методов управления и защиты энергетической инфраструктуры, так как используемые технические решения в данной области не всегда могут обеспечить необходимый уровень надежности и безопасности. Применяемые для этого системы и программы в разных секторах экономики, являющиеся сегодня передовыми, могут мгновенно потерять актуальность из-за появления методов их взлома и невозможности их своевременного усовершенствования. В статье рассматриваются технологии на основе искусственного интеллекта, внедрение которых может обеспечить комплексный подход к вопросам управления и защиты энергетического сектора страны. Представлен анализ технологий на базе искусственного интеллекта, внедрение которых может обеспечить защиту от внешних воздействий на критическую инфраструктуру страны.

Ключевые слова: критическая информационная инфраструктура, энергетика, контроль, искусственный интеллект.

Для цитирования: Исмаилов Э. М., Гавриленко Т. В. Проблема устойчивости критической инфраструктуры в условиях высокого уровня автоматизации. *Успехи кибернетики*. 2025;6(2):75–79.

Поступила в редакцию: 16.02.2025.

В окончательном варианте: 17.05.2025.

SUSTAINABILITY OF HIGHLY AUTOMATED CRITICAL INFRASTRUCTURE

E. M. Ismailov^{1,a}, T. V. Gavrilenko^{1,2,b}

¹ Surgut Branch of Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, Surgut, Russian Federation

² Surgut State University, Surgut, Russian Federation

^a ORCID: <https://orcid.org/0009-0003-9913-3283>, trol.8@inbox.ru

^b ORCID: <https://orcid.org/0000-0002-3243-2751>, taras.gavrilenko@gmail.com

Abstract: we examined the need for new approaches to managing and protecting energy infrastructure in the context of rapid scientific and technological advancement and evolving geopolitical conditions. The technical solutions currently employed in this sector fail to ensure the required level of reliability and security. New systems and programs introduced across various economic sectors quickly become obsolete due to the continuous emergence of novel methods for bypassing protective mechanisms and the inherent difficulty in promptly detecting such threats.

Keywords: critical IT infrastructure, power industry, monitoring, artificial intelligence.

Cite this article: Ismailov E. M., Gavrilenko T. V. Sustainability of Highly Automated Critical Infrastructure. *Russian Journal of Cybernetics*. 2025;6(2):75–79.

Original article submitted: 16.02.2025.

Revision submitted: 17.05.2025.

Введение

С 2014 года на Российскую Федерацию активно накладываются разного рода санкции, которые в том числе ограничивают доступ к технологиям, кроме этого, в последнее время стремительно растет количество кибератак из-за рубежа, а также атак телефонных мошенников с применением профессиональных психологических подходов, которые могут нанести серьезный ущерб не только личной безопасности граждан, но и предприятиям критической инфраструктуры, сотрудники которых также подвергаются указанному воздействию. Однако при должном внимании к этим проблемам усилия,

направленные на их решение, могут не только обезопасить важные для страны объекты, но и дать мощный толчок развитию собственной научной и технической базы.

Ввиду сложившихся обстоятельств многие зарубежные предприятия, которые обеспечивали техническое оснащение и программное обеспечение для защиты объектов критической инфраструктуры, заявили о закрытии своих дочерних компаний и прекращении любого сотрудничества с Россией. Как следствие, информационные системы, информационно-телекоммуникационные сети, автоматизированные системы управления, работающие практически во всех секторах экономики (энергетика, оборонная промышленность, банковская сфера, транспорт, связь и т. д.), остались без программной и технической поддержки со стороны зарубежных производителей, чье оборудование и программное обеспечение используется в том числе и на объектах критической информационной инфраструктуры страны (КИИ), а это значит, что все отрасли КИИ стали более уязвимы перед несанкционированным вмешательством в работу их систем.

В итоге промышленный шпионаж приобретает более циничный характер, направленный не только на кражу финансовой и интеллектуальной собственности, но и на разные виды диверсий с целью нанесения максимально возможного ущерба механизмам и технологическим процессам.

Указанные обстоятельства влекут за собой технические, политические, военные, юридические, коммерческие, производственные и иные виды рисков, реализация которых может привести к тяжелейшим последствиям для жизни и здоровья граждан и экономики страны [1, 2].

Общеизвестным примером внешнего вмешательства в объекты КИИ и последствий таких воздействий являются события в Венесуэле и Иране.

Венесуэла

В 2019 году Венесуэла подверглась кибератаке, в ходе которой вышла из строя самая большая в стране гидроэлектростанция.

По словам министра информации страны Хорхе Родригеса, «кибератака стала причиной самого масштабного в истории Венесуэлы отключения электроэнергии». Атака была направлена на автоматическую систему контроля ГЭС «Гури», которая контролировала процесс выработки электроэнергии. В ходе данной кибератаки пропало электроснабжение таких важных секторов КИИ, как больницы, банки и т.д. [3].

Иран

В 2010 году в Иране в ходе кибератаки с применением компьютерного червя Stux.net, причинен вред объектам ядерной промышленности – заводам по обогащению урана и АЭС, что нанесло огромный урон ядерной программе Ирана [4], развитие которой было отброшено на многие годы назад.

Основной задачей вредоносного червя было испортить все обратные связи между системами управления на базе программно-логического комплекса и системами мониторинга текущего состояния установок по обогащению урана.

При этом необходимо отметить, что в приведенных выше примерах были атакованы системы управления на базе программно-логического комплекса от ведущих производителей и их системы защиты продемонстрировали свою неэффективность перед внешним вмешательством.

Очевидно, что существует острая необходимость защитить объекты КИИ, это важно для обеспечения безопасности государства и стабильности ключевых сфер жизнедеятельности.

Состояние проблемы

Перспективным направлением для решения данной задачи представляется применение технологии искусственного интеллекта (ИИ) и ее активное внедрение в сектор КИИ.

Федеральный закон Российской Федерации от 24.04.2020 № 123-ФЗ дает определение искусственному интеллекту: комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получить при выполнении конкретных задач результаты, сопоставимые как минимум с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру (в том числе информационные системы, информационно-телекоммуникационные сети, иные технологические средства обработки информации), программное обеспечение (в том числе то, в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений [5].

В рамках управления системами безопасности на сегодняшний день технологии ИИ применяются, например, в системах видеонаблюдения и контроля безопасности путем сбора данных с камер видеонаблюдения, охранных систем и их передачи в адрес оператора [5].

Перспективной сферой применения ИИ являются предприятия среднего и малого бизнеса. Облачные системы ML Space от Sber Cloud, Azure от Microsoft и Auto ML от Google предоставляют все необходимые инструменты для создания собственных алгоритмов с целью автоматического ответа на запрос или прогнозирования спроса [6]. Такие компании могут стать активными участниками создания и внедрения искусственного интеллекта в секторе КИИ.

Сегодня для управления любым сектором экономики используют комплекс аппаратных и программных средств. При разделении любого крупного процесса управления на подпроцессы мы увидим, что существует точечное управление (локальное по месту) и обширное (общее управление из операторной). Процессы такого масштаба необходимо рассматривать как динамические системы с изменяющимися во времени параметрами. На рисунке 1 (модель черного ящика) представлена упрощенная система управляемого объекта (S), например насоса по перекачке нефтепродукта, с входящими ($U(t)$, $V(t)$) и выходящим ($Y(t)$) векторными параметрами.

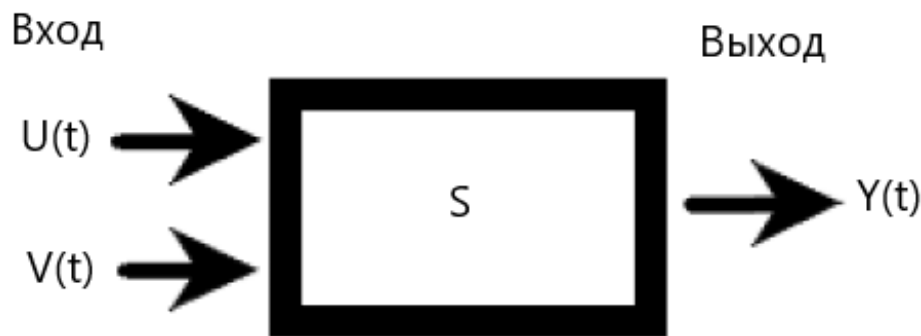


Рис. 1. Модель черного ящика

Рассматривая векторный выход системных параметров $Y(t)$, мы понимаем, что это может быть реакция системы на управление в нормальном режиме работы по входящему каналу $U(t)$ или неуправляемому каналу $V(t)$ при вмешательстве с целью нанесения вреда системе, в которой объект (S) служит исполнительным органом. При этом существует вероятность множественных возможных значений (X) и элементов упорядоченного множества (T), моментов времени (t). Таким образом, вероятность событий $T \rightarrow X$, $X(t) \in X^T$, $t \in T$.

На сегодняшний день велика вероятность того, что при внешнем воздействии или активации вируса в системе управления и мониторинга параметров персонал при удаленном нахождении от объекта управления не сможет распознать (услышать, увидеть), что объект больше не поддается регулированию и, по сути, является неуправляемой системой, что может привести к негативным последствиям. Существует нерешенная задача по созданию саморазвивающейся модели, которая сможет в режиме реального времени проводить самообучение, тем самым являясь гарантом оперативного пресечения негативных воздействий и при необходимости — возвращения системы в штатный режим работы.

Проведя анализ часто применяемого электрооборудования, систем управления и передачи данных, можно выдвинуть предположение, что ИИ был бы востребован как дополнительная (автономная) система мониторинга состояний электрооборудования, систем контроля и передачи данных [7–10], а также для сравнения значений сигналов, которые приходят от электрооборудования по сетям контроля и передачи данных в контроллер и от него — на верхний уровень, с сигналами, которые получает ИИ от тех же датчиков, что и система управления [11]. При возникновении разницы между входящим и уходящим сигналом контроллера ИИ мог бы мгновенно формировать реестр ошибочных сигналов и передавать его оператору, тем самым давая возможность оперативному персоналу объекта КИИ предпринять необходимые действия для проверки и устранения неисправностей. На рисунке 2 представлена разработанная нами схема высокой автоматизации, в которую встроен ИИ, условные обозначения и изображения рисунка 2 представлены на рисунке 3.

Применение ИИ может позволить сформировать независимый контур сбора и анализа параметров состояния систем объекта КИИ, в том числе на базе оценки больших данных о косвенных

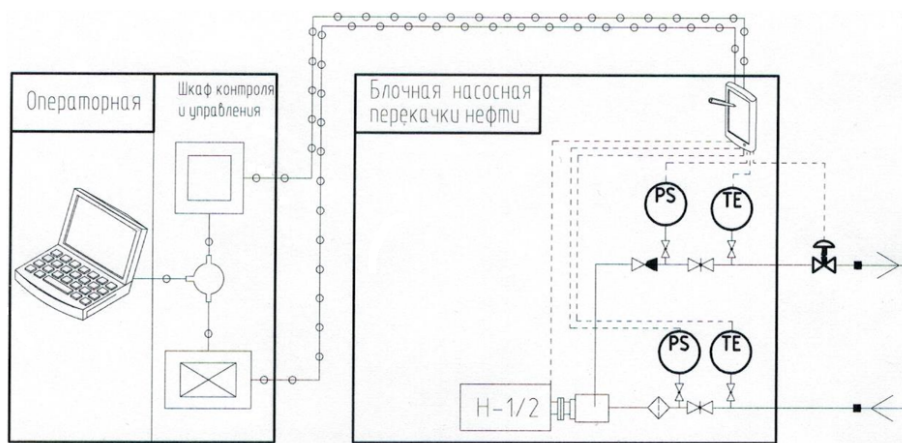


Рис. 2. Схема высокой автоматизации

УСЛОВНЫЕ ОБОЗНАЧЕНИЯ И ИЗОБРАЖЕНИЯ

Обозначение и изображение	Характеристика
	Интерфейс передачи данных (ETHERNET)
	Интерфейс передачи данных (RS-485)
	Нефтепровод технологический
	Клапан регулирующий с электроприводом
	Направление выхода жидкости
	Направление входа жидкости
	Задвижка ручная
	Комбинированный клапан
	Датчик давления
	Датчик температуры
	Сетчатый фильтр
	Насосный агрегат
	Станция управления насосной (по месту)
	Клапан обратный
	Рабочее место оператора
	Контроллер
	Искусственный интеллект
	Маршрутизатор

Рис. 3. Условные обозначения и изображения

признаках работы основного оборудования, например, таких как небольшие отклонения от обычных значений температуры оборудования, потребления электроэнергии и т.п. [12]. Такой дополнительный, внешний по отношению к основным системам контур безопасности позволит снизить уязвимость объектов КИИ к вредоносным воздействиям, являясь вспомогательной надстройкой к существующим программам, сетям и оборудованию.

Заключение

Таким образом, система управления на базе ИИ дает возможность дополнительного мониторинга состояния существующих систем без вмешательства в действующую программу управления электрооборудованием, при этом оперативный персонал получает полностью автономную систему наблюдения, которая по косвенным параметрам с датчиков позволяет обнаружить и пресечь любые внешние воздействия на системы управления на объекте.

Алгоритмы ИИ предоставляют в режиме реального времени возможность анализировать ключевые параметры работы оборудования, и при их отклонении от штатных показателей система на базе ИИ оказывает интеллектуальную поддержку оператору в форме актуальных алгоритмов, направленных на стабилизацию производственного процесса и его возвращение к штатному режиму работы.

Эффективным решением для обеспечения информационной и промышленной безопасности будет внедрение в существующие системы защиты и управления КИИ технологий на основе искусственного интеллекта. В этой связи разработка интеллектуальных моделей для применения в управлении КИИ представляется одной из наиболее актуальных задач.

ЛИТЕРАТУРА

1. Востров В. А., Гавриленко Т. В., Исмаилов Э. М. Некоторые аспекты разработки интеллектуальной модели для финансового риск-менеджмента нефинансовых организаций. *Успехи кибернетики*. 2023;4(4):74-80. DOI: 10.51790/2712-9942-2023-4-4-07.
2. Wen L., Li X., Gao L., Zhang Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Transactions on Power Delivery*. 2017;65(7):5990-5998. DOI: 10.1109/TIE.2017.2774777.
3. Мадуро заявил об атаках на энергосистему Венесуэлы изнутри. *Информационное агентство Интерфакс*. Режим доступа: <https://www.interfax.ru/world/653598>.
4. Ромашкина Н. П., Махукова А. В. Компьютерная вредоносная атака на ядерную программу Ирана. *Информационные войны*. 2013;(4):40–50.
5. О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации — городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона «О персональных данных». *Федеральный закон от 24.04.2020 № 123-ФЗ*. Режим доступа: <http://www.kremlin.ru/acts/bank/45475>.
6. Облачные сервисы: что это такое, какими они бывают и кому полезны. *Skillbox Media*. Режим доступа: <https://skillbox.ru/media/code/oblachnye-servisy-cto-takoe-kakimi-byvayut-i-komu-polezny/>.
7. Masa'd F. M., Al-maaitah T. A., Al-maaitah D. A., Qawasmeh E. F., Qatawneh N. A. Harnessing Artificial Intelligence for Human Resources Management: Tools, Advantages, and Risks in the Energy Sector. *E3S Web of Conferences*. 2024;541(4). DOI: 10.1051/e3sconf/202454102004.
8. Lee H. J., Ahn B. S., Park Y. M. A Fault Diagnosis Expert System for Distribution Substations. *IEEE Transactions on Power Delivery*. 2000;15(1):92-97. DOI: 10.1109/61.847234.
9. Sun J., Qin S. Y., Song Y. H. Fault Diagnosis of Electric Power Systems Based on Fuzzy Petri Nets. *IEEE Transactions on Power Delivery*. 2004;19(4):2053-2059. DOI: 10.1109/TPWRS.2004.836256.
10. Yongli Z., Limin H., Jinling L. Bayesian Networks-Based Approach for Power Systems Fault Diagnosis. *IEEE Transactions on Power Delivery*. 2006;21(2):634-639. DOI: 10.1109/TPWRD.2005.858774.
11. Iqbal R., Maniak T., Doctor F., Karyotis C. Fault Detection and Isolation in Industrial Processes Using Deep Learning Approaches. *IEEE Transactions on Power Delivery*. 2019;15(5):3077-3084. DOI: 10.1109/tii.2019.2902274.
12. Lo N. G., Flaus J. M., Adrot O. Review of Machine Learning Approaches in Fault Diagnosis Applied to IoT Systems. *International Conference on Control, Automation and Diagnosis*. 2019:1-6. DOI: 10.1109/ICCAD46983.2019.9037949.

ДЕЦЕНТРАЛИЗОВАННЫЕ ТЕХНОЛОГИИ В КАДАСТРОВОМ УЧЕТЕ НЕДВИЖИМОСТИ

С. В. Мещеряков^а, В. Ф. Шуршев^б*Астраханский государственный технический университет, г. Астрахань, Российская Федерация*^а  *Serg-93@list.ru*, ^б *v.shurshev@mail.ru*

Аннотация: современные системы учета недвижимости сталкиваются с проблемами безопасности, прозрачности и скорости обработки данных. Предложена инновационная блокчейн-модель кадастрового учета, обеспечивающая децентрализованное хранение данных, защиту от несанкционированных изменений и автоматизированную верификацию сделок. Разработан механизм регулирования комиссий в зависимости от численности населения и интеграции блокчейна с государственными реестрами. Представлены модели транзакций, алгоритмы работы системы и перспективы внедрения.

Ключевые слова: блокчейн, учет недвижимости, децентрализация, смарт-контракты, кадастр, IPFS.

Для цитирования: Мещеряков С. В., Шуршев В. Ф. Децентрализованные технологии в кадастровом учете недвижимости. *Успехи кибернетики*. 2025;6(2):80–85.

Поступила в редакцию: 11.04.2025.

В окончательном варианте: 22.04.2025.

DECENTRALIZED TECHNOLOGIES FOR CADASTRAL SYSTEMS

S. V. Meshcheryakov^а, V. F. Shurshev^б*Astrakhan State Technical University, Astrakhan, Russian Federation*^а  *Serg-93@list.ru*, ^б *v.shurshev@mail.ru*

Abstract: we analyzed the limitations of modern real estate accounting systems in terms of security, transparency, and data processing speed. To address these challenges, we proposed an innovative blockchain-based model for cadastral accounting that ensures decentralized data storage, protection against unauthorized modifications, and automated transaction verification. We developed a mechanism for regulating transaction fees based on population size and for integrating the blockchain system with state registries. We also presented transaction models, system operation algorithms, and potential implementation scenarios.

Keywords: blockchain, real estate registration, decentralization, smart contracts, cadastral system, IPFS.

Cite this article: Meshcheryakov S. V., Shurshev V. F. Decentralized Technologies for Cadastral Systems. *Russian Journal of Cybernetics*. 2025;6(2):80–85.

Original article submitted: 11.04.2025.

Revision submitted: 22.04.2025.

Введение

Традиционные кадастровые системы имеют централизованную структуру, что делает их уязвимыми для манипуляций, ошибок и взломов. Такие недостатки снижают доверие к системе учета недвижимости, усложняют процесс подтверждения прав собственности и создают препятствия для быстрого принятия решений при осуществлении сделок. Проблемы, связанные с дублированием записей, отсутствием прозрачности и длительными бюрократическими процедурами, требуют поиска инновационных решений [1–3].

Применение технологии блокчейн позволяет радикально изменить процесс учета недвижимости, обеспечивая неизменяемость данных, защиту от несанкционированного доступа и автоматизацию сделок. Использование смарт-контрактов минимизирует влияние человеческого фактора, снижая вероятность ошибок. Внедрение децентрализованной модели учета недвижимости способствует повышению эффективности кадастрового администрирования, создавая более прозрачную, безопасную и оперативную среду для всех участников рынка недвижимости [4, 5].

Блокчейн представляет собой распределенный реестр, который состоит из последовательности блоков, содержащих записи о транзакциях. Основное преимущество этой технологии заключается в неизменяемости данных: после добавления информации в блок ее нельзя удалить или подделать без изменения всей цепочки блоков, это достигается благодаря криптографическим методам защиты и механизмам консенсуса [6, 7].

В системе блокчейна каждая транзакция подписывается цифровой подписью владельца, после чего проверяется сетью узлов на предмет подлинности. Успешно подтвержденные транзакции объединяются в блоки, которые добавляются в цепочку с привязкой к предыдущему блоку, что исключает возможность фальсификации данных. Среди ключевых механизмов, обеспечивающих безопасность и функционирование системы, можно выделить [8]:

1. Механизм консенсуса — алгоритм, определяющий, как узлы сети достигают единого решения по валидации новых блоков (например, Proof-of-Stake или Practical Byzantine Fault Tolerance).
2. Криптографическое хеширование — применение алгоритмов хеширования (SHA-256) для создания уникальных идентификаторов блоков.
3. Децентрализация — отсутствие единого центра хранения данных, что повышает устойчивость системы к внешним атакам и сбоям.
4. Смарт-контракты — программируемые алгоритмы, автоматически выполняющие условия сделок при соблюдении заданных параметров.

Была предложена модель учета недвижимости с помощью технологий блокчейн, которую можно представить в виде кортежа:

$$\text{Блокчейн} = \{K, T, B, P\}, \quad (1)$$

где:

K — механизм консенсуса (PoS, PBFT), обеспечивающий верификацию транзакций;

T — набор транзакций, фиксирующих изменения прав собственности;

B — цепочка блоков, содержащих зашифрованные данные кадастра;

P — правила валидации данных и механизмы защиты.

В предложенной системе комиссия за обработку транзакций динамически изменяется в зависимости от численности населения региона. Это решение обусловлено необходимостью балансировки нагрузки на сеть, повышения скорости обработки сделок в мегаполисах и обеспечения доступности сервиса в малонаселенных регионах.

Разработана формула для баланса комиссий:

$$C = \alpha + \beta \times V \times P_{\text{пор}} \quad (2)$$

где:

C — комиссия за транзакцию;

α — фиксированная базовая комиссия за проведение транзакции;

β — коэффициент масштабирования, зависящий от объема данных;

V — объем данных, внесенных в кадастр;

$P_{\text{пор}}$ — коэффициент приоритета, связанный с численностью населения города, откуда отправляется транзакция.

Данный подход позволяет гибко настраивать систему, обеспечивая высокую пропускную способность там, где она необходима, и снижая затраты на проведение транзакций в регионах с низкой активностью сделок. В густонаселенных районах коэффициент $P_{\text{пор}}$ увеличивает комиссию, ускоряя обработку записей, а в малонаселенных регионах снижает, поддерживая доступность системы для пользователей.

Коэффициент приоритета рассчитывается по формуле:

$$P_{\text{пор}}(i) = \frac{N_i}{N_{\text{max}}}, \quad (3)$$

где:

$P_{\text{пор}}(i)$ — коэффициент приоритета;

N_i — население города, откуда отправлена транзакция i ;

N_{max} — население самого большого города в системе.

Данная формула добавляет приоритетность транзакциям из крупных городов, так как $P_{\text{пор}}(i)$ принимает значение от 0 до условно 100, где 100 соответствует самому крупному городу. Таким образом, комиссия C будет выше в густонаселенных районах, обеспечивая ускоренную обработку записей, и ниже в малонаселенных регионах, поддерживая доступность системы для всех пользователей [9].

Дополнительно система предусматривает динамическое изменение коэффициента α в зависимости от текущей нагрузки сети. В пиковые моменты параметр увеличивается, ускоряя обработку данных, а в периоды низкой активности — снижается, уменьшая стоимость транзакций.

В блокчейне модель адаптации комиссий выглядит следующим образом:

$$B = H(\phi + \tau + \gamma + \nu + \zeta) + (\alpha + \beta \times V \times P_{\text{pop}}), \quad (4)$$

где:

B — блокчейн;

H — криптографическая хеш-функция;

ϕ — предыдущий блок в цепи блоков;

τ — список транзакций, включаемых в текущий блок;

γ обозначает временную метку, указывающую время создания блока;

ν обозначает уникальный номер блока в цепи блоков;

ζ обозначает число (Нонс), которое изменяется для поиска подходящего хеша.

Формула, описывающая блокчейн как результат хеширования структурных данных блока, является основой понимания его работы и ключевым элементом, объясняющим механизмы его неизменяемости, безопасности и децентрализации. Включение в формулу таких элементов, как хеш предыдущего блока, транзакции, метка времени, номер блока и нонс, подчеркивает ее прикладное значение в реализации технологии. Данные элементы не только формируют структуру, но и задают поведенческую модель системы, что особенно важно для понимания, как блокчейн достигает своих уникальных характеристик.

Интеграция блокчейна с государственными реестрами

Внедрение технологии блокчейн в государственные реестры недвижимости представляет собой важный этап цифровой трансформации кадастровой системы, позволяющий устранить множество системных проблем, характерных для традиционных централизованных методов управления данными. Процесс интеграции блокчейна с государственными кадастровыми системами базируется на использовании API-интерфейсов, которые обеспечивают двусторонний обмен данными между существующими централизованными структурами и децентрализованным блокчейн-реестром. В рамках такой архитектуры все внесенные изменения фиксируются в блокчейне в форме неизменяемых записей, что предотвращает возможность несанкционированных изменений и манипуляций с информацией. Использование криптографических методов аутентификации позволяет гарантировать подлинность записей, сохраняя высокую степень защиты от несанкционированного доступа [10, 11].

Особое значение в системе интеграции приобретают смарт-контракты, которые автоматизируют процессы проверки данных и их юридической верификации. При изменении прав собственности или наложении обременений смарт-контракты выполняют алгоритмический анализ условий сделки, проверяют соответствие данных нормативно-правовым требованиям и фиксируют все изменения в блокчейне без необходимости ручного подтверждения со стороны операторов кадастра. Данная технология позволяет существенно снизить нагрузку на государственные структуры, исключая необходимость в промежуточных проверках и бумажном документообороте.

Дополнительным преимуществом интеграции блокчейна с кадастровыми реестрами является возможность обеспечения прозрачности сделок и их оперативной верификации. Поскольку вся история операций с объектами недвижимости фиксируется в блокчейне, заинтересованные стороны, включая нотариусов, банковские структуры и государственных регистраторов, могут получить мгновенный доступ к актуальной информации о статусе конкретного объекта. Это не только ускоряет процессы принятия решений, но и исключает возможность фальсификации данных или дублирования записей.

Государственные реестры, интегрированные с блокчейн-системой, получают значительные преимущества с точки зрения безопасности и отказоустойчивости. Традиционные централизованные базы данных подвержены рискам взлома, коррупции и технических сбоев, тогда как блокчейн, благодаря децентрализованной структуре, минимизирует вероятность разрушения или модификации данных из-за несанкционированного вмешательства. Даже в случае выхода из строя отдельных узлов сети блокчейн вся информация остается доступной благодаря множеству копий, распределенных среди участников сети. Кроме того, интеграция блокчейна с государственными кадастрами может значительно упростить

процессы аудита и контроля со стороны регулирующих органов. Поскольку все транзакции, связанные с объектами недвижимости, хранятся в неизменяемом виде, проверяющие инстанции могут оперативно анализировать историю владения, правовые обременения и юридическую чистоту сделок, исключая возможность манипуляций и ошибок, характерных для традиционных бумажных архивов и централизованных цифровых баз данных.

Таким образом, применение блокчейна в кадастровом учете не только обеспечивает прозрачность и достоверность данных, но и кардинально изменяет принципы взаимодействия между участниками рынка недвижимости. Снижение количества бюрократических процедур, повышение уровня автоматизации, уменьшение количества посредников и ускорение всех процессов, связанных с регистрацией и передачей прав собственности, делает блокчейн перспективной технологией для внедрения в государственные системы управления недвижимостью. В долгосрочной перспективе подобные изменения могут способствовать созданию глобальной децентрализованной сети кадастровых реестров, обеспечивая единый стандарт учета недвижимости и трансграничного взаимодействия между различными государственными системами. Блокчейн-реестр должен взаимодействовать с государственными кадастрами через API-интерфейсы, обеспечивая автоматическую сверку данных и их неизменяемость. Это взаимодействие реализуется через безопасные каналы передачи информации, где данные верифицируются в режиме реального времени, исключая человеческий фактор и снижая вероятность ошибок.

Механизм интеграции:

1. Передача данных в блокчейн — при внесении новой информации в государственный кадастр ее хеш фиксируется в блокчейне, что обеспечивает подтверждение подлинности.
2. Использование смарт-контрактов — каждая операция с недвижимостью (передача права собственности, наложение обременений) фиксируется с помощью смарт-контрактов, которые автоматически проверяют корректность изменений и соответствие законодательным требованиям.
3. Обратная связь с кадастром — при запросах из государственных реестров система сверяет данные в блокчейне, обеспечивая их актуальность и предотвращая возможные дублирования или расхождения в информации.
4. Проверка правового статуса объектов — блокчейн хранит историю всех транзакций, что позволяет мгновенно получать подтверждение о статусе объекта недвижимости без необходимости ручной проверки архивных данных.

Интеграция блокчейна с государственными реестрами создает единую систему контроля и регистрации недвижимости, повышая доверие к сделкам и ускоряя процедуры юридического оформления. Блокчейн-реестр должен взаимодействовать с государственными кадастрами через API-интерфейсы, обеспечивая автоматическую сверку данных. Смарт-контракты фиксируют изменения прав собственности и выполняют проверку на предмет обременений, исключая необходимость ручной обработки заявок, это позволяет значительно ускорить процессы регистрации, уменьшить количество ошибок и повысить прозрачность сделок. Данные процессы возможно отобразить в виде модели работы децентрализованной системы учета недвижимости на основе блокчейна (рис.).

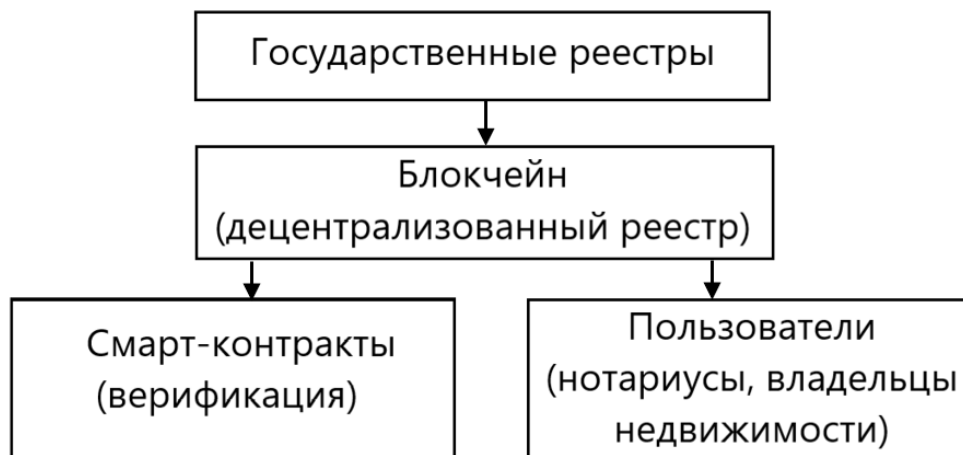


Рис. Модель работы децентрализованной системы учета недвижимости на основе блокчейна

Оптимизация хранения данных с использованием IPFS

Хранение полных данных об участках недвижимости (особенно пространственных) в блокчейне может быть неэффективным из-за его ограничений по объему. Для решения данной проблемы могут быть предложены технологии IPFS (межпланетная файловая система). В блокчейне сохраняются только ссылки на данные, хранящиеся в распределенной файловой системе. Это позволяет значительно снизить объем хранимой информации и ускорить операции [12].

Оптимизированное время обработки транзакций может быть рассчитано по следующей формуле:

$$T_{\text{opt}} = \frac{\sum_{i=1}^n d_i v_i P_{\text{пор}}}{\sum_{i=1}^n w_i}, \quad (5)$$

где:

T_{opt} — оптимизированное время обработки;

d_i — размер транзакции i ;

v_i — скорость обработки данных узлом;

w_i — вес транзакции i .

Данная формула учитывает влияние распределения данных в сети IPFS и блокчейне, а также приоритет транзакций. Хранение лишь хеш-ссылок на кадастровые записи позволяет не только экономить пространство, но и улучшать производительность за счет сокращения объема передаваемых данных и ускорения доступа к ним. В результате использования IPFS значительно уменьшается объем данных, которые необходимо хранить непосредственно в блокчейне, что делает обработку транзакций более эффективной и масштабируемой.

Заключение

Разработанная блокчейн-модель учета недвижимости сочетает в себе прозрачность, безопасность и автоматизацию процессов. Применение динамического регулирования комиссий позволяет повысить эффективность обработки сделок, а интеграция с государственными реестрами обеспечивает юридическую значимость записей. Технология блокчейн не только снижает риски мошенничества и несанкционированного доступа к данным, но и предоставляет уникальные возможности для оптимизации кадастровых процессов, сокращая время обработки сделок и уменьшая финансовые затраты.

Внедрение предложенной модели создает условия для полноценного перехода к цифровому кадастру, минимизируя влияние человеческого фактора и снижая административные издержки. В перспективе это может привести к созданию единой глобальной системы кадастрового учета, способной интегрировать данные между различными государственными структурами, что упростит трансграничные операции с недвижимостью. Таким образом, предложенная блокчейн-модель является не только технологическим решением, но и основой для трансформации кадастрового учета в новую, более совершенную форму управления недвижимостью.

ЛИТЕРАТУРА

1. Павлова В. А., Уварова Е. Л. Применение информационно-коммуникационных технологий в современной российской кадастровой системе. *Геодезия и картография*. 2019;2:57–63. DOI: 10.22389/0016-7126-2019-944-2-57-63.
2. World Bank Group. *Blockchain and Emerging Digital Technologies for Enhancing Post-2020 Climate Markets*. Washington, DC: World Bank; 2018.
3. Swan M. *Blockchain: Blueprint for a New Economy*. Sebastopol: O'Reilly Media; 2015. ISBN: 978-1-4919-1524-2.
4. Мещеряков С. В., Шуршев В. Ф. Применение технологий блокчейн для учета и регистрации земельных участков. *Инженерно-строительный вестник Прикаспия*. 2023;(4):112–115.
5. Мещеряков С. В., Шуршев В. Ф. Децентрализованный учет недвижимости: блокчейн и его применение. *Современные задачи и перспективные направления инновационного развития науки: материалы международной научно-практической конференции*. 2025:84–86.

6. Kshetri N. Can Blockchain Strengthen the Internet of Things? *IT Professional*. 2017;19(4):68–72. DOI: 10.1109/MITP.2017.3051335.
7. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Payment System. Режим доступа: <https://bitcoin.org/bitcoin.pdf>.
8. Swan M. The Blockchain: Brief History, Risks, and Opportunities. *IEEE Computer*. 2017;50(9):58–61. DOI: 10.1109/MC.2017.3571052.
9. Мещеряков С. В., Шуршев В. Ф. Адаптация комиссий в блокчейн-сети недвижимости. *Информационные системы и технологии*. 2024;(6):69–74.
10. Tapscott D., Tapscott A. *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World*. New York: Penguin; 2016. ISBN: 978-1-101-90363-2.
11. Grigg I. *Triple Entry Accounting*. Режим доступа: https://iang.org/papers/triple_entry.html.
12. Helmy K., Zaki M. Blockchain Technology for Accounting and Registration of Land Plots: Advantages and Challenges. *International Journal of Advanced Computer Science and Applications*. 2020;11(3):24–37. DOI: 10.14569/IJACSA.2020.0110304.

АНАЛИЗ МЕТОДОВ ИЗВЛЕЧЕНИЯ КЛЮЧЕВЫХ СЛОВ ИЗ ТЕКСТОВЫХ ДАННЫХ

Ш. И. Мутаиров^а, К. И. Бушмелева^б*Сургутский государственный университет, Сургут, Российская Федерация*^а *✉usermage@gmail.com*^б *bkiya@yandex.ru*

Аннотация: статья посвящена сравнительному анализу методов извлечения ключевых слов из текстовых данных с использованием статистических и нейросетевых подходов, таких как YAKE, и языковых моделей на основе трансформеров (например, BERT). В работе исследуются их точность, производительность и применимость к различным типам текстов. Основное внимание уделено разработке рекомендаций по оптимизации этих методов для повышения эффективности обработки текстов в условиях роста объемов данных. Результаты основаны на экспериментах с реальными текстами из научных статей.

Ключевые слова: обработка естественного языка, извлечение ключевых слов, YAKE, BERT, точность, производительность.

Для цитирования: Мутаиров Ш. И., Бушмелева К. И. Анализ методов извлечения ключевых слов из текстовых данных. *Успехи кибернетики*. 2025;6(2):86–91.

Поступила в редакцию: 14.03.2025.

В окончательном варианте: 02.04.2025.

COMPARATIVE ANALYSIS OF KEYWORD EXTRACTION STRATEGIES

Sh. I. Mutairov^а, K. I. Bushmeleva^б*Surgut State University, Surgut, Russian Federation*^а *✉usermage@gmail.com*^б *bkiya@yandex.ru*

Abstract: we conducted a comparative analysis of keyword extraction methods based on statistical and neural network approaches, including YAKE and transformer-based language models such as BERT. We evaluated their accuracy, performance, and suitability for different text types. Our primary goal was to give recommendations for optimizing these methods to enhance text processing efficiency as data volumes increase. We performed experiments with real scientific texts.

Keywords: natural language processing, keyword extraction, YAKE, BERT, accuracy, performance.

Cite this article: Mutairov Sh. I., Bushmeleva K. I. Comparative Analysis of Keyword Extraction Strategies. *Russian Journal of Cybernetics*. 2025;6(2):86–91.

Original article submitted: 14.03.2025.

Revision submitted: 02.04.2025.

Введение

Рост объемов данных в цифровой среде сопровождается увеличением потребности в автоматизированных методах анализа текстов [1]. Извлечение ключевых слов находит широкое применение в различных областях:

- научные исследования: упрощение индексации статей и поиска релевантной информации;
- информационные системы: повышение качества поиска и классификации документов;
- маркетинг и аналитика: анализ текстовых данных для выявления трендов и предпочтений пользователей.

Традиционные подходы к извлечению ключевых слов, такие как статистические методы (TF-IDF, YAKE), успешно справляются с задачей в условиях ограниченных ресурсов, однако их возможности ограничены при анализе сложных текстов с глубоким контекстом [2], [3]. С развитием языковых моделей (ЯМ) на основе трансформеров (например, BERT) появилась возможность учитывать семантические связи и контекст, что значительно повышает качество извлечения. Однако такие модели требуют больших вычислительных ресурсов и времени, что ограничивает их применение в реальных задачах с большими объемами данных [4].

Целью данного исследования является сравнительный анализ методов извлечения ключевых слов — статистического метода YAKE и языковой модели BERT — с точки зрения их точности, скорости и применимости. В статье также рассматриваются пути оптимизации этих методов для

достижения баланса между качеством и производительностью.

Сравнительный анализ методов извлечения ключевых слов

Существующие методы различаются по подходам: статистические алгоритмы (например, YAKE, RAKE) быстры и эффективны для коротких текстов, тогда как модели глубокого обучения (BERT) лучше справляются с длинными и сложными текстами, требующими учета контекста. Выбор подходящего метода зависит от задачи, объема данных и доступных ресурсов, что делает сравнительный анализ особенно актуальным.

В таблице 1 представлено сравнение характеристик трех подходов к извлечению ключевых слов.

Таблица 1

Сравнение подходов к извлечению ключевых слов

Характеристика	RAKE	YAKE	Языковая модель
Подход	Статистический	Статистический + контекст	Нейронные сети, семантический анализ
Точность	Низкая	Средняя	Высокая
Контекст	Не учитывается	Учитывается частично	Глубокий анализ контекста
Скорость	Очень высокая	Высокая	Низкая
Ресурсы	Минимальные	Низкие	Высокие
Мультиязычность	Ограничена	Частично	Полная (при поддержке модели)
Простота настройки	Простая	Простая	Сложная
Применение	Для коротких текстов	Для коротких и средних текстов	Для сложных и длинных текстов

В текущем исследовании рассмотрены варианты использования YAKE и языковой модели BERT. RAKE не подходит, так как его низкая точность ограничивает его применимость.

Метод YAKE

YAKE (Yet Another Keyword Extractor) — это статистический метод, использующий локальные особенности текста для ранжирования ключевых слов [5, 6]. Он основан на комбинации нескольких метрик, объединенных в итоговую оценку:

$$f_{score}(k_i) = TF(k_i) \cdot Pos(k_i) \cdot Spread(k_i) \cdot PC(k_i) \cdot Case(k_i), \quad (1)$$

где $TF(k_i)$ — частота термина в тексте;

$Pos(k_i)$ — взвешенная позиция слова (слова в начале текста считаются более значимыми);

$Spread(k_i)$ — разброс вхождений слова по тексту;

$PC(k_i)$ — плотность соседних слов (учитывает контекст);

$Case(k_i)$ — чувствительность к регистру.

Чем ниже значение f_{score} , тем выше значимость фразы k_i .

Метод на основе BERT

BERT (Bidirectional Encoder Representations from Transformers) — это языковая модель, использующая трансформеры для создания контекстных эмбеддингов [4, 7]. Алгоритм извлечения ключевых слов с помощью BERT включает следующие шаги:

1. Получение векторов: создание эмбеддингов для текста и отдельных слов/фраз.
2. Выбор кандидатов: извлечение потенциальных ключевых слов (n-грамм).
3. Сравнение: вычисление косинусного сходства между вектором текста и векторами кандидатов.
4. Ранжирование: выбор слов с наибольшим сходством.

В данном исследовании применялась модель `deeppavlov/rubert-base-cased` (180М параметров), адаптированная для русского языка.

Экспериментальное исследование

Для анализа было обработано 18 научных статей. В таблицах 2 и 3 представлены результаты извлечения данных по 6 статьям. Каждая статья сопровождалась списком ключевых слов, заданных авторами, что позволило оценить точность методов. Тестирование проводилось на следующих параметрах:

- процент совпадений извлеченных слов с авторскими ключевыми словами;
- среднее время обработки одного файла в секундах.

Таблица 2

Результаты извлечения ключевых слов из статей с помощью языковой модели

№	Тема статьи	Ключевые слова (автор статьи)	Языковая модель (deeppavlov/rubert-based – 180M)
1.	RFM-анализ работы автозаправочных станций в системе Loginom	RFM-анализ, автозаправочные станции, система Loginom, low-code платформа	решения автоматизации, технологическое решение, student scientific, основными драйверами, rfm analysis
2.	Применение искусственных нейронных сетей в идентификации значимых электрокардиографических параметров операторов технологических установок	искусственные нейронные сети, ЭКГ-параметры, значимые диагностические признаки	операторов технологических, способами статистическая, ai обработка, биофизики, electrocardiographic parameters
3.	Метод расчета коэффициента достоверности на уровне заголовков сегментов и пакетов	коэффициент достоверности данных, заголовки сегмента и пакета, расчет, безопасность данных	исследования сетевого, модификацию трафика, зашифрованного, использование коэффициента, detect data
4.	Многоразрядные вычисления как область приложения модулярной арифметики	многоразрядные вычисления, модулярная система счисления, модулярная система счисления квадратичного диапазона, модульные операции, математическое моделирование, программирование на языке Python, оценка сложности алгоритмов	решения автоматизации, вычислительной теории, уровне машинных, компьютере выполнить, проблем прикладной
5.	Прототип роботизированной системы транспортировки деталей	роботизированная система, Robotino, транспортировка, компьютерное зрение	робототехники основанный, prototype of, wiki openrobotino, scientific supervisor, transportation computer
6.	Проектирование информационно-аналитической системы «Достижения студентов СурГУ»	информационно-аналитическая система, функциональная модель, информологическая и физическая модели	информационно-аналитическая, возможностей системы, диаграммы функциональной, образовательным учреждениям, типов данных

Таблица 3

Результаты извлечения ключевых слов с помощью YAKE

№	Тема статьи	Ключевые слова (автор статьи)	YAKE
1.	RFM-анализ работы автозаправочных станций в системе Loginom	RFM-анализ, автозаправочные станции, система Loginom, low-code платформа	RFM, LOGINOM, клиентов, RFM ANALYSIS, анализа, АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ УДК
2.	Применение искусственных нейронных сетей в идентификации значимых электрокардиографических параметров операторов технологических установок	искусственные нейронные сети, ЭКГ-параметры, значимые диагностические признаки	ОПЕРАТОРОВ ТЕХНОЛОГИЧЕСКИХ УСТАНОВОК, ЭКГ-параметров операторов технологических, ЭКГ-параметров операторов, обработки ЭКГ-параметров операторов, ИДЕНТИФИКАЦИИ ЗНАЧИМЫХ ЭЛЕКТРОКАРДИОГРАФИЧЕСКИХ
3.	Метод расчета коэффициента достоверности на уровне заголовков сегментов и пакетов	коэффициент достоверности данных, заголовки сегмента и пакета, расчет, безопасность данных	коэффициент достоверности данных, достоверности данных сетевого трафика, компьютерной сети предприятия, РАСЧЕТА КОЭФФИЦИЕНТА ДОСТОВЕРНОСТИ
4.	Многоразрядные вычисления как область приложения модулярной арифметики	многоразрядные вычисления, модулярная система счисления одинарного диапазона, модулярная система счисления квадратичного диапазона, модульные операции, математическое моделирование, программирование на языке Python, оценка сложности алгоритмов	МСС квадратичного диапазона, МСС квадратичного, квадратичного диапазона, МСС, квадратичного, диапазона
5.	Прототип роботизированной системы транспортировки деталей	роботизированная система, Robotino, транспортировка, компьютерное зрение	Иван Сергеевич Курочкин, ДЕТАЛЕЙ Иван Сергеевич, Сургутский государственный университет, ПРОТОТИП РОБОТИЗИРОВАННОЙ СИСТЕМЫ, ПРОТОТИП РОБОТИЗИРОВАННОЙ
6.	Проектирование информационно-аналитической системы «Достижения студентов СурГУ»	информационно-аналитическая система, функциональная модель, инфологическая и физическая модели	Сургутский государственный университет, ДОСТИЖЕНИЯ СТУДЕНТОВ СУРГУ, ДОСТИЖЕНИЯ СТУДЕНТОВ, кафедры АСОИУ, СТУДЕНТОВ СУРГУ

В таблице 4 приведена оценка экспериментальных данных по ранее представленным методам.

Таблица 4

Оценка экспериментальных данных

Характеристика	YAKE	Языковая модель
Точность	40%. Демонстрирует более высокую точность благодаря статистическому подходу, который лучше выявляет часто встречающиеся термины	30%. Показывает меньшую точность, так как склонен извлекать слова с глубоким контекстным значением, которые не всегда совпадают с авторскими ожиданиями
Производительность	0.30 сек.	1.52 сек.

Анализ экспериментальных данных, представленный в таблице 4, показал, что YAKE лучше справляется с текстами, где ключевые слова представляют собой часто повторяющиеся термины или фразы (например, «RFM», «коэффициента достоверности данных», «прототип роботизированной системы»). Это делает его более подходящим для статей с технической направленностью, где терминология играет ключевую роль.

BERT показывает потенциал в извлечении контекстно-зависимых слов, что может быть полезно для текстов с более абстрактной тематикой или сложной структурой, однако в данном эксперименте его результаты оказались менее точными из-за несоответствия с авторскими ожиданиями.

Оптимизация методов

Для получения преимущества обоих методов и достижения баланса между точностью и производительностью предлагается 2 варианта гибридного алгоритма (таблица 5).

Таблица 5

Описание вариантов комбинации методов для извлечения ключевых слов

№	Вариант комбинации	Преимущества	Недостатки	Вывод
1.	YAKE формирует кандидатов, ЯМ пересортирует и берет топ-N	быстрый отбор кандидатов (YAKE); семантическая точность от ЯМ; снижение шума	зависит от качества кандидатов YAKE; возможна потеря редких, но важных слов	хороший баланс скорости и точности, но ограничен первым шагом
2.	Средний вес от YAKE и KeyBERT, затем ранжирование	баланс статистики и семантики; учет сильных сторон обоих методов; высокая точность	сложность нормализации весов; возможные искажения при усреднении	лучший подход по точности и гибкости, но чуть сложнее в реализации

В первом варианте метод YAKE как статистический алгоритм быстро анализирует текст и выделяет потенциальные ключевые слова на основе частоты, позиции и других локальных характеристик. Это позволяет минимизировать начальные вычисления и сосредоточиться на наиболее вероятных терминах. Далее с помощью языковой модели пересчитываются веса кандидатов с использованием косинусного сходства между векторами текста и отдельных слов или фраз. Это улучшает точность за счет учета семантического контекста.

Во втором варианте метод объединяет статистические веса от YAKE с контекстными эмбедами от языковой модели. Средний вес позволяет сбалансировать точность и релевантность. Ранжирование на основе среднего веса снижает влияние ошибок одного из методов. Например, если YAKE переоценит частотный термин, а языковая модель скорректирует его семантическую значимость, итоговый результат будет более надежным.

Вариант 1 предпочтителен для сценариев с ограниченными ресурсами или необходимостью быстрой обработки (например, обработка потоков данных в реальном времени).

Вариант 2 лучше подходит для задач, где приоритет отдается максимальной точности и устойчивости, таких как анализ научной литературы или написание аннотаций.

Заключение

Экспериментальные данные показывают, что YAKE является более эффективным методом для извлечения ключевых слов из научных текстов с технической направленностью, обеспечивая точность 40% и высокую скорость обработки (0.30 сек.). BERT, несмотря на меньшую точность (30%) и большее время обработки (1.52 сек.), демонстрирует потенциал для задач, требующих глубокого контекстного анализа.

Ограничения эксперимента связаны с выборкой текстов: все статьи имеют техническую направленность, что могло повлиять на результаты в пользу YAKE. В дальнейшем целесообразно протестировать методы на текстах с более абстрактной тематикой, таких как гуманитарные или художественные произведения, чтобы оценить эффективность BERT в условиях сложного контекста.

Комбинированные подходы, объединяющие преимущества YAKE и BERT, позволяют достичь оптимального баланса между точностью и производительностью, что делает их перспективными для дальнейшего развития и применения в системах обработки текстов.

Результаты исследования могут быть использованы в системах автоматической индексации, научных поисковиках и аналитических платформах. В дальнейшем планируется расширить эксперименты и интегрировать методы с полнотекстовым и семантическим поиском.

ЛИТЕРАТУРА

1. Campos R., Mangaravite V., Pasquali A., Jorge A., Nunes C., Jatowt A. YAKE! Keyword Extraction from Single Documents Using Multiple Local Features. *Information Sciences*. 2020;509:257–289.
2. Ванюшкин А. С., Гращенко Л. А. Оценка алгоритмов извлечения ключевых слов: инструментарий и ресурсы. *Новые информационные технологии в автоматизированных системах*. 2017;20:95–102. Режим доступа: <https://cyberleninka.ru/article/n/otsenka-algoritmov-izvlecheniya-klyuchevyh-slov-instrumentariy-i-resursy>.
3. Ванюшкин А. С., Гращенко Л. А. Методы и алгоритмы извлечения ключевых слов. *Новые информационные технологии в автоматизированных системах*. 2016;19:85–93. Режим доступа: <https://cyberleninka.ru/article/n/metody-i-algoritmy-izvlecheniya-klyuchevyh-slov>.
4. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805.
5. Горлевич Д. Алгоритмы для выделения ключевых слов: Rake, YAKE!, TextRank. *NewTechAudit*. Режим доступа: <https://newtechaudit.ru/algoritmy-dlya-vydeleniya-klyuchevyh-slov-rake-yake-textrank>.
6. Мутаиров Ш. И., Бушмелева К. И. Алгоритмы обработки и вычисления сходства текстовых данных пользователей социальных сетей. *Успехи кибернетики*. 2023;4(1):33–38.
7. Sharma P., Li Y. Self-Supervised Contextual Keyword and Keyphrase Retrieval with Self-Labelling. *Preprints.org*. Режим доступа: <https://www.preprints.org/manuscript/201908.0073/v1>.

МОДЕЛИРОВАНИЕ ЗВУКА В СИСТЕМАХ ВИРТУАЛЬНОГО ОКРУЖЕНИЯ

И. П. Саблин^{1,a}, М. В. Михайлюк^{1,b}, Д. В. Омельченко^{1,c}, Д. А. Кононов^{2,d}¹ Федеральное государственное автономное учреждение «Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Москва, Российская Федерация² Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук, г. Москва, Российская Федерация^a ✉ sablinivan97@gmail.com^b ORCID: <http://orcid.org/0000-0002-7793-080X>, mix@niisi.ras.ru^c omelchenko_dv@mail.ru^d ORCID: <http://orcid.org/0000-0002-6059-5590>, dmitrykon52@gmail.com

Аннотация: в работе представлена технология и методы реализации подсистемы звука в системах виртуального окружения. Основное внимание уделено вопросам внедрения акустических компонентов в 3D-сцены и разработке программного обеспечения для управления звуком. В качестве решения предлагается метод подготовки объектов типа «Источник звука» в системе трехмерного моделирования 3DS MAX с последующим экспортом в целевую платформу. Для реализации звука в реальном времени разработано программное обеспечение, обеспечивающее динамический расчет позиций источников звука и слушателя. Предложена классификация источников звука на стационарные (гудок, сирена, рупор и т.д.), динамические (работа лифта, холостой ход и звук двигателя робота под нагрузкой и др.) и природные (например, дождь или град различной интенсивности), а также методы определения моментов звукового сопровождения и проигрывания звуковых файлов. Для стационарных источников предложена методика создания виртуальных пультов управления с функциями включения и выключения звука, регулировки громкости и паузы. Апробация предлагаемых методов и подходов была проведена в созданном комплексе виртуального окружения VirSim. Результаты апробации показали адекватность предлагаемых решений и их применимость для моделирования звука в системах виртуального окружения.

Ключевые слова: система виртуального окружения, источник звука, пульт управления, библиотека Miniaudio.

Благодарности: публикация выполнена в рамках государственного задания НИЦ «Курчатовский институт» — НИИСИ по теме № FNEF-2024-0002 «Математическое моделирование многомасштабных динамических процессов и системы виртуального окружения».

Для цитирования: Саблин И. П., Михайлюк М. В., Омельченко Д. В., Кононов Д. А. Моделирование звука в системах виртуального окружения. *Успехи кибернетики*. 2025;6(2):92–99.

Поступила в редакцию: 29.04.2025.

В окончательном варианте: 15.05.2025.

SOUND MODELING IN VIRTUAL ENVIRONMENT SYSTEMS

I. P. Sablin^{1,a}, M. V. Mikhailyuk^{1,b}, D. V. Omelchenko^{1,c}, D. A. Kononov^{2,d}¹ Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, Moscow, Russian Federation² V. A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russian Federation^a ✉ sablinivan97@gmail.com^b ORCID: <http://orcid.org/0000-0002-7793-080X>, mix@niisi.ras.ru^c omelchenko_dv@mail.ru^d ORCID: <http://orcid.org/0000-0002-6059-5590>, dmitrykon52@gmail.com

Abstract: we presented technologies and methods for implementing a sound subsystem within virtual environment systems. We focused on integrating acoustic components into 3D scenes and developing sound management software. As a solution, we proposed a method for preparing Sound Source objects in the 3DS MAX three-dimensional modeling system, followed by export to the target platform. To enable real-time sound implementation, we developed software that dynamically calculates the positions of sound sources and the listener. We introduced a classification of sound sources into immovable (e.g. horn, siren, megaphone),

dynamic (e.g. elevator operation, idling or loaded robot engine sounds), and ambient (e.g. varying rain or hail intensity), and outlined methods for determining the timing of sound accompaniment and playback of sound files. For stationary sources, we proposed a method for creating virtual control panels with functions such as sound activation, volume control, and pause. We tested the proposed methods and approaches within the virtual environment system VirSim. The results confirmed the adequacy of the proposed solutions and their applicability to sound modeling in virtual environments.

Keywords: virtual environment system, sound source, control panel, Miniaudio librar.

Acknowledgements: this is a part of the FNEF-2024-0002 Simulation of Multiscale Dynamic Processes and Virtual Environments government contract granted to the Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”.

Cite this article: Sablin I. P., Mikhailyuk M. V., Omelchenko D. V., Kononov D. A. Sound Modeling in Virtual Environment Systems. *Russian Journal of Cybernetics*. 2025;6(2):92–99.

Original article submitted: 29.04.2025.

Revision submitted: 15.05.2025.

Введение

Одним из важных направлений в современных исследованиях является создание и усовершенствование систем виртуального окружения. Данные системы применяются во многих сферах, в том числе при создании тренажеров (медицинских, авиационных, строительных, архитектурных и др.), для создания игр, виртуальных выставок и т.д. Для более глубокого погружения в виртуальную реальность и большей реалистичности в ней необходимо влиять на различные органы чувств человека. Следовательно, систему виртуальной реальности удобно разделить на подсистемы, отвечающие за каждый орган соответственно. В данной статье описывается разработка одной из этих подсистем — подсистемы звука.

К решению данной задачи может быть несколько подходов. В работах [1] и [2] рассматривалась возможность синтеза звука непосредственно во время работы системы виртуальной реальности. Данный метод помогает сэкономить на общем объеме системы, так как при синтезе звука отпадает необходимость в хранении самих аудиофайлов, но, с другой стороны, такой подход увеличит время расчета каждого шага моделирования. Другим подходом может быть разработка подсистемы звука на основе уже существующих библиотек для работы с аудиофайлами. Например, в статье [3] делается обзор аудио плагинов различных разработчиков, использующих игровой движок Unity. В диссертации [4] описывается создание сервера 3D звука для системы виртуальной реальности на основе аудио библиотеки OpenAL. Но библиотеки, использовавшиеся в двух последних работах, имеют ограничения по применению, поэтому в наших разработках используется бесплатная библиотека Miniaudio.

Miniaudio — это кроссплатформенная библиотека пространственного аудио, предназначенная для эффективного рендеринга многоканального позиционного аудио, написанная на языке программирования C. Miniaudio оформлена в виде одного файла с расширением h, содержащего в себе код, а также комментарии к нему. Для установки данной библиотеки необходимо скачать этот файл с аккаунта ее разработчиков в GitHub [5] и подключить к проекту C++, прописав в одном из файлов с расширением cpp следующие строчки:

```
#define MINIAUDIO_IMPLEMENTATION
#include "miniaudio.h"
```

Более подробная документация и примеры использования возможностей данной библиотеки находятся на сайте разработчиков [6].

Система виртуального окружения VirSim

Система VirSim [7] состоит из оболочки, а также подсистем управления, визуализации, динамики и звука (см. рис. 1).

Виртуальная сцена создается в системе трехмерного моделирования 3DS MAX и загружается в подсистемы динамики и визуализации. Оператор управляет с помощью реальных или виртуальных пультов управления динамическими объектами и эффектами, и эти управляющие сигналы передаются через оболочку в подсистемы динамики и визуализации. Подсистема динамики вычисляет новые координаты и ориентации объектов сцены и передает их через оболочку в подсистему визуализации, которая выводит на монитор оператора вид сцены после изменений. Этот цикл занимает не более 40

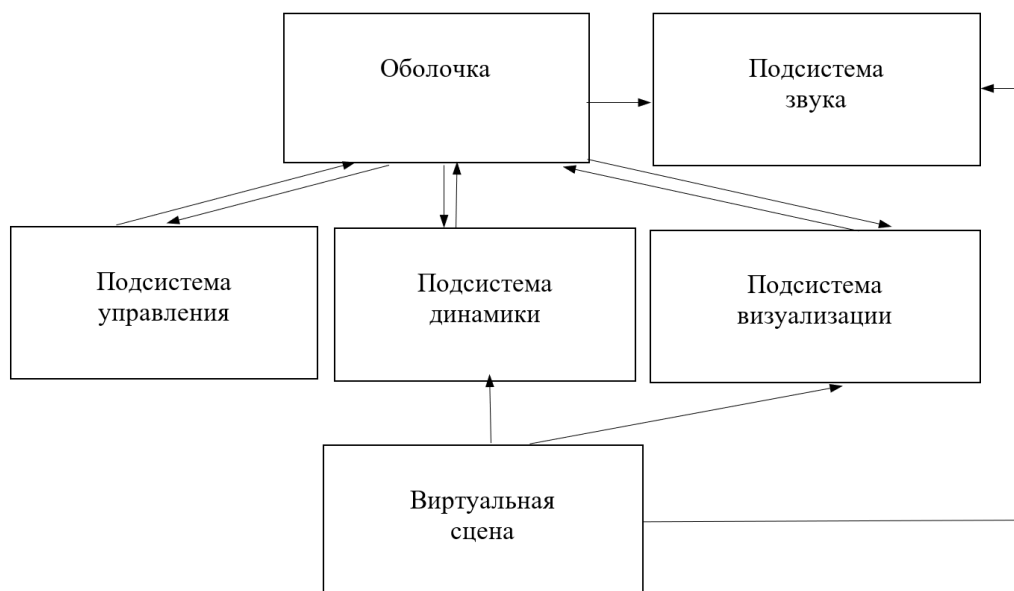


Рис. 1. Блок-схема системы VirSim

мсек, т. е. работа системы происходит в масштабе реального времени. Работа подсистемы звука в системе VirSim организована таким же образом. В подсистему звука загружается виртуальная сцена, а необходимые для ее работы данные от других подсистем передаются через оболочку.

Библиотека Miniaudio работы со звуком

Библиотека Miniaudio предоставляет возможность работы с низкоуровневым и высокоуровневым API. Низкоуровневый API дает прямой доступ к аппаратным ресурсам системы для обработки звука. Он позволяет управлять такими процессами, как буферизация данных, частота дискретизации, обработка многоканального звука, а также работа с конкретными устройствами вывода и ввода. Высокоуровневый аудио API предлагает решения для выполнения стандартных задач, таких как воспроизведение звуков, управление эффектами, пространственное позиционирование и так далее.

Для работы со звуком надо инициализировать ряд структур, в частности, «движок» (*engine*), который инициализируется функцией

ma_engine_init(NULL, &engine).

Параметры источника звука. Для инициализации источника звука из звукового файла можно использовать функцию

ma_sound_init_from_file(&engine, имя файла, 0, NULL, NULL, &sound).

Здесь источник звука *sound* является структурой, в которую запишутся параметры файла. Остальные параметры заданы значением *NULL* — это означает, что будут использованы их значения по умолчанию.

Источник звука имеет следующие основные параметры: координаты положения в мировой системе координат (МСК), вектор направления, громкость и звуковые конусы. Положение *S* источника звука задается функцией

ma_sound_set_position(&sound, S_x, S_y, S_z),

а вектор *F* направления — при помощи функции

ma_sound_set_direction(&sound, F_x, F_y, F_z).

Громкость звука изменяется в зависимости от расстояния от источника звука. Miniaudio поддерживает несколько моделей затухания звука в пространстве, в том числе модель Inverse Distance Clamped Model, описанную в документации библиотеки OpenAL [8] для работы с аудио. Для нее операторами

ma_sound_set_min_distance(&sound, D_{min})

ma_sound_set_max_distance(&sound, D_{max})

ma_sound_set_rolloff(&sound, k)

можно задать минимальное ($D_{\min}$) и максимальное ($D_{\max}$) расстояния от источника звука, а также коэффициент k затухания в области между этими границами (см. рис. 2). В области от источника звука до границы $D_{\min}$ громкость звука будет совпадать с громкостью V_S в точке S , т. е. коэффициент затухания $g_d = 1$. В любой точке P между этими границами, находящейся на расстоянии d от источника звука, коэффициент затухания громкости вычисляется по формуле:

$$g_d = \frac{D_{\min}}{D_{\min} + k_d(d - D_{\min})},$$

так что громкость в этой точке будет равна:

$$V_P = g_d V_S.$$

После границы $D_{\max}$ коэффициент затухания в любой точке равен $g_{D_{\max}}$.

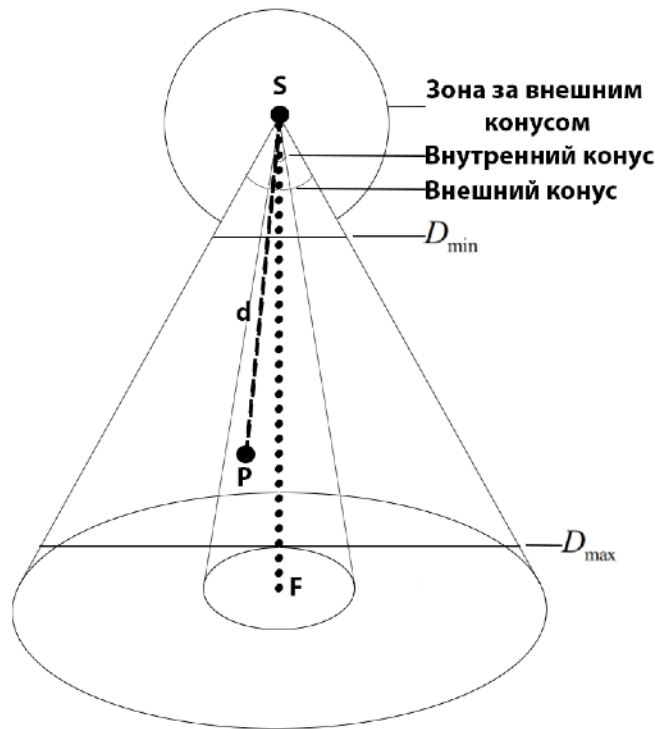


Рис. 2. Схема затухания звука

Затухание звука происходит также в зависимости от угла отклонения от направления источника звука. Для этого задаются углы *inAng* внутреннего и *outAng* внешнего звуковых конусов и коэффициент k_a затухания вне внешнего конуса с помощью оператора

`ma_sound_set_cone(&sound, inAng, outAng, k_a).`

Внутри внутреннего конуса коэффициент затухания $g_a = 1$, т. е. затухания звука по направлению не будет. Вне внешнего конуса $g_a = k_a$, а между конусами g_a будет линейно интерполироваться от 1 до k_a .

Общий коэффициент затухания равен произведению коэффициентов затухания по расстоянию и по углу, т. е. $g = g_a g_d$, а громкость звука в точке P будет вычисляться по формуле

$$V_P = g_a g_d V_S.$$

Параметры слушателя. В общем случае библиотека *Miniaudio* поддерживает обработку нескольких слушателей, но в нашем случае в сцене может быть только один слушатель (ему присваивается номер 0). Основными параметрами слушателя являются координаты положения в МСК, вектор направления и вектор «верха», а также звуковые конусы. Положение L слушателя задается при помощи функции

`ma_engine_listener_set_position(&engine, 0, L_x, L_y, L_z).`

Вектор V направления и вектор U «верха» слушателя задаются при помощи операторов

$ma_engine_listener_set_direction(\&engine, 0, V_x, V_y, V_z)$
 $ma_engine_listener_set_world_up(\&engine, 0, U_x, U_y, U_z).$

Как и для источника звука, для слушателя задаются звуковые конусы углами $inAL$, $ouAL$ и коэффициент k_L при помощи оператора

$ma_engine_listener_set_cone(\&engine, 0, inAL, ouAL, k_L).$

Звук от источника попадает к слушателю по прямой линии (направлению), соединяющей источник и слушателя. Коэффициент затухания звука g_L в зависимости от направления вычисляется по такому же принципу, как и у источника звука. А именно: если направление попадает во внутренний угол слушателя, то $g_L = 1$, между внутренним и внешним углами g_L линейно интерполируется от 1 до k_L , а вне внешнего угла $g_L = k_L$. Таким образом, громкость звука в точке L будет вычисляться по формуле:

$$V_L = g_L g_{a,L} g_{d,L} V_S.$$

Реализация звука в системе виртуального окружения VirSim

Создание источника звука. Для создания объекта типа «Источник звука» в системе трехмерного моделирования 3DS MAX разработан новый плагин, который позволяет через диалоговое окно (см. рис. 3) задать параметры источника звука в сцене. Параметры для всех источников звука в сцене экспортируются в бинарный файл с расширением sls, который сохраняется в папке, где находятся остальные конфигурационные файлы сцены.

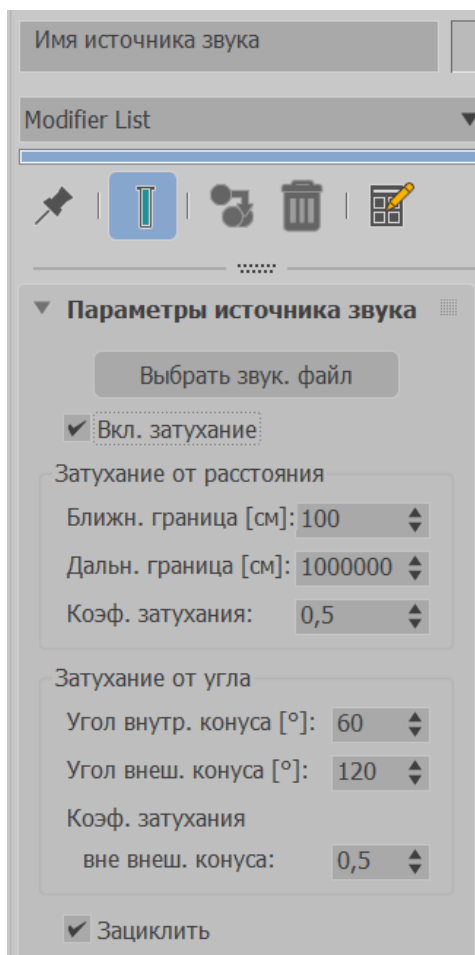


Рис. 3. Параметры источника звука в 3DS MAX

Также в данной папке мы будем хранить все заранее подготовленные аудиофайлы, которые будут воспроизводиться в данной сцене. Аналогично создаются и сохраняются в отдельных папках источники звука для каждого робота (звук работающего мотора, поворота манипулятора, вращения гусеницы и др.).

Создание программы обработки аудио. Программа для реализации звука в системе VirSim создана в виде динамической библиотеки (dll), в которую включены (с помощью директивы `#include`) библиотека `Miniaudio.h`, файлы `Sound.h` и `Sound.cpp` с объявлением и реализацией класса источника звука (в терминах C++), а также файлы `SoundLib.h` и `SoundLib.cpp`, содержащие объявления и тела экспортных функций, которые будут вызываться в процессе моделирования. Класс слушателя не создается, т. к. слушатель единственный.

Инициализация источников звука и слушателя. Инициализация значений всех параметров производится один раз в начале работы системы VirSim.

1. Загружаются и сохраняются параметры всех источников звука. Для этого используется экспортная функция `LoadSoundObjects`, в которую передается путь к sls-файлу, а также (для источников звука, расположенных на роботе) номер робота в сцене. В ней также создается список `SndObjectList` объектов класса `SndSource`, в каждый из которых записываются данные из sls-файла, относящиеся к одному источнику звука.

2. Инициализируется «движок» библиотеки `Miniaudio`, как описано в разделе *Библиотека Miniaudio работы со звуком*, с применением для этого экспортной функции `SoundInitialize`. В этой функции также задаются звуковые конусы слушателя при помощи функции, описанной в подразделе *Параметры слушателя*. Эти конусы не будут меняться в процессе работы системы, и для них устанавливаются следующие значения: угол внутреннего конуса — 30° , угол внешнего — 90° , коэффициент затухания за внешним конусом — 0.5.

Затем инициализируется источник звука и задаются его начальные параметры. Для этого создана функция `SetSoundSource`, которая, в свою очередь, циклично вызывается в функции `SoundInitialize`.

Моделирование работы источников звука. При моделировании звука во время работы системы VirSim мы различаем три типа источников звука: стационарные (положение и направление которых не изменяется, например, гудок, сирена, рупор и др., см. рис. 4), динамические (меняющие свои положение и направление) и природные (отличаются от стационарных только тем, что громкость их звука не зависит от затухания по расстоянию и углового затухания). На каждом шаге моделирования для любого стационарного источника звука с помощью его пульта управления можно включить, выключить, поставить на паузу или изменить громкость звука. Пример такого пульта показан на рис. 5. С помощью callback функции `PanelSignalDataHandler` эти данные передаются в оболочку, которая с помощью экспортной функции `CtrlPanelData` передает их в подсистему звука.

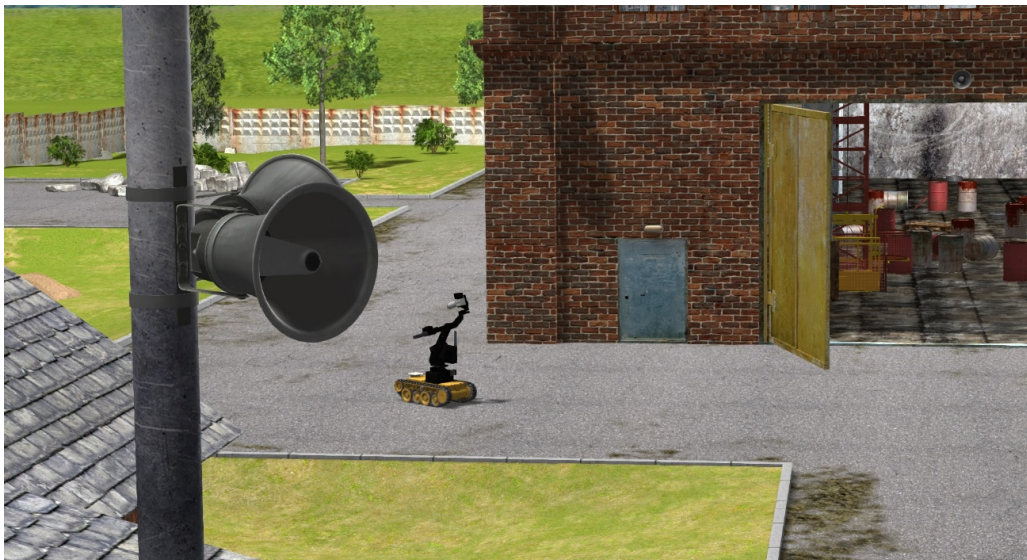


Рис. 4. Сирена в сцене «Полигон»

Для воспроизведения звука используются созданные ранее звуковые файлы и следующие функции, которые потребуются для работы со звуковыми файлами:

`ma_sound_start(&sound)` — воспроизводит звуковой файл;

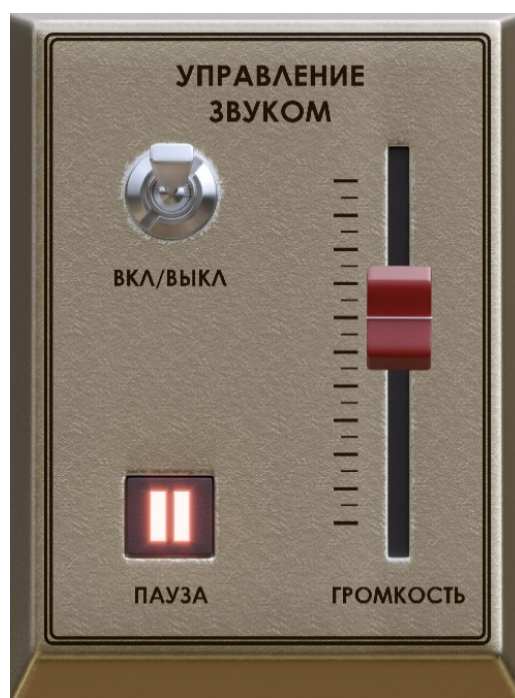


Рис. 5. Пульт управления звуком

ma_sound_stop(&sound) — останавливает воспроизведение звукового файла;

ma_sound_seek_to_pcm_frame(&sound, N) — начинает воспроизведение аудиофайла с кадра *N*;

ma_sound_set_looping(&Sound, isLoop) — зацикливает воспроизведение звукового файла.

Динамическими источниками звука являются источники, меняющие свое положение или ориентацию. Примерами таких источников являются звук работы лифта, звук холостого хода двигателя робота и под нагрузкой. В них громкость зависит не только от расстояния и угла, но также, например, от скорости движения. Для динамических источников звука подсистема динамики с помощью callback функции *DynDataHandler* передает в оболочку имя источника звука, его позицию и направление. Эти данные оболочка передает в подсистему звука с помощью экспортной функции *SndSourceCoords*. В ней работает цикл по объектам списка *SndObjectList*, в котором для каждого изменившего параметры объекта записываются его новые положение и направление.

Примером природного источника звука является дождь. Отличие звука дождя от статических и динамических звуков состоит в том, что его громкость не будет зависеть от каких-либо параметров затухания, т. к. он не локализован в одной точке. Для того чтобы отключить затухания звука в зависимости от расстояния и угла, при инициализации аудиофайла дождя в функции *ma_sound_init_from_file()* в качестве 3-го параметра необходимо передать флаг *MA_SOUND_FLAG_NO_SPATIALIZATION*.

В каждом шаге моделирования необходимо передать в подсистему звука также положение слушателя. В системе VirSim его положение и ориентация совпадают с положением и ориентацией виртуальной камеры, с помощью которой оператор наблюдает виртуальную сцену. Эти данные на каждом шаге моделирования подсистема визуализации передает с помощью callback функции *CameraDataHandler* в оболочку, которая, в свою очередь, при помощи экспортной функции *ListenerCoords* передает их в подсистему звука.

Заключительным шагом моделирования является установка и ориентация динамических объектов в соответствии с новыми данными, а также запуск звуковых файлов в соответствии с их текущими параметрами.

После окончания моделирования вызывается экспортная функция *DestroySoundObjects*, в которой все объекты списка *SndObjectList* очищаются в цикле при помощи функции *ma_sound_uninit(&sound)* и удаляются при помощи оператора *delete*, а сам список очищается при помощи оператора *clear*. «Движок» библиотеки *Miniaudio*, в свою очередь, очищается при помощи функции *ma_engine_uninit(&engine)*.

Заключение

В данной работе рассматривается реализация подсистемы звука в системах виртуального окружения. Предлагается механизм подготовки источников звука для виртуальной сцены, способ вычисления позиции слушателя, определения моментов активизации и проигрывания звуковых файлов с использованием библиотеки Miniaudio. Разработанные методы объединены в рамках подсистемы работы со звуком. Апробация этих методов в системе виртуального окружения VirSim показала их адекватность поставленным задачам.

ЛИТЕРАТУРА

1. Russ M. *Sound Synthesis and Sampling*. Focal Press; 2012. 464 p. DOI: 10.4324/9780080481043.
2. Cornelis Pieter van den Doel. *Sound Synthesis for Virtual Reality and Computer Games*. Ph.D. thesis, University of British Columbia; 1998. 186 p. DOI: 10.14288/1.0051161.
3. Nuora J. *Introduction to Sound Design for Virtual Reality Games: a Look into 3D Sound, Spatializer Plugins and Their Implementation in Unity Game Engine*. Bachelor's thesis, Tampere University of Applied Sciences; 2018. 47 p.
4. Schreier M. *Audio Server for Virtual Reality Applications*. A dissertation submitted in partial fulfilment of the requirements for the degree of Master of Science. Brunel; 2002. 75 p.
5. *Github Miniaudio*. Режим доступа: <https://github.com/mackron/miniaudio?tab=readme-ov-file#readme>.
6. *Документация Miniaudio*. Режим доступа: <https://miniaud.io/docs/>.
7. Михайлюк М. В., Мальцев А. В., Тимохин П. Ю., Страшнов Е. В., Крючков Б. И., Усов В. М. Система виртуального окружения VirSim для имитационно-тренажерных комплексов подготовки космонавтов. *Пилотируемые полеты в космос*. 2020;37(4):72-95. DOI: 10.34131/MSF.20.4.72-95.
8. *Документация OpenAL*. Режим доступа: <https://openal.org/documentation/>.

МЕТОДИКА ОЦЕНКИ И ВЫБОРА ОПТИМАЛЬНОЙ МОДЕЛИ, ПРОГНОЗИРУЮЩЕЙ ОКОНЧАНИЕ ПЕРИОДА ЗАМОРОЗКОВ

В. А. Солозобов^а, С. А. Лысенкова^б

Сургутский государственный университет, г. Сургут, Российская Федерация

^а ✉ solo.val.al@yandex.ru

^б ORCID: <http://orcid.org/0000-0003-1007-7610>, lsa1108@mail.ru

Аннотация: одной из первых задач при разработке математической модели прогнозирования является определение критерия, по которому будет оцениваться и выбираться оптимальная модель из ряда возможных. При этом решаемые задачи моделирования могут требовать различных методов оценки в зависимости от поставленной цели. Именно от того, каким образом будет оцениваться модель прогнозирования, зависит конечный итог. В данном исследовании была разработана методика оценки и выбора модели, прогнозирующей окончание периода заморозков в весенне-летний период.

Ключевые слова: методы оценки, метрика качества, нейронная сеть, прогноз погоды, заморозок.

Для цитирования: Солозобов В. А., Лысенкова С. А. Методика оценки и выбора оптимальной модели, прогнозирующей окончание периода заморозков. *Успехи кибернетики*. 2025;6(2):100–107.

Поступила в редакцию: 19.05.2025.

В окончательном варианте: 04.06.2025.

ASSESSMENT AND SELECTION OF AN OPTIMAL MODEL FOR FORECASTING THE END OF THE FROST PERIOD

V. A. Solozobov^а, S. A. Lysenkova^б

Surgut State University, Surgut, Russian Federation

^а ✉ solo.val.al@yandex.ru

^б ORCID: <http://orcid.org/0000-0003-1007-7610>, lsa1108@mail.ru

Abstract: one of the first steps in developing a mathematical forecasting model is to define the criteria for evaluating and selecting the optimal model from among several candidates. Since modelling objectives can vary, the choice of evaluation method must align with the specific goal of the study. The accuracy and usefulness of the final result depend heavily on how the model is assessed. In this study, we developed a procedure for evaluating and selecting a model to forecast the end of the frost period in the spring–summer season.

Keywords: assessment methods, quality metric, neural network, weather forecast, frost.

Cite this article: Solozobov V. A., Lysenkova S. A. Assessment and Selection of an Optimal Model for Forecasting the End of the Frost Period. *Russian Journal of Cybernetics*. 2025;6(2):100–107.

Original article submitted: 19.05.2025.

Revision submitted: 04.06.2025.

Введение

Прогноз завершения периода отрицательных температур и наступления устойчивой положительной температуры в весенне-летний период является важным и определяющим для большого количества видов деятельности человека, таких как сельскохозяйственный, строительный, транспортный и другие. Для расчета такого прогноза требуется создание модели прогнозирования окончания периода заморозков (ОПЗ). Один из вариантов такой модели — строить прогноз на каждый день и выдавать вероятность завершения периода заморозков на этот день. При этом высокая вероятность завершения заморозков будет говорить о том, что с этого дня будет устойчивая положительная температура воздуха. Данная информация полезна для уменьшения рисков и увеличения эффективности деятельности человека.

Одними из первых задач, стоящих перед разработкой такой модели, являются выбор и определение критерия оценки качества, по которому будет определяться оптимальность этой модели [1]. Однако выбор оценочных критериев, или, по-другому, метрик качества, не является простой задачей. Нужно учитывать, к какому типу относится решаемая проблема, свойство входных данных, желаемое соотношение ложноположительной ошибки к ложноотрицательной ошибке, подходы к обучению и настройке модели, тип модели и другое [2, 3]. С помощью правильно разработанной методики оценки качества можно будет сравнивать различные модели.

Таким образом, цель данного исследования — определить методику оценки моделей, по которой будет происходить их сравнение и выбор оптимальной модели прогнозирования ОПЗ. Для достижения поставленной цели нужно решить следующие задачи:

- 1) рассмотреть метрики качества, используемые в подобного вида задачах;
- 2) определить основные свойства применяемого класса моделей и характеристики входных и выходных данных;
- 3) выбрать или разработать свою методику оценки и выбора оптимальной модели прогнозирования ОПЗ.

Методы оценки моделей, прогнозирующих явления

Разрабатываемая модель решает задачу классификации. Она должна определить на день прогноза, наступило ли ОПЗ в данный день, на основе предшествующей метеобстановки. Данные метеобстановки представляют собой временные ряды со значениями метеопараметров за последние несколько недель или месяцев. То есть модель классифицирует данные временные ряды на предмет того, относятся ли они к периоду заморозков или в них появились признаки наступления устойчивой положительной погоды.

Основные методы оценки качества при решении задачи классификации основаны на матрице ошибок (таблица 1). В нашем случае таких классов два: отрицательный (Negative) и положительный (Positive).

Таблица 1

Матрица ошибок

Классификация (прогнозная)	Классификация (действительная)	
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

К позитивному классу относятся случаи продолжения заморозков, а случай окончания периода заморозков относится к негативному классу соответственно. Таким образом, FN — это случай, когда модель прогнозирует завершение периода заморозков, но на самом деле это не так. Данный случай несет наибольший ущерб, и его следует минимизировать в первую очередь.

Основные метрики, по которым делается оценка качества бинарной классификации, рассчитываются по значениям в матрице ошибок [1]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

$$Precision = \frac{TP}{TP + FP}, \quad (2)$$

$$Recall = \frac{TP}{TP + FN}, \quad (3)$$

Коэффициент корреляции Мэтьюса:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4)$$

$$F_\beta = \left(1 + \beta^2\right) \frac{Precision \cdot Recall}{(\beta^2 Precision) + Recall}, \quad (5)$$

где F-мера при $\beta = 1$ — среднегармоническая мера.

Также используют сравнение площадей под ROC-кривой (ROC-AUC) (англ. Receiver Operating Characteristics Curve — Area Under the Curve), оценку с помощью построения Precision-Recall (PR) кривой и расчет площади под PR-кривой.

Отдельно стоит упомянуть логистическую функцию ошибки, которая будет минимизироваться в процессе обучения модели. Она рассчитывается с помощью логарифмирования расхождений, рассчитанных вероятностей, принадлежности к классу и фактических значений. Казалось бы, минимальная логистическая ошибка должна указывать на оптимальную модель, но это не так, как выяснится дальше в данном исследовании.

Логистическая функция ошибки:

$$\text{LogLoss} = \frac{-1}{l} \sum_{i=1}^l (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)). \quad (6)$$

Одна из проблем, которая мешает выбрать стандартную метрику качества, заключается в том, что объем выборки будет различаться для моделей с разными входными параметрами. Например, это происходит при изменении скользящего окна — количества дней во временном ряде, которые подаются на вход прогнозирующей модели. Так, при изменении временного окна с 15 дней до 60 обучающая выборка уменьшится на 45 единиц. При этом уменьшение произойдет в положительном классе, который требуется прогнозировать в первую очередь. Изменение объема в выборке не позволит сравнивать различные модели по стандартным метрикам, так как объем класса будет изменяться, и при одинаковом количестве ошибок лучше себя покажет тот, у которого объем выборки больше.

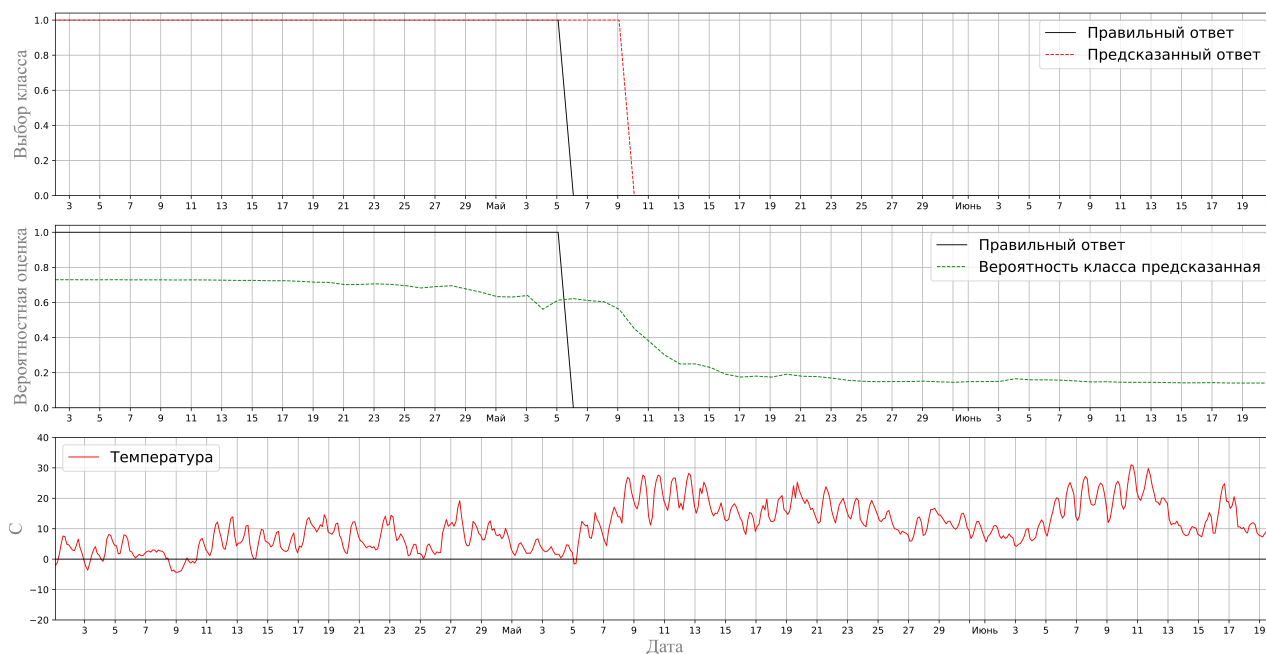


Рис. 1. Пример сравнения прогноза ОПЗ моделью и фактических данных

Еще одна причина того, почему стандартные метрики не подходят, — это неравномерное распределение ошибок по годовой выборке. Из рисунка 1 видно, что наиболее сложно для модели предсказать переходный период весна-лето, представляющий 20–40 дней, в котором она совершает ошибки. Начальная часть годовой выборки и конечная ее часть точно идентифицируются как зима и как лето, и от того, как много будет добавлено легких к прогнозированию дней (приходящихся на зиму или на лето), количество ошибок не изменится. Это еще одна причина, по которой из процесса вычисления метрики следует убрать общее количество экземпляров, по которым будет происходить обучение. Но и просто сравнивать количество ошибок нельзя, так как исчезнет возможность сравнивать модели, обученные на разных объемах выборок, ошибку следует привести к некоторому среднему показателю. Самый простой вариант — это рассчитывать отношение количества ошибок к количеству лет, участвующих в выборках. То есть мы получим, сколько в среднем за один прогнозируемый сезон модель совершит ошибок. Понятно, что некоторые года и их весенне-летний сезон прогнозируется хуже, чем другие, и на них будет приходиться больше ошибок, чем на другие. Но данная проблема будет решаться равномерным распределением сложных и простых случаев и использованием большего количества значений для обучения модели.

Еще одной важной проблемой является то, что ошибки первого и второго рода не равнозначны [4]. Следует в первую очередь снижать ошибку неверного прогнозирования ОПЗ. Но тут возможно возникновение проблемы: получение модели, у которой минимальная ошибка второго рода, но при этом ошибка первого рода настолько велика, что теряется смысл прогнозирования [3]. Требуется поиск сбалансированного решения. Поэтому при обучении модели будет производиться выбор по минимальной общей ошибке. А при сравнении моделей будет происходить отбор по минимальной ложнонегативной ошибке.

Для борьбы с переобучением, которое возможно, есть различные способы, но универсальный — это контроль по тестовой группе. Если по тестовой группе ошибка перестала падать и стала расти, начался процесс переобучения и запоминания выборки. Тут встает вопрос, какую все-таки выбрать модель: опираться на метрики тестовой модели или на средние метрики по обеим группам. В данном исследовании предлагается следующая система определения метрики качества по обучающей и тестовой выборке. Если тестовая ошибка выше обучающей, то она будет использоваться при оценивании модели. Если тестовая ошибка ниже обучающей, то будет находиться средняя ошибка по тестовой и обучающей выборкам, и она уже будет использоваться для оценки модели.

Также существует проблема наглядности и интерпретируемости метрик качества [5]. Хотя точность в 95% и полнота в 99% говорят о качественной модели, существует множество случаев, где таких цифр будет недостаточно в силу дисбаланса выборок как по классам, так и по количеству сложных случаев к общему объему выборки [6]. Также не всегда ясно, что выгоднее: снизить на полпроцента полноту, при этом увеличив на процент точность, или наоборот. Еще сложнее обстоят дела с другими метриками и сравнениями площадей под кривыми. С другой стороны, сравнивая именно количество ошибок, которые делает модель в среднем за сезон, ясно понимаешь возможности данной модели. И при регулировании порогового коэффициента сразу видно, как меняется баланс ложноположительной и ложноотрицательной ошибки, что дает четкое понимание результатов данной регуляции.

Методика оценки и выбора модели, прогнозирующей окончание периода заморозков

На основе вышеизложенного предлагается следующая методика сравнения и выбора оптимальной модели прогнозирования явления, в нашем случае ОПЗ:

1. Данные делятся на две выборки: обучающую и тестовую в пропорциях 2 к 1 соответственно.
2. В процессе обучения полученная матрица ошибок приводится к среднему количеству за один год-сезон.

$$FP_{\text{ср.тест}} = \frac{FP_{\text{тест}}}{Q_{\text{тест}}} \quad (7)$$

$$FP_{\text{ср.обуч}} = \frac{FP_{\text{обуч}}}{Q_{\text{обуч}}} \quad (8)$$

$$FN_{\text{ср.тест}} = \frac{FN_{\text{тест}}}{Q_{\text{тест}}} \quad (9)$$

$$FN_{\text{ср.обуч}} = \frac{FN_{\text{обуч}}}{Q_{\text{обуч}}}, \quad (10)$$

где $FP_{\text{тест}}$, $FP_{\text{обуч}}$ — общее количество ложноположительных результатов для тестовой и обучающей выборки соответственно,

$FN_{\text{тест}}$, $FN_{\text{обуч}}$ — общее количество ложнонегативных результатов для тестовой и обучающей выборки соответственно,

$Q_{\text{тест}}$ — количество лет в тесте,

$Q_{\text{обуч}}$ — количество лет в обучающей выборке.

3. Для расчета средней общей ошибки по обеим группам производится сравнение ошибок по обучающей и тестовой группе.

Общая ошибка модели считается равной ошибке по тестовой группе, если ошибка на тестовой группе больше или равна ошибке на обучающей группе. Считается, что произошло переобучение модели, поэтому оценка качества по обучающей группе не учитывается.

Общая ошибка модели равна средней ошибке по обеим группам (тестовой и общей), если ошибка тестовой группы меньше ошибки на обучающей группе. Это возможно в случае, если в обучающую

группу попали сложно классифицируемые случаи, а в тестовую группу — простые случаи. То есть требуется учет ошибок в обучающей группе.

$$FP_{\text{ср}} = \begin{cases} FP_{\text{тест}}, & FP_{\text{ср.тест}} \geq FP_{\text{ср.обуч}} \\ \frac{FP_{\text{тест}} + FP_{\text{обуч}}}{Q_{\text{тест}} + Q_{\text{обуч}}}, & FP_{\text{ср.тест}} < FP_{\text{ср.обуч}} \end{cases} \quad (11)$$

$$FN_{\text{ср}} = \begin{cases} FN_{\text{тест}}, & FN_{\text{ср.тест}} \geq FN_{\text{ср.обуч}} \\ \frac{FN_{\text{тест}} + FN_{\text{обуч}}}{Q_{\text{тест}} + Q_{\text{обуч}}}, & FN_{\text{ср.тест}} < FN_{\text{ср.обуч}} \end{cases} \quad (12)$$

4-1. В случае сравнения и выбора модели в процессе обучения будет выбираться модель с минимальной общей ошибкой: $FP_{\text{ср}} + FN_{\text{ср}}$. При равных величинах общей ошибки выбирается та, у которой меньше ошибка $FN_{\text{ср}}$. Если они тоже равны, выбирается та модель, у которой наименьшая логистическая функция ошибки.

4-2. В случае сравнения и выбора среди уже обученных моделей выбираться будет с наименьшей $FN_{\text{ср}}$. При равных значениях выбирается с наименьшей $FP_{\text{ср}}$. Если они тоже равны, выбирается та модель, у которой наименьшая логистическая функция ошибки.

Стоит отметить, что подход, описанный в пункте 3, применим и для других оценочных метрик: точность, полнота, F_1 мера, логистическая функция ошибки, расчет площади под ROC и PR. Но только при условии постоянства объема выборок, участвующих в обучении модели.

Параметры модели прогнозирования окончания периода заморозков

Выходным параметром модели является случай ОПЗ. Обучающая выборка размечается «0» и «1», где «0» — это признак завершения периода заморозков с этого дня, а «1» — заморозки еще будут.

Входными данными для модели прогнозирования ОПЗ являются временные ряды с различными метеопараметрами. Для данного исследования были выбраны следующие метеопараметры: температура, давление, относительная влажность, скорость и направление ветра, точка росы, процент облачности. Эти метеопараметры были выбраны как наиболее коррелирующие с определением ОПЗ в ходе обзора исследований на данную тему [7, 8]. В качестве длины временных окон были взяты 30, 45, 60 дней. То есть будут сравниваться модели с разным количеством входных параметров, как по длине временного окна, так и по конкретным параметрам.

Имеющийся архив в 20 лет был поделен на три части: обучающая (12 лет), тестовая (6 лет), валидационная (2 года) [9]. Разделение осуществлялось строго по годам, так как при бесконтрольном перемешивании будет происходить утечка данных из одной группы в другую, что может изменить конечные результаты.

В качестве модели прогнозирования используется искусственная нейронная сеть (ИНС) с двумя скрытыми слоями. Это продолжение исследования [7], в котором решалась аналогичная задача, только с помощью методов линейной регрессии. Поиск нелинейных зависимостей с помощью ИНС должен привести к построению более точной модели прогнозирования. Два скрытых слоя в ИНС выбраны как стартовая точка, с которой будет проводиться исследование. Архитектура ИНС и параметры ее обучения здесь не являются определяющими, так как требуется, в первую очередь, разработать и оценить методику, по которой будут сравниваться получаемые модели. Уже после выбора оценочных критериев будет производиться поиск оптимальной архитектуры ИНС. Для оценки метрик качества вводится валидационная часть выборки.

Результаты

Процесс проверки методики оценки качества моделей был разделен на два этапа.

На первом этапе оценивалось применение методики в процессе обучения модели. Модель обучалась на входных данных: температура, давление, влажность воздуха, направление ветра и значение точки росы за последние 60 дней. На этом этапе вполне можно пользоваться стандартными метриками качества для выбора оптимальной модели в процессе обучения. Но, как показывают эксперименты, разработанная методика уверенно выбирает оптимальную модель из возможных (табл. 2). В таблице выделены оптимальные значения для конкретной метрики и можно оценить, какие значения при

Таблица 2

Выборка моделей с оптимальными значениями метрик качества

Метрики качества	Варианты моделей прогнозирования ОПЗ					
	1	2	3	4	5	6
FN(L+T)	3.833	0.000	0.000	0.000	0.000	0.000
FP(L+T)	2.833	7.167	7.500	8.333	8.333	8.667
FP+FN(L+T)	6.667	7.167	7.500	8.333	8.333	8.667
FN(L+T+V)	3.222	0.333	0.444	0.000	0.000	0.000
FP(L+T+V)	2.556	8.111	8.333	9.333	9.333	9.556
FP+FN(L+T+V)	5.778	8.444	8.778	9.333	9.333	9.556
AUC_ROC(L+T)	0.992	0.991	0.991	0.997	0.997	0.997
AUC_ROC(L+T+V)	0.994	0.992	0.992	0.997	0.997	0.997
AUC_PR(L+T)	0.984	0.980	0.979	0.994	0.994	0.994
AUC_PR(L+T+V)	0.988	0.982	0.982	0.994	0.994	0.994
log_loss(L+T)	0.365	0.132	0.132	0.175	0.176	0.194
log_loss(L+T+V)	0.369	0.137	0.137	0.182	0.183	0.201
f_one(L+T)	0.932	0.932	0.929	0.922	0.922	0.919
f_one(L+T+V)	0.942	0.922	0.919	0.915	0.915	0.913
mcc(L+T)	0.900	0.901	0.897	0.887	0.887	0.883
mcc(L+T+V)	0.914	0.885	0.880	0.876	0.876	0.873

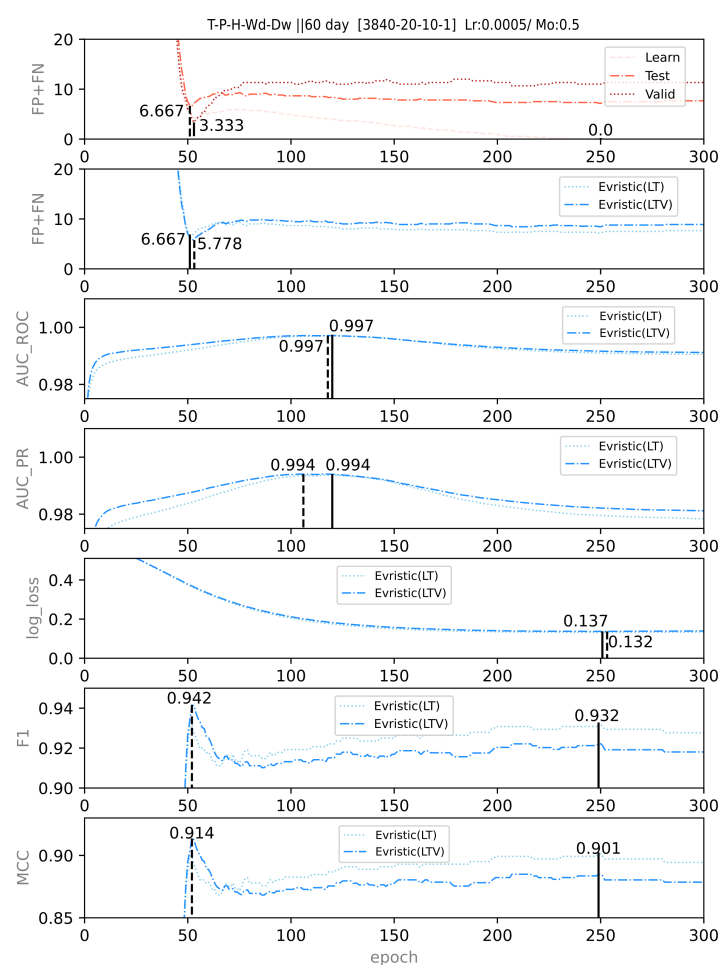


Рис. 2. Пример сравнения изменения значения метрик в процессе обучения модели

этом принимают другие метрики. На рисунке 2 представлены графики изменения значений различных метрик качества от эпох обучения. Важно отметить, что метрики MCC , F_1 , $FP_{cp} + FN_{cp}$ обладают схожим поведением и их оптимальные значения приходятся на модель со схожими коэффициентами и характеристиками. При этом оптимальные значения AUC_ROC и AUC_PR приходятся на модель с более высоким средним количеством ошибок, но нулевой ложноотрицательной ошибкой (табл. 2). То есть данные метрики также пригодны, если требуется снижение FN ошибки на стадии обучения. Однако причиной занижения FN является дисбаланс выборки в сторону негативного класса. Поэтому, если требуется снижение ложноотрицательной ошибки на стадии обучения, правильное будет использовать разработанную методику, минимизируя FN . Опять же это стоит делать внимательно, так как минимизация ошибки второго рода на этом этапе может сильно ухудшить общую ошибку.

На втором этапе производилось сравнение различных моделей с помощью разработанной методики (рис. 3). Были обучены модели с различным диапазоном временного окна: 15, 30, 45, 60 дней и различным набором входных данных: Т — температура, Р — давление, Н — относительная влажность воздуха, Ws — скорость ветра, Wd — направление ветра, Dw — точка росы, Cp — процент облачности, DTU — коэффициент дня по году. Значение Evristic(LT) — это оценка модели по разработанной методике на основе тестовой и обучающей выборки. Значение Evristic(LTV) — это оценка с учетом валидационной выборки. Из графика видно (рис. 3), что если мы обучим модель ИНС на основе данных температуры, давления, влажности воздуха и скорости ветра за 15 дней, то получим модель со средней общей ошибкой 9 случаев за один год. Но если использовать данные за 30 дней и вместо скорости ветра использовать направление ветра, то модель получится со средней ошибкой за один год, равной 5,2.

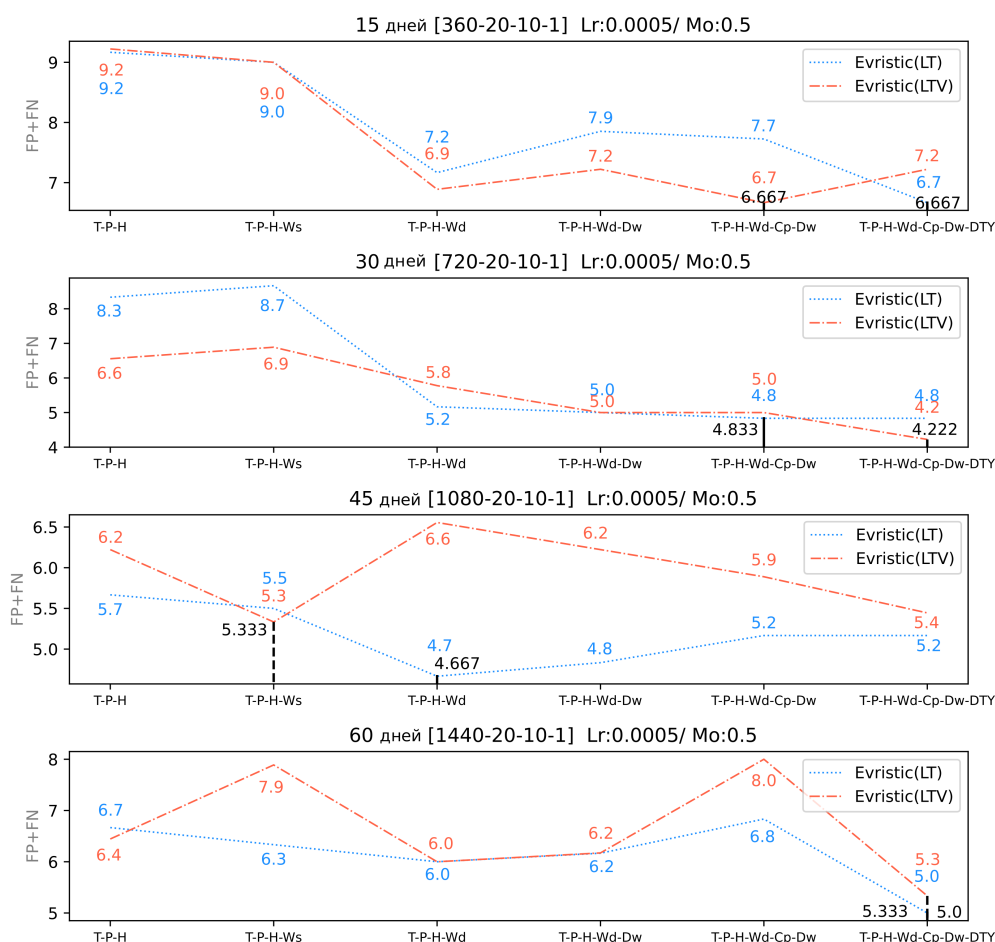


Рис. 3. Сравнение моделей, полученных с помощью разного набора входных данных

Отметим, что при 15- и 30-дневных временных окнах увеличение числа параметров приводит к снижению ошибки, что логично. Но при 45 и 60 днях это уже не наблюдается. Также можно

отметить, что в среднем ошибка увеличивается по сравнению с 30-дневным временным окном, хотя должно происходить иначе. Это можно объяснить тем, что увеличение временного окна приводит к быстрому переобучению, поэтому требуется изменение архитектуры ИНС и параметров обучения. В дальнейших исследованиях можно провести поиск оптимальной архитектуры и гиперпараметров, при которых 60-дневное временное окно будет увеличивать точность, а не снижать ее. Также стоит отметить схожесть поведения значений *Evristic(LT)* и *Evristic(LTV)*, что говорит о сравнимом результате методики, основанной на тестовой и обучающей выборках и основанной на валидационной выборке.

Заключение

Полученная методика оценки и выбора оптимальной модели окончания периода заморозков позволяет сравнивать модели на основе ИНС с различной архитектурой и различным набором входных данных. Стоит отметить, что данный способ оценки может быть использован и для сравнения моделей, полученных различными методами машинного обучения.

Отметим следующие важные моменты разработанной методики:

- наглядность и интерпретируемость метрик;
- возможность сравнения различных видов моделей;
- возможность сравнения влияния входных параметров на прогнозируемое явление;
- возможность сравнения влияния параметров и методов оптимизации обучения на качество получаемых моделей;
- учитывается возможность переобучения модели.

Разработанная методика оценки качества в дальнейшем будет использоваться для поиска оптимального набора входных данных и получения модели с высокой прогнозирующей способностью.

ЛИТЕРАТУРА

1. Hossin M., Sulaiman M. N. A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*. 2015;5(2):1–11. DOI: 10.5121/ijdkp.2015.5201.
2. Huang C. et al. Performance Metrics for the Comparative Analysis of Clinical Risk Prediction Models Employing Machine Learning. *Circulation: Cardiovascular Quality and Outcomes*. 2021;14(10):1076–1086. DOI: 10.1161/CIRCOUTCOMES.120.007526.
3. Vujović Ž. et al. Classification Model Evaluation Metrics. *International Journal of Advanced Computer Science and Applications*. 2021;12(6):599–606.
4. Архипов В. А. Сравнительный анализ метрик качества для моделей бинарной классификации на примере кредитного скоринга. *Вестник Алтайской академии экономики и права*. 2019;9(2):12–15. EDN: CFZQDC.
5. Hicks S. A. et al. On Evaluation Metrics for Medical Applications of Artificial Intelligence. *Scientific Reports*. 2022;12(1):5979. DOI: 10.1038/s41598-022-09954-8.
6. Owusu-Adjei M. et al. Imbalanced Class Distribution and Performance Evaluation Metrics: A Systematic Review of Prediction Accuracy for Determining Model Performance in Healthcare Systems. *PLOS Digital Health*. 2023;2(11):E0000290. DOI: 10.1371/journal.pdig.0000290.
7. Солозобов В. А. Поиск оптимальной регрессионной модели, предсказывающей завершение периода возвратных весенних заморозков. *Наука и инновации XXI века : материалы X Всероссийской конференции молодых ученых*. Сургут, 15 декабря 2023 года. В 3 ч. 2024;1:98–102. EDN: BHSRWN.
8. Солозобов В. А. Исследование возможности использования нейронных сетей для предсказания возвратных весенних заморозков, и поиск оптимального набора данных для этого вида прогнозирования. *Наука и инновации XXI века : сборник статей по материалам IX Всероссийской конференции молодых ученых*. В 4 томах. Сургут, 02 ноября 2022 года. 2023;1:150–155. EDN: ZXQFJS.
9. *Архив метеоданных*. Режим доступа: <https://tp5.ru/>.

ОБОГАЩЕНИЕ БАЗЫ ЗНАНИЙ С ПОМОЩЬЮ АВТОМАТИЧЕСКОГО ИЗВЛЕЧЕНИЯ КЛЮЧЕВЫХ СЛОВ И АННОТАЦИЙ

Э. Г. Тунян^{1,2,3a}, Р. С. Сазиков^{1,2,3,b}, С. А. Харламов^{1,2c}

¹ Сургутский государственный университет, г. Сургут, Российская Федерация

² ООО «ЕДРО», г. Сургут, Российская Федерация

³ Сургутский филиал федерального государственного автономного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Национального исследовательского центра «Курчатовский институт», г. Сургут, Российская Федерация

^a ORCID: <https://orcid.org/0009-0003-3260-1310>, tunyan@edro.su

^b ORCID: <https://orcid.org/0009-0005-0078-0013>, sazikov@edro.su

^c ORCID: <https://orcid.org/0009-0000-5605-0531>, harlamov_sa@surgu.ru

Аннотация: в работе рассматриваются разные способы автоматического извлечения терминов: от давно известных и технологически несложных — вроде TF-IDF, RAKE или TextRank — до более современных решений, основанных на трансформерах, включая KeyBERT и модели типа LLM. Применение методов выделения ключевых слов показало, что именно гибридные схемы проявляют себя наиболее эффективно: когда статистика и нейросетевые модели работают не по отдельности, а в связке, удается добиться как формальной релевантности, так и смысловой глубины в отборе терминов. Предложен поэтапный подход: сначала — чистка и разметка текста, потом — параллельный запуск нескольких алгоритмов, позволяющий свести к минимуму случайные отклонения, и уже затем — более тонкое ранжирование на основе нейросетевых моделей. Такой алгоритм, основанный на комплексном подходе к автоматическому построению баз знаний, позволяет существенно улучшить качество автоматического выделения релевантных терминов и значительно повысить как точность, так и полезность извлекаемой информации. Автоматизация анализа больших текстовых массивов без необходимости предварительной разметки и обучения под конкретную предметную область дает возможность быстро масштабировать подход как для исследовательских задач, так и для прикладных решений в сфере цифровизации, например, в медицине, образовании, документообороте.

Ключевые слова: обогащение баз знаний, извлечение терминов, ключевые слова, аннотация, TF-IDF, RAKE, TextRank, KeyBERT, LLM, интеллектуальный поиск, автоматизация обработки текста, семантический анализ.

Для цитирования: Тунян Э. Г., Сазиков Р. С., Харламов С. А. Обогащение базы знаний с помощью автоматического извлечения ключевых слов и аннотаций. *Успехи кибернетики*. 2025;6(2):108–113.

Поступила в редакцию: 01.06.2025.

В окончательном варианте: 16.06.2025.

AUTOMATIC EXTRACTION OF KEYWORDS AND SUMMARIES FOR KNOWLEDGE BASE POPULATION

E. G. Tunyan^{1,2,3,a}, R. S. Sazikov^{1,2,3,b}, S. A. Kharlamov^{1,2,c}

¹ Surgut State University, Surgut, Russian Federation

² LLC Edro, Surgut, Russian Federation

³ Surgut Branch of Scientific Research Institute for System Analysis of the National Research Centre “Kurchatov Institute”, Surgut, Russian Federation

^a ORCID: <https://orcid.org/0009-0003-3260-1310>, tunyan@edro.su

^b ORCID: <https://orcid.org/0009-0005-0078-0013>, sazikov@edro.su

^c ORCID: <https://orcid.org/0009-0000-5605-0531>, harlamov_sa@surgu.ru

Abstract: we studied a range of automatic term extraction methods, from well-established techniques such as TF-IDF, RAKE, and TextRank to recent transformer-based approaches, including KeyBERT and large language models (LLMs). Our findings show that hybrid approaches — combining statistical and neural methods — achieve superior performance by ensuring both formal relevance and semantic depth in term selection. We proposed a multi-stage pipeline: initial text preprocessing and annotation, followed by the parallel application of several extraction algorithms to minimize stochastic variation, and final refinement through neural network-based ranking. This integrated algorithm significantly improves the quality of term extraction and enhances both the accuracy and practical utility of the retrieved information. The proposed

method enables scalable analysis of large unstructured text corpora without the need for manual annotation or domain-specific training, making it suitable for a wide range of research and applied digitalization tasks, including applications in medicine, education, and document management.

Keywords: knowledge enrichment, term extraction, keywords, abstract, TF-IDF, RAKE, TextRank, KeyBERT, LLM, intelligent search, automated text processing, semantic analysis.

Cite this article: Tunyan E. G., Sazikov R. S., Kharlamov S. A. Automatic Extraction of Keywords and Summaries for Knowledge Base Population. *Russian Journal of Cybernetics*. 2025;6(2):108–113.

Original article submitted: 01.06.2025.

Revision submitted: 16.06.2025.

Введение

В современном мире технологии искусственного интеллекта (ИИ) стремительно развиваются и интегрируются во все сферы деятельности человека. Развитие будет набирать еще большие обороты, что подтверждается объемом инвестиций. Инвестиции США, Европы и Израиля в 2024 году на развитие ИИ и облачных технологий увеличились на 27% — с 62,5 млрд до 79,2 млрд долларов. В 2025 году OpenAI, Oracle и SoftBank инвестировали 500 млрд долларов, Национальный фонд Китая — 138,01 млрд долларов.

С развитием ИИ векторные базы данных требуют эффективных методов структурирования и поиска информации. В качестве метаданных документа служат аннотации и ключевые слова. На фоне экспоненциального роста текстовых данных, лишенных четкой структуры, что, впрочем, закономерно в условиях ускоряющейся цифровизации и активного внедрения ИИ-инструментов, постепенно обостряется нехватка средств, способных не просто обрабатывать, но и вычленять из потока существенные смысловые фрагменты. В частности, термины и аннотации, без которых ни поиск, ни генерация, ни построение когнитивных карт не могут считаться полноценными.

Поскольку ручное выделение ключевых слов и индексирование является слишком трудоемким и времязатратным, используется автоматическое извлечение из текста слов и фраз, наиболее подходящих по смыслу (Automatic Keyword Extraction). Подобные данные могут применяться для обогащения баз знаний, к примеру, в качестве тегов для документов или атрибутов для баз знаний.

Классические подходы выделения ключевых фраз используют принцип статистического анализа, то есть берут термины, которые чаще всего встречаются. Принцип Term Frequency — Inverse Document Frequency (TF-IDF) дает каждому термину свой вес пропорционально его частоте в документе и обратно пропорционально его встречаемости в совокупности текстов. Этот метод впервые был описан в работе Karen Spärck Jones в 1972 году [1], которая стала основой для множества поисковых систем.

В последнее время было разработано множество эффективных методов выделения основных фраз без учителя (unsupervised), такими алгоритмами являются Rapid Automatic Keyword Extraction (RAKE) и TextRank. Алгоритм RAKE, представленный S. Rose с соавторами в 2010 году [2], разделяет весь текст на слова, при этом отбрасывает все стоп-слова, выделяет всех кандидатов, строит граф связей между словами, определяет важность, основываясь на частоте встречаемости слова. Алгоритм TextRank был предложен в 2004 году Mihalcea R. и Tarau P. [3]. Это графовый алгоритм, который основывается на алгоритме ссылочного ранжирования PageRank, используемом в том числе Google. TextRank строит граф слов или целых предложений, где вершины — это слова или фразы, а в качестве ребер выступают связи этих слов и фраз. После этого вычисляются веса вершин по принципу: наиболее важные узлы считаются словами или предложениями.

В последние годы модели на основе трансформеров (BERT) и больших языковых моделей (LLM) начали применять для задач выделения ключевых слов, так как они извлекают более осмысленные фразы. К примеру, метод KeyBERT использует эмбединги BERT для формирования словосочетаний по смыслу. Также генеративные модели, такие как GPT4, DeepSeek, могут выделять ключевые слова без дополнительного обучения (zero-shot), просто получив запрос от пользователя на естественном языке. Кроме того, такие модели могут генерировать аннотации, что также полезно для построения баз знаний.

Цель текущего исследования — сравнить разные подходы к автоматическому выделению ключевых слов и оценить их эффективность для построения баз знаний.

Материалы и методы

В данном разделе описаны основные алгоритмы и методы автоматического выделения ключевых слов, фраз, аннотаций из документов для формирования баз: статические (TF-IDF), графовые (TextRank), RAKE — метод на основе правил, с использованием трансформеров (BERT) и методов применения больших языковых моделей (LLM).

TF-IDF — один из самых первых и широко используемых методов для извлечения главных терминов из документа. Задача — оценить значимость слова на основе того, как часто встречается слово в документе (TF) и обратно пропорционально его встрече в совокупности текстов (IDF).

Формула веса:

$$TFIDF(t, d, D) = TF(t, d) \cdot IDF(t, D), \quad (1)$$

где:

t — слово,

d — конкретный документ,

D — совокупность документов,

TF — частота термина в документе,

IDF — обратная частота документа.

Формула частоты термина в документе:

$$TF(t, d) = \frac{\hat{f}_{t,d}}{\sum_{t' \in d} \hat{f}_{t',d}}, \quad (2)$$

где:

$\hat{f}_{t,d}$ — количество вхождений термина t в документ d ,

$\sum_{t' \in d} \hat{f}_{t',d}$ — общее количество слов в документе.

Обратная частота документа:

$$IDF(t, D) = \log \left(\frac{|D|}{1 + |d \in D : t \in d|} \right), \quad (3)$$

где:

$|D|$ — общее количество документов в корпусе,

$|d \in D : t \in d|$ — количество документов, содержащих термин t ,

1 нужно для того, чтобы избавиться от деления на 0.

К примеру, для научной статьи TF-IDF выделит термины, относящиеся к определенной области науки, игнорируя простые слова. Несмотря на относительную простоту реализации, TF-IDF остается одним из эффективных методов быстрой индексации текстов [4], но, так как не учитывается порядок слов, ключевые слова могут быть пропущены, также не учитывается семантика и контекст употребления слова, что негативно влияет на качество результата.

Алгоритм RAKE создан для быстрого автоматического, то есть без предварительной обработки текста, выделения ключевых слов из конкретного текста. RAKE использует тот факт, что ключевые слова, в основном, это существительные, прилагательные и редко являются стоп-словами. По этой причине текст разбивается на слова с использованием знаков препинания и стоп-слов как разделителей. Полученные слова выступают в качестве кандидатов и оцениваются по алгоритму, учитывающему частоту слов. В первоначальной реализации каждому уникальному слову задается оценка, которая равна степени, деленной на частоту. Степень слова — это количество других слов, с которыми идет сравнение. Частота — это общее количество, сколько раз встречается кандидат в тексте. Затем оценка фразы считается суммой оценок слов, которые в нее входят.

RAKE обладает рядом преимуществ, одним из которых является его непривязанность к конкретной предметной области. Не требуется заранее подготовленный текст для обучения, и не нужны внешние словари, легко настраивается под предметную область, обладает высокой скоростью вычисления, которая растет пропорционально размеру текста и равна $O(n)$.

Алгоритм TextRank — это графовый алгоритм, созданный для ранжирования важности отрывков текста, предложенный исследователями Mihalcea R. и Tarau P. в 2004 году. Он применяется как для сортировки текста, так и для извлечения ключевых слов или фраз. Для выделения фраз берется k слов

с самыми высокими весами, после чего по ним, находя в тексте смежные слова, восстанавливаются фразы. Разработчики показали, что TextRank выдавал результат, сопоставимый с лучшими на то время системами на таких тестах, как Inspec и Semeval.

Основным преимуществом TextRank является учет структуры текста. Если два слова часто встречаются рядом или связаны через стоп-слова, то алгоритм рассматривает это как усиление связи. Так находятся качественные словосочетания, даже если слова часто встречаются. Если фраза состоит из двух слов, то оба слова могут выйти в топ. Чтобы решить эту проблему, используется постобработка и группировка смежных слов. Несмотря на имеющиеся недостатки, такие как чувствительность к параметрам, высокая трудоемкость, TextRank стал основой для многих алгоритмов.

Контекстуальные методы, основанные на BERT, и появление его модификаций дали новые возможности для извлечения ключевых слов с учетом контекста. Одним из самых эффективных инструментов является KeyBERT, который был придуман M. Grootendorst в 2020 году [5]. Модель создает вектор всего документа при помощи BERT, а далее находит слова и словосочетания, векторы которых максимально близки к эмбедингу документа. В отличие от TF-IDF, BERT находит слова, наиболее точно отражающие смысл, учитывает синонимы и контекст поиска. KeyBERT, кроме того, позволяет записывать выражения разными словами, задавать желаемую длину фраз и применять механизм Maximal Marginal Relevance (MMR) для повышения разнообразия ключевых слов, чтобы они не были слишком похожи друг на друга. Эксперименты показывают, что даже без обучения на заранее подготовленных данных метод BERT обходит традиционные алгоритмы.

Методы, основанные на LLM, такие как GPT, DeepSeek, Mistral, используют самые последние достижения для работы с текстом. Данные модели показали возможность решать новые задачи без предварительного обучения и дообучения, применяя текстовый запрос (prompt). За счет ответов на естественном языке на уровне специализированных моделей LLM дают возможность выделять ключевые слова достаточно качественно. В отличие от предыдущих алгоритмов, LLM не опираются просто на правила, а воспринимают содержимое по смыслу, даже если тема указана неявно. Основным недостатком данных моделей является то, что они работают по принципу черного ящика, то есть ответ могут дополнить лишними словами и фразами. У каждой модели есть свои особенности, и в совокупности они дают максимальную гибкость, могут выделять ключевые термины, писать аннотации, даже делать поиск по заданному тексту, с дополнением и обогащением баз знаний при помощи больших языковых моделей, могут служить универсальным интерфейсом для необработанного текста, из которого будет получена структурированная база знаний, однако имеется проблема с оценкой результата этого метода из-за возможных галлюцинаций.

Результаты и их обсуждение

Многоэтапный подход для обогащения базы знаний, который при помощи LLM автоматически извлекает ключевые слова с последующей аннотацией, будет показывать хорошие результаты. Идея в том, чтобы не просто выделять из документа термины, а пополнять их информацией, которая связана по смыслу, и затем добавлять эти данные в базу. Такой подход значительно улучшит поиск по базе знаний. Предлагаемый метод (рис.) представляет собой многоэтапный алгоритмический конвейер, основанный на интеграции классических статистических подходов с современными нейросетевыми технологиями.

На первом этапе исходный текст подвергается глубокой предварительной обработке. В рамках данной стадии осуществляется очистка документа от лишних символов и HTML-тегов, разбиение текста на предложения, токенизация, лемматизация (приведение слов к их базовой форме) и выделение именованных сущностей (NER). Такой комплексный анализ обеспечивает создание корректного представления исходных данных для дальнейшей обработки.

Далее, используя обработанный текст, параллельно запускаются несколько алгоритмов для выделения кандидатов в ключевые слова. Среди них классический метод RAKE, основанный на анализе стоп-слов и частотном анализе, графовый алгоритм TextRank, применяющий принципы ранжирования по аналогии с PageRank, а также современные подходы на базе трансформеров — KeyBERT и LLM-модуль, способный в режиме zero-shot извлекать ключевые термины посредством генеративных моделей. Полученные списки кандидатов объединяются с последующим удалением дублирующих элементов и уточняются посредством финального ранжирования, выполняемого с помощью LLM, что позволяет оценить релевантность каждого термина с учетом контекстуальных связей.

Заключительным этапом является автоматическая генерация аннотаций для выделенных ключевых слов. Для каждого термина формируется краткое определение, описание и при необходимости указываются ссылки на внешние источники или соответствующие элементы онтологии. Полученные метаданные интегрируются в базу знаний, что не только обеспечивает ускоренное индексирование, но и повышает семантическую связанность информационных систем. Примечательно, что этап аннотирования здесь рассматривается не как формальное приложение, а, скорее, как звено, связывающее смысловую обработку с онтологическим представлением.

Таким образом, представленный метод, сочетающий традиционные алгоритмы (например, TF-IDF, RAKE и TextRank) с современными нейросетевыми подходами (KeyBERT, LLM), позволяет су-

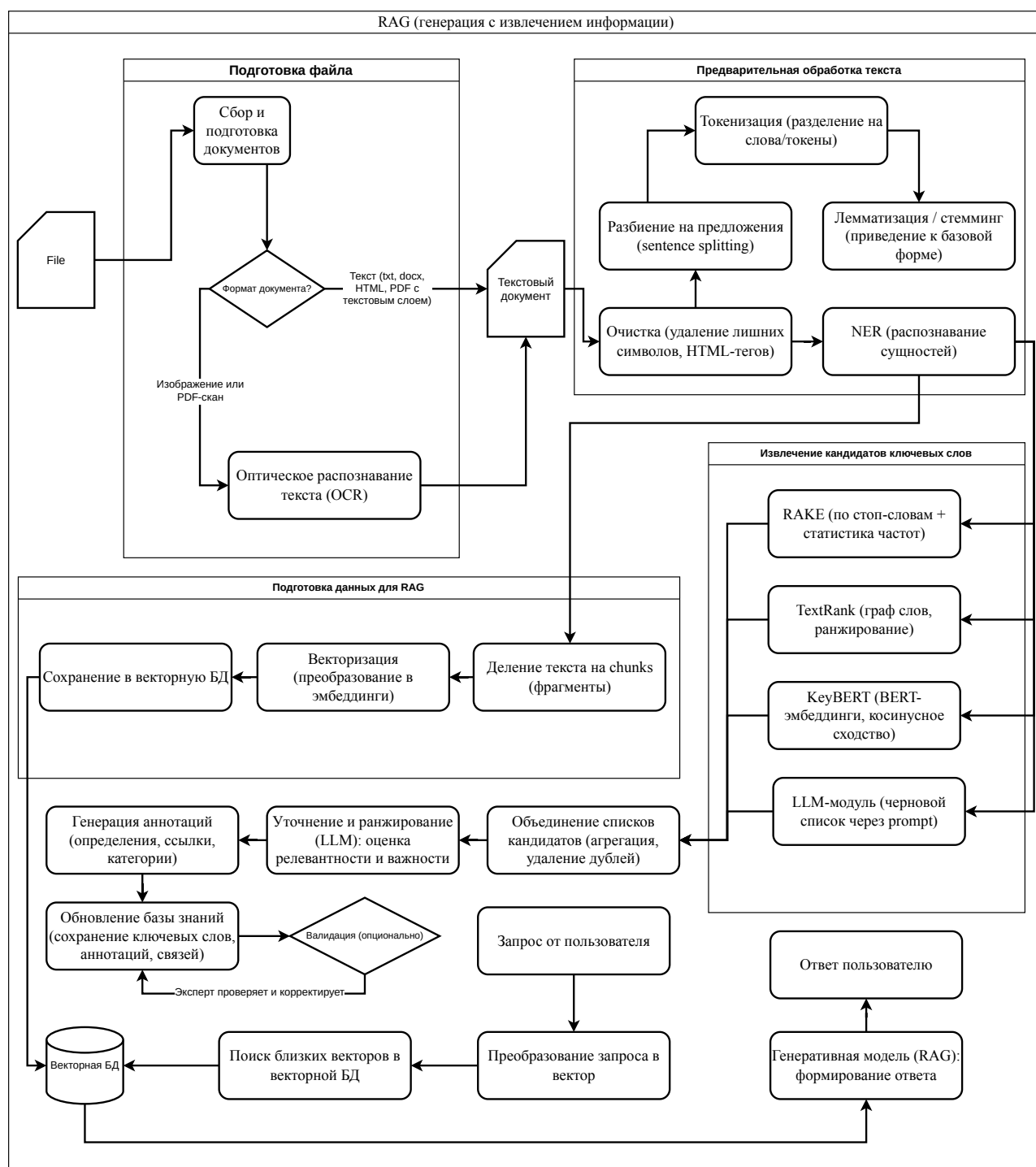


Рис. Схема комплексного автоматического выделения ключевых слов, аннотирования и поддержки RAG

щественно улучшить качество автоматического выделения релевантных терминов, повысить точность поиска и обеспечить оперативное обновление баз знаний. Если говорить о применимости, то метод оказывается достаточно гибким — особенно в тех случаях, когда требуется подстроиться под конкретную задачу или корпус. В определенных условиях — особенно при наличии предварительной адаптации под специфику задачи — метод способен обеспечить более точное смысловое считывание данных, что, в свою очередь, может оказаться полезным при разработке или эксплуатации систем, ориентированных на аналитическую поддержку принятия решений.

Заключение

Проведенное исследование подтвердило, что комплексный подход к автоматическому построению баз знаний, действительно, может значительно повысить как точность, так и полезность извлекаемой информации. Применение методов выделения ключевых слов — от классических алгоритмов вроде TF-IDF, RAKE и TextRank до современных решений, построенных на трансформерах (KeyBERT, LLM), — показало, что именно гибридные схемы работают наиболее эффективно. Когда статистика и нейросетевые модели работают не по отдельности, а в связке, удастся добиться как формальной релевантности, так и смысловой глубины в отборе терминов.

Предложенный многоэтапный метод, в котором на каждом шаге происходит фильтрация, параллельное извлечение терминов через разные алгоритмы и финальное ранжирование с помощью LLM, позволяет получить не просто набор ключевых слов, а по-настоящему осмысленную структуру знаний. Генерация аннотаций на основе этих слов усиливает связность информации: каждый термин получает краткое пояснение, контекст, что делает такие базы особенно пригодными для задач вопрос-ответ, семантического поиска или тематической навигации.

Практика показала, что именно подключение языковых моделей — в том числе zero-shot генерации через GPT и аналоги — открывает принципиально новые горизонты. Автоматизация анализа больших текстовых массивов без необходимости предварительной разметки и обучения под конкретную предметную область дает возможность быстро масштабировать подход под разные сферы — от научных архивов до внутренних корпоративных баз.

Разработанная архитектура обладает высокой гибкостью: алгоритмы можно комбинировать, заменять, подстраивать под особенности входных данных. Она подходит как для исследовательских задач, так и для прикладных решений в сфере цифровизации, например, в медицине, образовании, документообороте. В условиях растущих объемов данных и быстрого внедрения ИИ подобные инструменты становятся не вспомогательными, а ключевыми элементами в новых информационных экосистемах.

ЛИТЕРАТУРА

1. Sparck Jones K. A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *Journal of Documentation*. 1972;28:11–21.
2. Rose S., Engel D., Cramer N., Cowley W. Automatic Keyword Extraction from Individual Documents. *Text Mining: Applications and Theory*. 2010. DOI: 10.1002/9780470689646.ch1.
3. Mihalcea R., Tarau P. TextRank: Bringing Order into Text. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, Barcelona, Spain*. Association for Computational Linguistics; 2004:404–411.
4. Маннинг К. Д., Рагхаван П., Шютце Х. *Введение в информационный поиск*. Вильямс; 2020. 528 с. ISBN 978-5-907203-20-4.
5. Grootendorst M. *KeyBERT: Minimal Keyword Extraction with BERT*. Режим доступа: <https://www.maartengrootendorst.com/blog/keybert/>.

ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ «ОПТИМАЛЬНЫЙ УЧЕБНЫЙ ПЛАН»

М. В. Юрчишина

Сургутский государственный университет, г. Сургут, Российская Федерация

✉ nesterenko_mv@surgu.ru

Аннотация: в работе представлена функциональная модель системы поддержки принятия решений, предназначенной для формирования оптимального учебного плана в сфере высшего образования. Обоснована актуальность и значимость проектируемой системы поддержки принятия решений. Обозначены типы математических моделей и алгоритмов работы, которые будут составлять основу для решения задач принятия решений в проектируемой системе. Построение модели выполнено с использованием стандарта функционального моделирования IDEF0. В соответствии с выбранными математическими моделями и алгоритмами построена контекстная диаграмма, дающая представление о системе в целом, диаграмма декомпозиции первого уровня, отражающая модульную структуру системы и дающая представление о возможных аспектах оптимизации учебного плана, а также построены диаграммы декомпозиции второго уровня для каждого модуля системы, поясняющие функционирование системы.

Ключевые слова: оптимизация, учебный план, система поддержки принятия решений, функциональная модель, проектирование.

Для цитирования: Юрчишина М. В. Функциональная модель системы поддержки принятия решений «Оптимальный учебный план». *Успехи кибернетики*. 2025;6(2):114–125.

Поступила в редакцию: 09.04.2025.

В окончательном варианте: 17.04.2025.

FUNCTIONAL MODEL OF THE OPTIMAL CURRICULUM DECISION SUPPORT SYSTEM

M. V. Yurchishina

Surgut State University, Surgut, Russian Federation

✉ nesterenko_mv@surgu.ru

Abstract: we presented a functional model of a decision support system (DSS) designed to generate an optimal university curriculum. We demonstrated the importance of the proposed DSS. We identified the types of mathematical models and algorithms for solving decision-making problems within the system. We built the model using the IDEF0 functional modeling standard. Based on the selected models and algorithms, we developed a context diagram that outlines the system as a whole, a first-level decomposition diagram that reflects its modular structure and potential aspects of curriculum optimization, and second-level decomposition diagrams for each system module that explain its internal functioning.

Keywords: optimization, curriculum, decision support system, functional model, system design.

Cite this article: Yurchishina M. V. Functional Model of the Optimal Curriculum Decision Support System. *Russian Journal of Cybernetics*. 2025;6(2):114–125.

Original article submitted: 09.04.2025.

Revision submitted: 17.04.2025.

Введение

Проектирование учебных планов в системе высшего профессионального образования представляет собой задачу планирования, существенно зависящую от области науки, конкретного учебного заведения, текущей ситуации в стране. Например, тридцать лет назад можно было сформировать учебный план и не вносить корректировки в него несколько лет. Кроме того, стандарт, действующий в то время, содержал достаточно обширный набор составляющих самого учебного плана, в том числе более 20 названий учебных дисциплин для специальности, с четким перечнем их содержания [1]. В настоящее время ФГОС ВО содержит лишь минимальные фиксированные сведения о составе учебного плана, в основном интегрального характера [2]. Из-за отсутствия четких указаний каждое учебное заведение, реализующее образовательные программы, должно самостоятельно подбирать состав учебного плана в соответствии с ожидаемыми результатами в форме установленных компетенций. И хотя многие высшие учебные заведения в качестве основы ориентируются на те перечни, которые были изначально, трансформации за столько лет произошли очень существенные: изменились названия дисциплин,

добавились технологии, изменилась сама система высшего образования. Естественно, это привело к огромному количеству вариаций образовательных программ и учебных планов для них. При этом в силу высокой скорости изменений на рынке труда требования к выпускникам постоянно меняются, поэтому образовательные программы и, соответственно, учебные планы (УП) для них часто требуют корректировки. В настоящее время почти каждый год вуз вынужден по тем или иным причинам корректировать УП.

Формирование учебного плана зачастую производится не путем создания нового (хотя и такое возможно), а путем корректировки УП прошлого учебного периода или смежного направления/профиля. Учитывая наличие интегральных показателей, корректировки эти напоминают игру в «пятнашки». Но задача построения плана с указанием требований, очевидно, может быть поставлена математически и решаться в автоматизированном режиме. Указав все необходимые условия и ограничения, составитель УП может возложить на ЭВМ перебор всех вариантов и выбор допустимых из них, а при наличии критерия — и выбор оптимального варианта. Имеющиеся автоматизированные решения для создания УП [3] предлагают лишь оболочки для ручного ввода всех данных и автоматизированной проверки условий [4]. При этом все корректировки необходимо вводить в ручном режиме, а система лишь подсказывает, какое условие не выполнено. Всю работу по корректировке и вводу новых значений производит человек. Составитель должен понимать, как изменение повлияет на выполнение ограничений. При этом вопрос оптимизации даже не ставится — чаще всего первый же допустимый вариант является финальным.

Оптимизация при составлении УП для высшего образования в настоящее время рассматривается во многих научных работах. Имеется несколько основных подходов к построению оптимального УП. В работах [5–7] используется графовый подход, который учитывает последовательность изучения связанных дисциплин, трудоемкость дисциплин, допускает учет определенных ограничений, но также имеет ряд недостатков, как, например, сложный учет дисциплин, требующих непосредственной связи, учет современного набора универсальных дисциплин. В работах Сеньковской А.А. [8–10] используется оптимизация с точки зрения распределения нагрузки, но никак не принимается во внимание оптимизация для студентов. В работах [11–13] рассматривается подход к построению плана с оптимизацией его по нескольким критериям последовательно при планировании различных составляющих: расстановки учебных дисциплин по семестрам, распределения трудоемкости плана между дисциплинами, распределения трудоемкости дисциплины между видами деятельности студента. Такой подход позволяет не просто оптимизировать УП, а в автоматизированном режиме сформировать конечный документ для утверждения. При таком подходе составителю УП не нужно заполнять документ, корректировать его, а нужно установить все требуемые ограничения, и составляющие последовательно рассчитываются системой.

Некоторые аспекты формирования систем ограничений при построении математических моделей используют результаты, изложенные в [14–17].

Постановка задачи

Функциональное моделирование системы поддержки принятия решений (СППР) является одним из наиболее наглядных представлений системы. Проектируемая СППР, описываемая в работах [11–13], с точки зрения математического моделирования требует тщательной проработки функционального исполнения. Для построения модели выбран стандарт IDEF0, являющийся наиболее распространенным для создания функциональных моделей.

При построении функциональной модели необходимо учитывать модульный характер системы. Необходимо предусмотреть возможность пропуска работы любого из модулей. Также обязательно предусмотреть ручную корректировку результатов работы каждого модуля в соответствии с теорией принятия решений, которая предполагает ответственность лица, принимающего решение, за принятое решение.

Система должна иметь инструменты оптимизации 3 уровней: оптимизация расстановки дисциплин по учебным семестрам с учетом семантической связи, ограничений ФГОС, локальных нормативных документов, оптимизация распределения общей трудоемкости при известных значениях коэффициентов значимости дисциплин для выбранной расстановки дисциплин по учебным семестрам, оптимизация распределения академических часов при установленных значениях трудоемкости и ограничениях ФГОС и локальных нормативных актов, распределении дисциплин по учебным семестрам.

рам [12]. В данной работе в качестве параметров оптимизации выбраны коэффициенты значимости дисциплин. Такое понятие непосредственно в предметной области отсутствует, поэтому для установления значений этого коэффициента необходима подсистема, помогающая составителю установить их. Подход к построению такой подсистемы описан в [19]. Также необходим модуль формирования непосредственно документа УП по установленным требованиям.

Контекстная диаграмма

Построение контекстной диаграммы функциональной модели позволяет сформировать целостное представление о системе в целом. Данная диаграмма дает полное представление о взаимодействии пользователя с системой, результатах работы системы, а также обо всех управляющих факторах, обеспечивающих правильное выполнение функций системы (рис. 1).

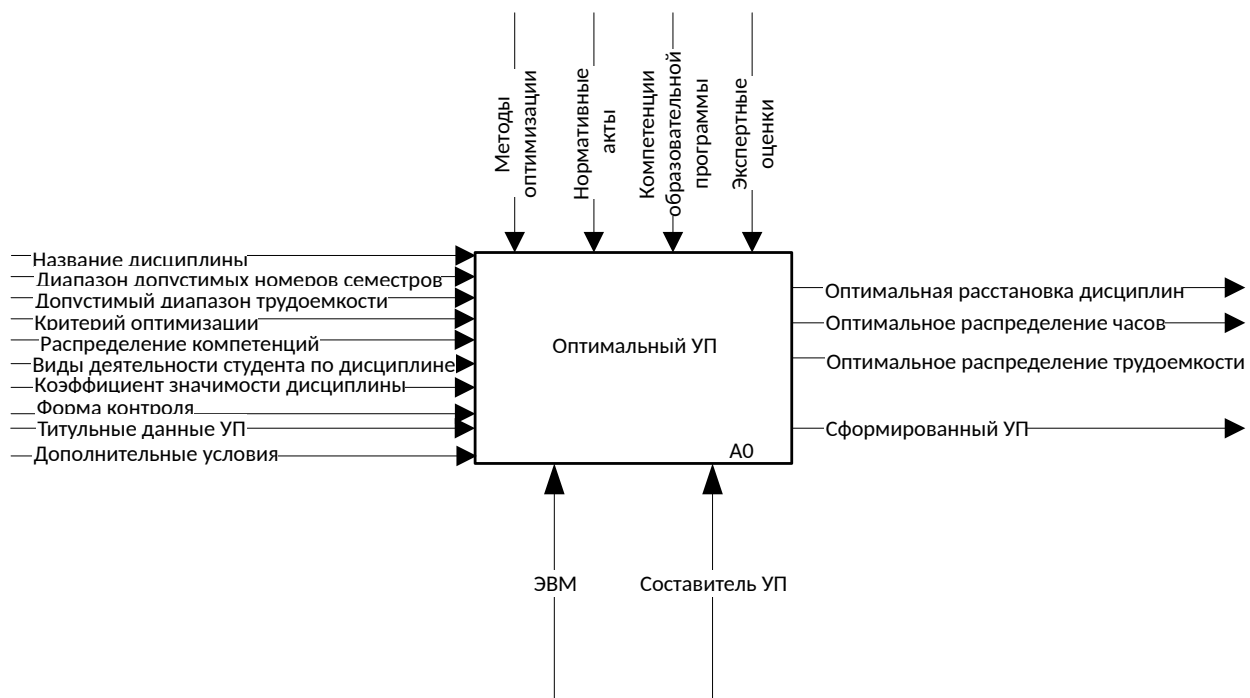


Рис. 1. Контекстная диаграмма функциональной модели СППР «Оптимальный УП»

Из рисунка 1 видно, что для оптимизации УП составителю необходимо ввести первичные сведения, которые определяют конкретную математическую постановку оптимизационной задачи, решаемой системой. Для запуска модулей системы необходимо вводить разные характеристики в различных комбинациях, общий перечень их следующий:

- список названий дисциплин;
- диапазон допустимых номеров семестров для каждой дисциплины;
- диапазон значений трудоемкости каждой дисциплины в зачетных единицах (з.е.);
- критерий оптимизации для каждой оптимизационной задачи;
- распределение компетенций по дисциплинам, формирующим их;
- перечень видов деятельности по каждой дисциплине;
- коэффициенты значимости дисциплин;
- формы контроля для каждой дисциплины;
- дополнительные условия для установления баланса часов между видами деятельности;
- титульные данные учебного плана [12].

Управляющими факторами являются:

- методы оптимизации;
- нормативные акты для УП (ФГОС, приказы, распоряжения и т.п.);
- компетенции образовательной программы;

- экспертные оценки значимости дисциплин.

Выходными данными являются:

- оптимальная по выбранному критерию расстановка дисциплин по семестрам;
- оптимальное распределение трудоемкости по дисциплинам;
- оптимальное распределение часов на виды деятельности при фиксированных трудоемкостях;
- документ в формате xls, содержащий учебный план.

Механизмами являются ЭВМ и составитель учебного плана.

Данная диаграмма наглядно отражает основные аспекты работы с системой при составлении оптимального учебного плана.

Диаграммы декомпозиции

Структура системы «Оптимальный УП» является модульной. Имеются четыре модуля (рис.2):

- оптимизации расстановки дисциплин «Расстановка»;
- оптимизации распределения трудоемкости «Трудоемкость»;
- оптимизации распределения часов «Часы»;
- формирования УП «Документ».

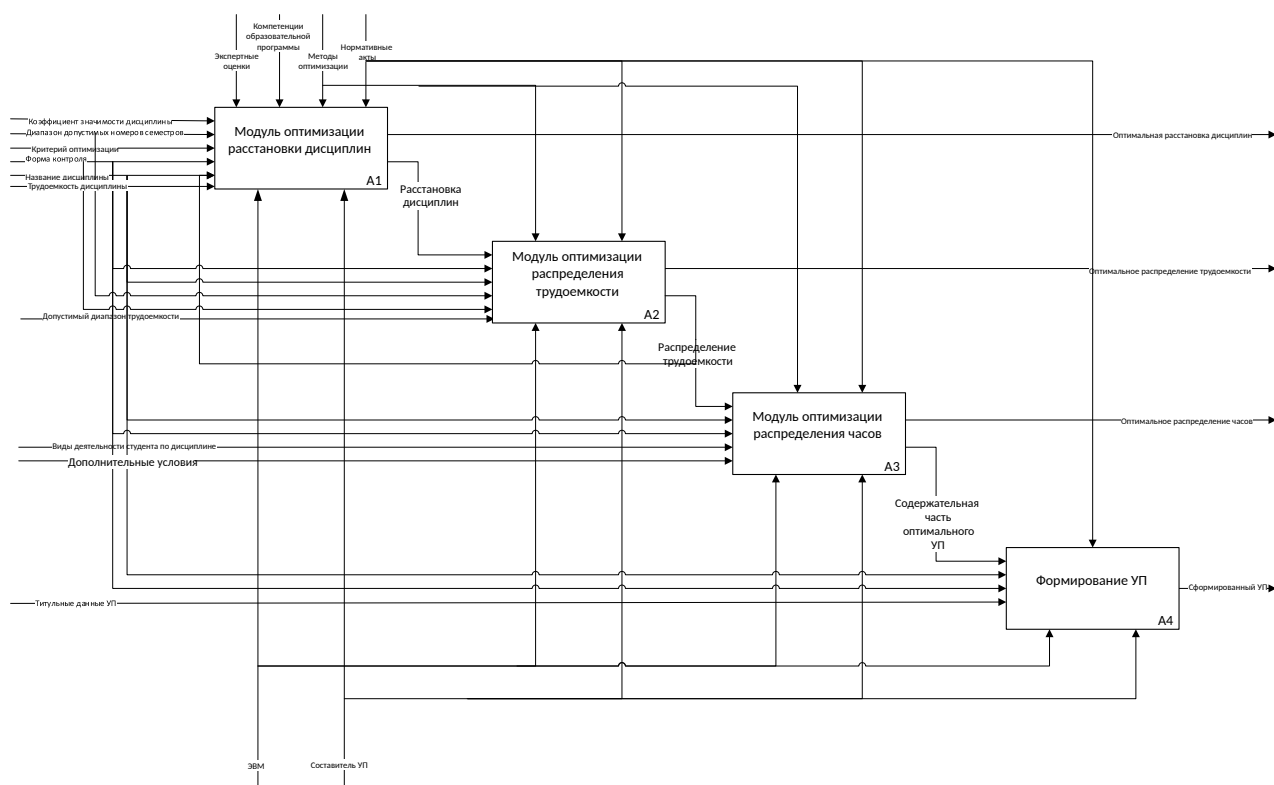


Рис. 2. Диаграмма декомпозиции функциональной модели СППР «Оптимальный УП»

Модуль оптимизации расстановки дисциплин «Расстановка» A1 предназначен для определения наилучшего варианта по выбранному критерию распределения всех дисциплин по семестрам. Входными данными данный модуль требует:

- перечень дисциплин;
- допустимые значения номеров семестров для каждой дисциплины;
- формы контроля всех дисциплин;
- критерий оптимизации (проектируемая система будет иметь несколько вариантов критерия в данной задаче);
- коэффициенты значимости дисциплин (весовые коэффициенты, являющиеся параметрами оптимизационной задачи);
- трудоемкости всех дисциплин (здесь необходимо установить минимальные или рекомендуемые значения для каждой дисциплины или распределение трудоемкости может быть получено из модуля

оптимизации трудоемкости) [12].

Управляющими факторами являются:

- метод оптимизации [11];
- нормативные документы (ФГОС, распоряжения, приказы и т.п.);
- компетенции образовательной программы;
- экспертные оценки (необходимы для помощи составителю в подборе значений коэффициентов значимости дисциплин) [12].

С помощью составителя и ЭВМ в качестве механизмов данный модуль рассчитывает оптимальное решение и выводит расстановку дисциплин по семестрам.

Модуль оптимизации распределения трудоемкости «Трудоемкость» А2 предназначен для подбора оптимальных значений трудоемкости для каждой дисциплины из набора УП. Решение о распределении трудоемкости находится при фиксированной расстановке дисциплин по семестрам. Предполагается, что значение трудоемкости может допускать вариативность. Для тех дисциплин, которые не допускают, достаточно установить одинаковые наибольшее и наименьшее значения.

Входные данные для модуля А2:

- список дисциплин;
- распределение дисциплин по семестрам;
- допустимый диапазон трудоемкости каждой дисциплины;
- формы контроля для каждой дисциплины;
- коэффициенты значимости дисциплин;
- критерий оптимизации.

Управление осуществляется выбранным методом оптимизации и нормативными актами. Механизмы выполнения функции данного блока — ЭВМ и составитель УП.

Результатом работы данного блока — выходными данными является оптимальное распределение общей трудоемкости УП по всем дисциплинам.

Модуль оптимизации распределения трудоемкости дисциплины по видам деятельности «Часы» А3 предназначен для подбора оптимальных значений количества часов на лекции, практические занятия, лабораторные работы, самостоятельную работу для каждой дисциплины в рамках фиксированной расстановки дисциплин по семестрам и установленных значений трудоемкости в з.е. [12]. Несмотря на то, что значения определяются для каждой дисциплины отдельно, существуют ограничения на количество часов занятости студента в неделю, которое потребует установления баланса.

Входными данными для модуля А3 являются:

- список названий дисциплин;
- расстановка дисциплин по семестрам;
- распределение трудоемкости УП по дисциплинам;
- формы контроля для каждой дисциплины;
- перечень видов деятельности студента по каждой дисциплине (лекции, лабораторные работы и т.д.);
- критерий оптимизации;
- количество учебных недель в каждом семестре;
- дополнительные условия по распределению трудоемкости (например, требуется лекций в 2 раза больше, чем лабораторных работ, и т.п.).

Управляющими факторами для блока являются выбранный метод оптимизации и нормативные акты (форма учебного плана), механизмами — ЭВМ и составитель УП.

Результатом работы данного блока — выходными данными являются оптимальные значения распределения учебных часов между всеми видами деятельности по каждой дисциплине.

После обработки информации всеми тремя модулями системы получается полная содержательная часть УП.

Последний модуль А4 «Документ» — модуль формирования документа, содержащего учебный план, соответствующий нормативным актам.

Для работы блока необходимы входные данные:

- титульные данные УП (направление обучения, профиль, код набора, срок обучения и т.д.);
- названия всех дисциплин;

- формы контроля;
- содержательная часть УП (расстановка дисциплин, распределение трудоемкости и т.д.).

Управляющими факторами для работы модуля является полный набор нормативных актов (ФГОС, приказы, распоряжения вуза и др.), механизмами выполнения функции — ЭВМ и составитель УП.

В качестве результата работы данного модуля выступает сформированный документ УП по установленной форме, пригодный для использования в работе.

Каждый модуль описываемой СППР «Оптимальный учебный план» представляет собой отдельную подсистему. Они могут быть использованы для получения промежуточного результата для построения УП или в комплексе для поэтапного формирования оптимального УП. Переход от одного модуля может производиться к следующему, но иногда может понадобиться просто вернуться и изменить параметры, до тех пор, пока не будет получен результат, полностью удовлетворяющий запросам составителя УП.

Учитывая сложность системы, необходимо построение диаграмм декомпозиции второго уровня. На рисунке 3 представлена диаграмма декомпозиции блока A1 «Расстановка».

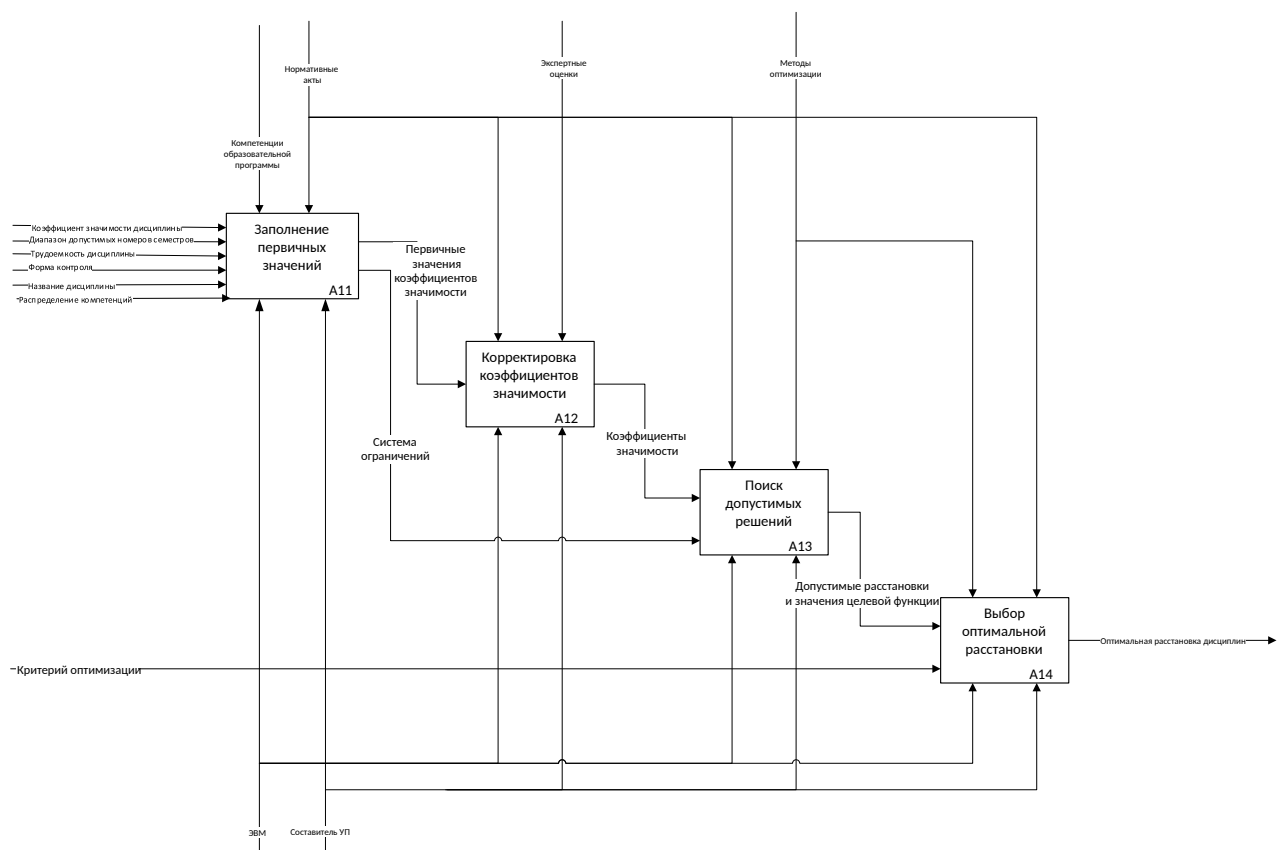


Рис. 3. Диаграмма декомпозиции блока A1 «Расстановка» функциональной модели СППР «Оптимальный учебный план»

Блок A1 «Расстановка» декомпозирован на 4 функциональных блока.

Блок A11 «Заполнение первичных значений» предназначен для первичного заполнения значений коэффициентов значимости дисциплин в соответствии с формируемыми компетенциями и формами контроля (курсовые проекты и работы). Входными данными данный блок требует:

- коэффициент значимости дисциплины (если пользователь хочет установить его самостоятельно);
- диапазон допустимых номеров семестров для каждой дисциплины;
- трудоемкость каждой дисциплины;
- наличие в качестве формы контроля курсового проекта или работы;
- название дисциплины;

- формируемые компетенции для каждой дисциплины.

Выходными данными для этого блока являются первичные значения коэффициентов значимости для каждой дисциплины, основанные только на компетенциях, формируемых дисциплинами, или мнении пользователя.

Управляющим фактором является набор правил о присвоении коэффициента значимости в соответствии с формируемыми компетенциями и формами контроля, а также полный набор компетенций для установления градации.

С помощью составителя и ЭВМ в качестве механизмов формируется набор значений коэффициентов значимости дисциплин [12].

Блок A12 «Корректировка коэффициентов значимости» предназначен для корректировки первичного варианта коэффициентов значимости дисциплин с использованием интеллектуальной подсистемы, построенной на основании экспертного оценивания.

Входными данными данный блок требует набор, сформированный в блоке A11 значений коэффициентов значимости.

Выходными данными являются уточненные в соответствии с экспертным мнением коэффициенты значимости.

Управляющими факторами являются:

- значимость типов компетенций;
- экспертные оценки дисциплин.

Механизмами являются ЭВМ и составитель УП.

Блок A13 «Поиск допустимых решений» предназначен для перебора всех вариантов расстановки дисциплин в допустимые семестры и проверки соответствия установленным ограничениям.

Входными данными данный блок требует набор коэффициентов значимости дисциплин и критерий оптимизации.

Выходными данными являются установленные соответствия между дисциплинами и оптимальными номерами семестров для их изучения, удовлетворяющие системе ограничений.

Управляющими факторами являются:

- метод оптимизации;
- нормативные акты, содержащие нормы.

Механизмами являются ЭВМ и составитель УП.

Блок A14 «Выбор оптимальной расстановки» предназначен для оценки по выбранному критерию всех полученных допустимых решений в блоке A13 и выбора наилучшего из них.

Входными данными данный блок требует набор допустимых решений, полученный в блоке A13, и критерий оптимальности.

Выходными данными являются установленные оптимальные соответствия между дисциплинами и семестрами их изучения в соответствии с выбранным критерием оптимальности.

Управляющими факторами являются:

- метод оптимизации;
- указания из нормативно-правовых актов.

Механизмами являются ЭВМ и составитель УП.

Диаграмма декомпозиции блока A2 «Трудоемкость» представлена на рисунке 4.

Диаграмма декомпозиции блока A2 «Трудоемкость» функциональной модели СППР «Оптимальный учебный план»

Блок A2 «Трудоемкость» декомпозирован на 4 функциональных блока.

Блок A21 «Заполнение первичных данных» предназначен для первичного заполнения расстановки дисциплин с соответствующими значениями коэффициентов значимости, установки диапазонов допустимых значений трудоемкости для каждой дисциплины, установленных форм контроля. Все эти значения используются в качестве постановки оптимизационной задачи по распределению общей трудоемкости между дисциплинами. Входными данными данный блок требует:

- коэффициент значимости каждой дисциплины;
- диапазон допустимых значений трудоемкости для каждой дисциплины;
- расстановку дисциплин по семестрам;
- перечень дисциплин;

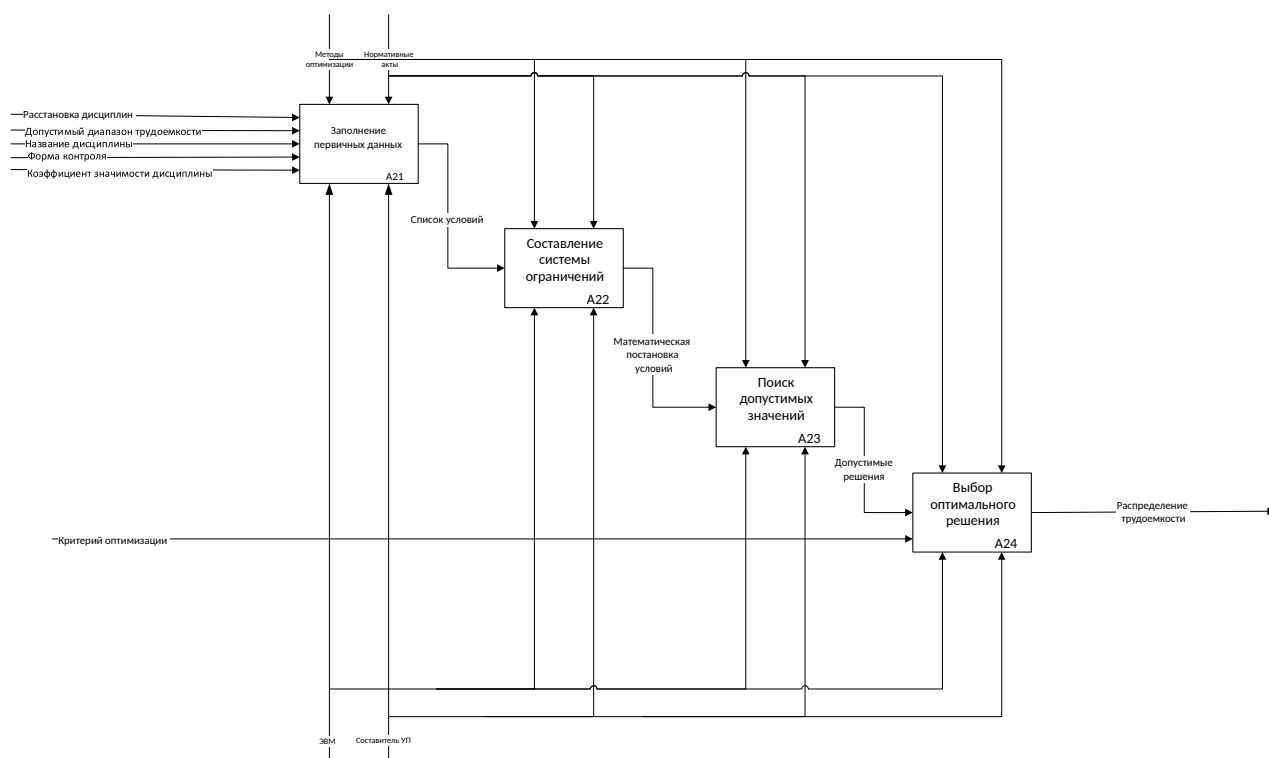


Рис. 4. Диаграмма декомпозиции блока A2 «Трудоемкость» функциональной модели СПМП «Оптимальный учебный план»

- формы контроля по всем дисциплинам.

Выходными данными для этого блока является сформированный список (система) ограничений для оптимизационной задачи.

Управляющими факторами являются математический метод (модель, для которой его можно применить) и сведения из нормативных документов, содержание ограничения на суммарную трудоемкость на один учебный год, допустимые минимальное и максимальное значения на одну дисциплину, допустимый баланс трудоемкости между семестрами одного года и другие.

С помощью составителя и ЭВМ в качестве механизмов формируется список условий для оптимизационной задачи по распределению трудоемкости [12].

Блок A22 «Составление системы ограничений» предназначен для записи формализованных математических ограничений оптимизационной задачи на основании списка условий и нормативных документов.

Входными данными данный блок требует список условий, сформированный в блоке A21.

Выходными данными является математическая система ограничений.

Управляющими факторами являются:

- метод оптимизации;
- нормативные документы, содержащие ограничения (ФГОС ВО и локальные акты).

Механизмами являются ЭВМ и составитель УП.

Блок A23 «Поиск допустимых значений» предназначен для перебора всех вариантов распределения трудоемкости и выбора из них тех, которые соответствуют заданным ограничениям.

Входными данными данный блок требует систему ограничений оптимизационной задачи.

Выходными данными являются допустимые варианты распределения трудоемкости учебного плана между дисциплинами в соответствии со всеми ограничениями.

Управляющими факторами являются:

- метод оптимизации;
- система ограничений.

Механизмами являются ЭВМ и составитель УП.

Блок А24 «Выбор оптимального решения» предназначен для оценки всех допустимых решений, полученных в блоке А23, и выбора наилучшего из них по установленному критерию.

Входными данными данный блок требует перечень всех допустимых решений, полученных в блоке А23, и критерий оптимизации.

Выходными данными являются установленные оптимальные соответствия между дисциплинами и их трудоемкостью.

Управляющими факторами являются:

- метод оптимизации;
- нормативные акты.

Механизмами являются ЭВМ и составитель УП.

Диаграмма декомпозиции блока А3 «Часы» представлена на рисунке 5.

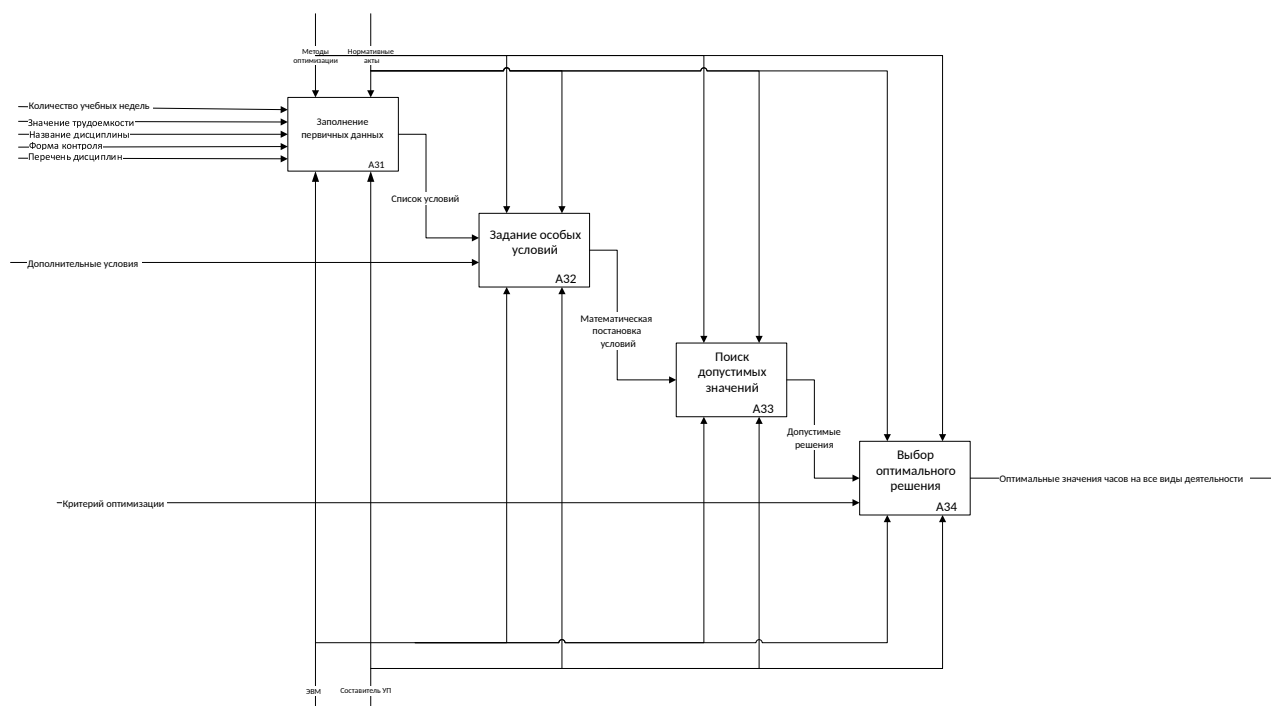


Рис. 5. Диаграмма декомпозиции блока А3 «Часы» функциональной модели СППР «Оптимальный учебный план»

Блок А3 «Часы» декомпозирован на 4 функциональных блока.

Блок А31 «Заполнение первичных данных» предназначен для заполнения всех необходимых сведений, требуемых при решении задачи по оптимизации распределения часов между видами деятельности студента. Входными данными данный блок требует:

- количество учебных недель;
- значения трудоемкости для каждой дисциплины;
- формы контроля по каждой дисциплине;
- названия дисциплин;
- имеющиеся виды деятельности (лекции, практики, лабораторные и т.д.) по каждой дисциплине;
- расстановку дисциплин по семестрам.

Выходными данными для этого блока являются условия, необходимые для записи математической модели задачи.

Управляющими факторами являются математический метод (модель, для которой его можно применить) и сведения из нормативных документов, содержащие нормативные соотношения между различными видами деятельности, нормы выбора количества часов по видам деятельности и т.д.

Механизмами для блока являются составитель УП и ЭВМ [12].

Блок А32 «Задание особых условий» предназначен для добавления особых условий, отсутствующих в нормативных актах, например, если по дисциплине требуется непропорциональное распре-

деление часов или количество аудиторных часов не должно быть пропорционально трудоемкости и т.д.

Входными данными данный блок требует список условий, сформированный в блоке А31, и ввод дополнительных условий для выбранных дисциплин.

Выходными данными является математическая постановка оптимизационной задачи.

Управляющими факторами являются:

- метод оптимизации;
- нормативные документы, содержащие ограничения (ФГОС ВО и локальные акты).

Механизмами являются ЭВМ и составитель УП.

Блок А33 «Поиск допустимых значений» предназначен для перебора всех вариантов распределения часов между установленными видами деятельности в соответствии с системой ограничений.

Входными данными данный блок требует математическую постановку задачи.

Выходными данными являются допустимые перечни распределения часов между видами деятельности по каждой дисциплине в соответствии со всеми ограничениями.

Управляющими факторами являются:

- метод оптимизации;
- нормативные документы.

Механизмами являются ЭВМ и составитель УП.

Блок А34 «Выбор оптимального решения» предназначен для оценки всех допустимых решений, полученных в блоке А33, и выбора наилучшего из них.

Входными данными данный блок требует перечень всех допустимых решений, полученных в блоке А33, и критерий оптимизации.

Выходными данными являются оптимальные соответствия между дисциплинами и распределениями часов между видами деятельности студентов в соответствии с постановкой задачи.

Управляющими факторами являются:

- метод оптимизации;
- нормативные акты.

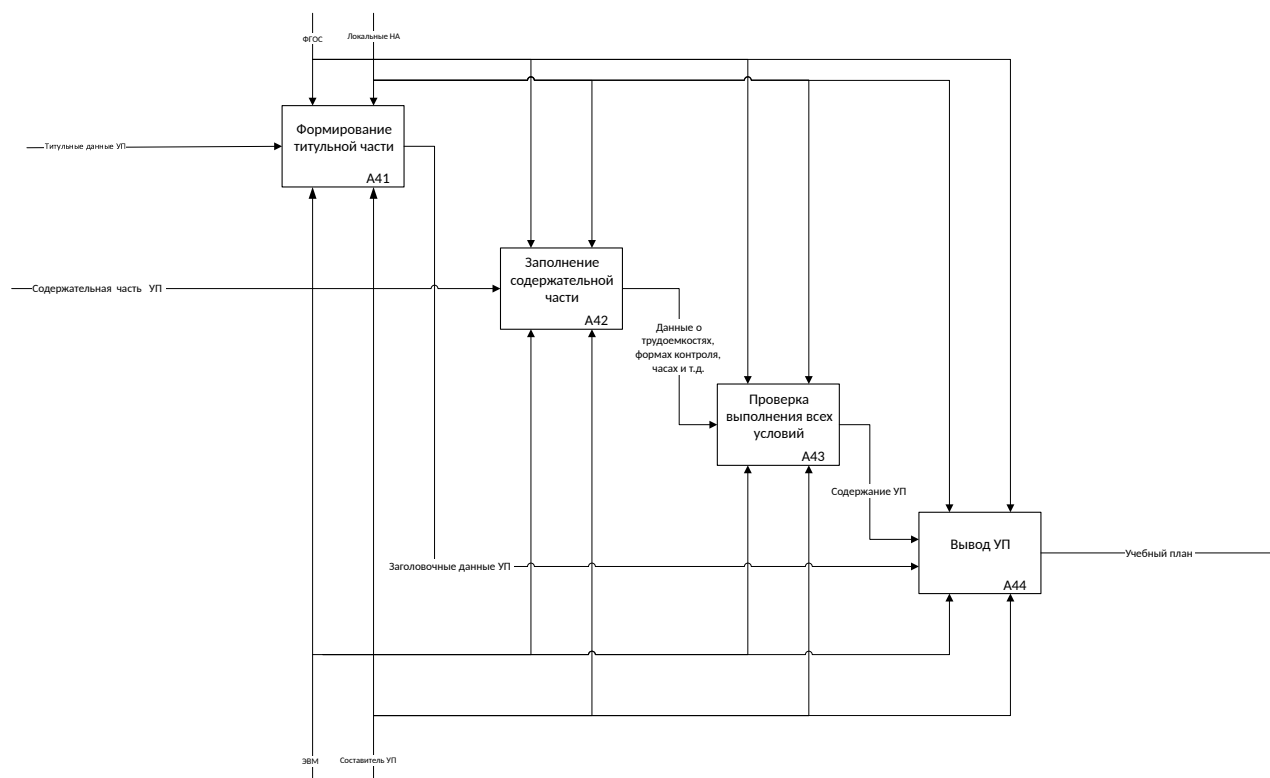


Рис. 6. Диаграмма декомпозиции блока А4 «Документ» функциональной модели СПМР «Оптимальный учебный план»

Механизмами являются ЭВМ и составитель УП.

Диаграмма декомпозиции блока А4 «Документ» представлена на рисунке 6.

Блок А4 «Документ» декомпозирован на 4 функциональных блока.

Блок А41 «Формирование титульной части» предназначен для заполнения всех необходимых сведений для титульного листа УП (направление, учебный год, профиль и т.д.).

Входными данными данный блок требует сведения для заполнения титульного листа УП.

Выходными данными для этого блока являются заголовочные данные УП.

Управляющими факторами являются ФГОС ВО и локальные нормативные акты вуза.

С помощью составителя и ЭВМ в качестве механизмов формируется титульная страница УП.

Блок А42 «Заполнение содержательной части» предназначен для заполнения основной части УП, т.е. названий дисциплин, их местоположения в плане, трудоемкости, распределения часов и т.д.

Входными данными данный блок требует содержание УП. Именно эти данные можно получить в результате решения оптимизационных задач в блоках А1-А3.

Выходными данными являются заполненные страницы УП с содержанием.

Управляющими факторами являются:

- ФГОС;
- локальные нормативные документы.

Механизмами являются ЭВМ и составитель УП.

Блок А43 «Проверка выполнения всех условий» предназначен для финальной проверки соответствия построенного плана всем требованиям ФГОС ВО и локальных нормативных актов.

Входными данными данный блок требует математическую постановку задачи.

Выходными данными являются заполненные страницы содержательной части УП.

Управляющими факторами являются:

- ФГОС ВО;
- локальные нормативные документы.

Механизмами являются ЭВМ и составитель УП.

Блок А44 «Вывод УП» предназначен для окончательного формирования документа «Учебный план».

Входными данными данный блок требует заполненные титульный лист и все страницы содержательной части УП.

Выходными данными являются все заполненные страницы учебного плана.

Управляющими факторами являются:

- ФГОС ВО;
- локальные нормативные документы.

Механизмами являются ЭВМ и составитель УП.

Таким образом, описаны все функциональные возможности проектируемой системы для формирования оптимального учебного плана.

Заключение

Построена функциональная модель СППР «Оптимальный учебный план». Контекстная диаграмма модели дает достаточное представление о необходимых данных, методах, правилах, необходимых для построения системы. Также диаграмма дает представление об ожидаемых результатах системы. Диаграмма декомпозиции первого уровня показывает модульную структуру системы и взаимосвязь между модулями. Представленные четыре диаграммы декомпозиции второго уровня детально описывают функциональные составляющие модулей системы.

ЛИТЕРАТУРА

1. Государственный образовательный стандарт высшего профессионального образования. Требования к минимуму содержания и уровню подготовки инженеров по специальности 220200 — Автоматизированные системы обработки информации и управления. Минобрнауки России; 1995. 88 с.

2. *Федеральный государственный стандарт высшего образования — бакалавриат по направлению подготовки 09.03.01 Информатика и вычислительная техника от 19 сентября 2017 г. № 929*. Минобрнауки России; 2017. 21 с.
3. *Конструктор учебных планов*. Режим доступа: <https://edsoo.ru/konstruktor-uchebnyh-planov>.
4. *АС Планы. ММИС Лаборатория*. Режим доступа: <https://www.mmis.ru/Portals/0/Download/Plany/InstrPIDemo.htm>.
5. Тришкова А. В., Магомедова М. К., Томашевский Ю. Б. Сравнение учебных планов компетентностно-ориентированных основных профессиональных образовательных программ высшего образования на основе графовых моделей. *Инновационные технологии в обучении и производстве: материалы XVI Всероссийской заочной научно-практической конференции*. Волгоград; 2021:165–168.
6. Зыкова Т. В., Вайнштейн Ю. В., Носков М. В. О графовой модели учебного плана образовательной программы. *Информатизация образования и методика электронного обучения: цифровые технологии в образовании: материалы VIII Международной научной конференции*. Красноярск; 2024:109–113.
7. Кузьмина Е. А., Низамова Г. Ф. Формирование учебного плана на основе графовой модели. *Информатика и образование*. 2020;5(314):33–43. DOI: 10.32517/0234-0453-2020-35-5-33-43.
8. Сеньковская А. А. *Математическое моделирование и методы оптимизации учебного процесса вуза (на примере задач формирования рабочих учебных планов)*: дис. канд. техн. наук. Омск; 2021. 154 с.
9. Сеньковская А. А., Фураева И. И. Алгоритмы оптимизации рабочих учебных планов. *Математическое и компьютерное моделирование: сборник материалов IV Международной научной конференции*. Омск; 2016:91–93.
10. Сеньковская А. А., Фураева И. И. Моделирование и оптимизация процесса распределения человеческих ресурсов и аудиторного фонда вуза на основе анализа учебных планов. *Математические структуры и моделирование*. 2017;2(42):119–126. DOI: 10.24147/2222-8772.2017.2.119-126.
11. Юрчишина М. В., Бушмелева К. И. Математическое моделирование процесса оптимизации учебного плана высшего образования. *Фундаментальные, поисковые, прикладные исследования и инновационные проекты: сборник трудов Национальной научно-практической конференции*. Москва; 2022:141–144.
12. Юрчишина М. В., Бушмелева К. И. Алгоритмическая модель СППР «Оптимальный учебный план». *Моделирование систем и процессов*. 2024;4(17):84–95. DOI: 10.12737/2219-0767-2024-17-4-84-95.
13. Юрчишина М. В. Обобщенный алгоритм работы системы поддержки принятия решений «Оптимальный учебный план». *Инновационные, информационные и коммуникационные технологии: сборник XXI Международной научно-практической конференции*. Сочи; 2024:393–397.
14. Алексахин С. В., Николаев А. Б., Строганов В. Ю. Моделирование связности дисциплин учебного плана в системе дистанционного образования. *Информационные технологии в образовании (ИТО-2001). Секция: 3. ИТ в открытом образовании: материалы XI Международ. конф.-выставки*. М.: МИФИ; 2001.
15. Лавлинская О. Ю. Ранжирование учебных дисциплин с использованием экспертных оценок. *Моделирование систем и информационные технологии: межвуз. сб. науч. тр.* Воронеж: Научная книга; 2006. Вып. 3. Ч. 2. С. 80–83.
16. Бабкина О. М., Бабкин Е. А. Об оптимизации учебных планов. *Вестник МГПУ. Серия: Информатика и информатизация образования*. 2006;7:19-20.
17. Кривицкая М. А. Проектирование автоматизированной системы синтеза рабочего учебного плана на направления высшего образования. *Вестник Сургутского государственного университета*. 2013;2:20-22.
18. Юрчишина М. В., Бушмелева К. И. Подход к формированию групп связанных дисциплин для обучающихся бакалавриата по направлению «Информатика и вычислительная техника». *Успехи кибернетики*. 2023;4(2):54–60. DOI: 10.51790/2712-9942-2023-4-2-08.
19. Сапожник К. В. Концепция подсистемы подбора коэффициента важности дисциплин. *Фундаментальные, поисковые, прикладные исследования и инновационные проекты: сборник трудов Национальной научно-практической конференции*. Москва; 2023:698–701.

Сетевое издание. Полные тексты статей размещаются на официальном сайте издания jsyb.ru.

Адрес учредителя и издателя:

Федеральное государственное автономное учреждение «Федеральный научный центр
Научно-исследовательский институт системных исследований Национального исследовательского
центра «Курчатовский институт»,
117218, г. Москва, Нахимовский просп., дом 36, корпус 1. Телефон: +7 (495) 718-21-10.

Дата выхода в свет 30.06.2025.

Формат 60 × 84/8.

Усл. печ. л. 8,2. Уч.-изд. л. 12,2.

Цена свободная.