

ВЕСТНИК КИБЕРНЕТИКИ

PROCEEDINGS IN CYBERNETICS

12+



**БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
ХАНТЫ-МАНСИЙСКОГО АВТОНОМНОГО ОКРУГА – ЮГРЫ
«СУРГУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**

**ВЕСТНИК КИБЕРНЕТИКИ
PROCEEDINGS IN CYBERNETICS**

Научный журнал

ТОМ 23, № 3

**Сургут
2024**

Учредитель и издатель
Бюджетное учреждение высшего образования Ханты-Мансийского автономного округа – Югры
«Сургутский государственный университет»

Главный редактор
Острейковский Владислав Алексеевич, д. т. н., профессор

Заместитель главного редактора
Тараканов Дмитрий Викторович, к. т. н., доцент

Ответственный секретарь
Брагинский Михаил Яковлевич, к. т. н., доцент

Члены редакционной коллегии:
Советов Б. Я., академик РАН и РАО, д. т. н., профессор
Юсупов Р. М., член-корреспондент РАН, д. т. н., профессор
Абрамов О. В., д. т. н., профессор
Анохин А. Н., д. т. н., профессор
Антонов А. В., д. т. н., профессор
Бурханов Р. А., д. филос. н., профессор
Бушмелева К. И., д. т. н., профессор
Гетман А. Ф., д. т. н., профессор
Григорьев Л. И., д. т. н., профессор
Древс Ю. Г., д. т. н., профессор
Ельников А. В., д. ф.-м. н., профессор
Инютин С. А., д. т. н., профессор
Каштанов В. А., д. ф.-м. н., профессор
Керимов Т. Х., д. филос. н., профессор
Крамаров С. О., д. ф.-м. н., профессор
Леонов Д. Г., д. т. н., профессор
Малышев Д. С., д. ф.-м. н., профессор
Мельников А. В., д. т. н., профессор
Певзнер Л. Д., д. т. н., профессор
Увайсов С. У., д. т. н., профессор
Цибульский В. Р., д. т. н., профессор
Юрков Н. К., д. т. н., профессор

Выпускающий редактор
Хасанова А. Ш.
Редактор
Черных О. Л.
Верстка
Чорньенского С. И.
Переводчик
Бенская М. О.

Рисунок на обложке:
Визуализация магнитогидродинамических расчетов для точного решения внутри сферы
© Галкин В. А., Гореликов А. В.

Решением Высшей аттестационной комиссии с 29 мая 2017 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». С 29.03.2022 издание принимает к публикации статьи по следующим научным специальностям: 2.3.1. Системный анализ, управление и обработка информации; 1.2.2. Математическое моделирование, численные методы и комплексы программ; 2.2.2. Электронная компонентная база микро- и нанoeлектроники, квантовых устройств.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций. Свидетельство о регистрации средства массовой информации
Эл № ФС77-83015 от 31.03.2022.

Издается с 2002 года. Выпускается 4 раза в год.

Адрес редакции:
БУ ВО «Сургутский государственный университет»,
628412, Россия, Ханты-Мансийский автономный округ – Югра, г. Сургут, пр. Ленина, 1
Тел.: +7 (3462) 76-29-88, факс: +7 (3462) 76-29-29, e-mail: proceed.cyber@mail.ru. Сайт: vestcyber.ru

Founder and Publisher
Surgut State University

Chief Editor

Prof. Vladislav A. Ostreikovskiy, Doctor of Sciences (Engineering)

Vice Chief Editor

Dmitry V. Tarakanov, Candidate of Sciences (Engineering), Docent

Executive Editor

Mikhail Ya. Braginsky, Candidate of Sciences (Engineering), Docent

Editorial Board:

Prof. Sovetov B. Ya., Doctor of Sciences (Engineering), member of the Russian Academy of Sciences (RAS), the Russian Academy of Education (RAE)

Prof. Yusupov R. M., Doctor of Sciences (Engineering), member of the Russian Academy of Sciences (RAS)

Prof. Abramov O. V., Doctor of Sciences (Engineering)

Prof. Anokhin A. N., Doctor of Sciences (Engineering)

Prof. Antonov A. V., Doctor of Sciences (Engineering)

Prof. Burkhanov R. A., Doctor of Sciences (Philosophy)

Prof. Bushmeleva K. I., Doctor of Sciences (Engineering)

Prof. Getman A. F., Doctor of Sciences (Engineering)

Prof. Grigoryev L. I., Doctor of Sciences (Engineering)

Prof. Drevs Yu. G., Doctor of Sciences (Engineering)

Prof. Elnikov A. V., Doctor of Sciences (Phys&Math)

Prof. Inyutin S. A., Doctor of Sciences (Engineering)

Prof. Kashtanov V. A., Doctor of Sciences (Phys&Math)

Prof. Kerimov T. Kh., Doctor of Sciences (Philosophy)

Prof. Kramarov S. O., Doctor of Sciences (Phys&Math)

Prof. Leonov D. G., Doctor of Sciences (Engineering)

Prof. Malyshev D. S., Doctor of Sciences (Phys&Math)

Prof. Melnikov A. V., Doctor of Sciences (Engineering)

Prof. Pevzner L. D., Doctor of Sciences (Engineering)

Prof. Uvaisov S. U., Doctor of Sciences (Engineering)

Prof. Tsibulsky V. R., Doctor of Sciences (Engineering)

Prof. Yurkov N. K., Doctor of Sciences (Engineering)

Publishing Editor

Khasanova A. Sh.

Editor

Chernykh O. L.

Layout

Chornenkiy S. I.

Translator

Benskaya M. O.

Cover Image:

A visualization of magnetohydrodynamic analysis used to find an exact solution within a sphere

© Galkin V. A., Gorelikov A. B.

Since 29.05.2017 the journal is included in the List of Leading Peer-Reviewed Scientific Journals of the Higher Attestation Commission, which publishes main scientific results of Doctor's and Candidate's theses.

Since March 29, 2022, the journal is included in the List on the following subject groups:

- 2.3.1. System Analysis, Data Management and Processing; 1.2.2. Mathematical Modeling, Calculus and Program Complexes;
2.2.2. Electronic Component Base of Micro- and Nanoelectronics, Quantum Devices.

The journal is registered with the Federal Service for Supervision of Communications, Information Technology and Mass Media. Mass media registration certificate El No. FS77-83015 dated on 31.03.2022.

Published since 2002. 4 issues per year.

Editorial Board Address:

Surgut State University, Russia 628412, Khanty-Mansi Autonomous Okrug – Ugra, Surgut, Lenina pr., 1

Tel.: +7 (3462) 76-29-88, fax: +7 (3462) 76-29-29, e-mail: proceed.cyber@mail.ru. Web: vestcyber.ru

СОДЕРЖАНИЕ

ТЕХНИЧЕСКИЕ НАУКИ

Антонюк П. В., Острейковский В. А., Сорочкин А. В. Оценка свойства безотказности элементов системы верхнего привода по результатам эксплуатации на нефтяных месторождениях Западной Сибири.....	6
Вережкин А. И., Лысенкова С. А. Проектирование информационной системы для диагностики аппаратной части компьютера	16
Детков А. А., Воронина В. А., Гарифуллина Ю. В., Корепанов А. М., Вишнякова А. Ю. Сравнительный анализ метрик векторного расстояния растровых изображений.....	22
Кузнецов К. В., Лысенкова С. А. Краткое аннотирование новостей из области информационных технологий.....	31
Носков С. И., Медведев А. П., Овсянников И. В. Анализ срабатываний предикторов в функции Леонтьева для объема финансирования отделения социального фонда	40
Пирогов А. А., Хорошайлова М. В., Сёмка Э. В. Разработка архитектуры нейросетевого декодирования, основанной на системах стробирования и распределения весовых коэффициентов.....	46
Попов О. Р., Крамаров С. О. Оптимизация при вероятностном тематическом моделировании технологической прогностической информации.....	56
Турецкий А. В., Пирогов А. А., Свиридова И. В., Хорошайлова М. В., Сёмка Э. В. Анализ влияния шероховатости паяльной маски и формы трафарета на образование пустот в паяных соединениях	70
Чубарова О. В., Чубаров А. В., Ликсонова Д. И. Применение метода вейвлет-анализа данных для построения прогнозной модели	82

ФИЗИКО-МАТЕМАТИЧЕСКИЕ НАУКИ

Беспалов М. С., Фролов К. А. Кодирование информации линейными перестановками дискретного преобразования Уолша	90
Дёмко А. И., Алмазова Е. Г. Оценка минимального достижимого уровня боковых лепестков функции неопределенности радиоимпульсных последовательностей	96

CONTENTS

ENGINEERING

<i>Antonyuk P. V., Ostreikovskiy V. A., Sorochkin A. V.</i> Assessment of failure free properties of elements of the top drive system based on operating data of oil fields in Western Siberia.....	6
<i>Verevkin A. I., Lysenkova S. A.</i> Designing an information system for computer hardware diagnostics.....	16
<i>Detkov A. A., Voronina V. A., Garifullina Yu. V., Korepanov A. M., Vishnyakova A. Yu.</i> Comparative analysis of vector distance metrics for raster images.....	22
<i>Kuznetsov K. V., Lysenkova S. A.</i> Summary of news in information technology.....	31
<i>Noskov S. I., Medvedev A. P., Ovsyannikov I. V.</i> Analysis of predictor responses in the Leontief function for financing amount of the social fund department.....	40
<i>Pirogov A. A., Khoroshailova M. V., Syomka E. V.</i> Development of a neural network decoding architecture based on gating and weight distribution systems.....	46
<i>Popov O. R., Kramarov S. O.</i> Optimization in probabilistic topic modeling of technological predictive information	56
<i>Turetsky A. V., Pirogov A. A., Sviridova I. V., Khoroshailova M. V., Syomka E. V.</i> Analysis of the influence of solder mask roughness and stencil shape on the formation of voids in solder joints.....	70
<i>Chubarova O. V., Chubarov A. V., Liksonova D. I.</i> Application of wavelet data analysis for forecast model development	82

PHYSICS AND MATHEMATICS

<i>Bespalov M. S., Frolov K. A.</i> Information encoding by linear permutations of discrete Walsh transform.....	90
<i>Dyomko A. I., Almazova E. G.</i> Assessment of minimum achievable sidelobes level of ambiguity function for radio pulse sequences.....	96

ТЕХНИЧЕСКИЕ НАУКИ / ENGINEERING

Научная статья

УДК 622.242-8:622.276.7

<https://doi.org/10.35266/1999-7604-2024-3-1>



Оценка свойства безотказности элементов системы верхнего привода по результатам эксплуатации на нефтяных месторождениях Западной Сибири

Павел Викторович Антонюк¹, Владислав Алексеевич Острейковский²,
Андрей Викторович Сорочкин³

^{1, 2, 3}Сургутский государственный университет, Сургут, Россия

¹Pavel_Antoniuk@mail.ru

²ova@surgu.ru

³sorochkin_av@surgu.ru

Аннотация. Системы верхнего привода (СВП), смонтированные на мобильных буровых комплексах, применяются для ремонта нефтяных и газовых скважин. Собраны и проанализированы данные по причинам и времени простоев более 60 единиц систем верхнего привода, эксплуатируемых на месторождениях Западной Сибири. СВП эксплуатируются в неблагоприятных условиях внешней среды с высоким уровнем и нестационарностью действующих нагрузок и, как следствие, зачастую с недостаточным уровнем функциональной надежности, что приводит к большим объемам ремонтно-восстановительных работ с частой заменой отработавших свой ресурс элементов. Отказ элементов СВП приводит к остановке всего технологического процесса ремонта скважин и, как следствие, к большим экономическим потерям, поэтому одним из важных направлений в проблеме управления надежностью СВП является совершенствование методов обоснования и анализа показателей надежности по статистическим данным реальной эксплуатации. Выполнен расчет показателей надежности и определены законы распределения, позволяющие качественно оценить функциональную надежность элементов СВП, что позволит принимать обоснованные решения при обеспечении требуемого уровня надежности.

Ключевые слова: система верхнего привода, надежность, мобильные буровые комплексы

Финансирование: работа выполнена при поддержке РФФИ (объект 18-07-391).

Для цитирования: Антонюк П. В., Острейковский В. А., Сорочкин А. В. Оценка свойства безотказности элементов системы верхнего привода по результатам эксплуатации на нефтяных месторождениях Западной Сибири // Вестник кибернетики. 2024. Т. 23, № 3. С. 6–15. <https://doi.org/10.35266/1999-7604-2024-3-1>.

Original article

Assessment of failure free properties of elements of the top drive system based on operating data of oil fields in Western Siberia

Pavel V. Antonyuk¹, Vladislav A. Ostreikovskiy², Andrey V. Sorochkin³

^{1, 2, 3}Surgut State University, Surgut, Russia

¹pavel_antoniuk@mail.ru

²ova@surgu.ru

³sorochkin_av@surgu.ru

Abstract. Top drive systems, mounted on mobile drilling complexes, are used for repairing oil and gas wells. Our team collected and analyzed data on the causes and downtime of more than 60 units of top drive systems operating in the fields of Western Siberia. Operators use top drive systems under adverse conditions that have a high level and unsteadiness of actual loads. Consequently, the functional reliability of these systems often falls short, leading to numerous repair and recovery activities and frequent replacement of elements that have exceeded their service life. Failure of top drive systems elements leads to the stop of the entire technological process of well repair and, as a result, to large economic losses. Therefore, one of the important directions in the problem of managing the reliability of top drive systems is the improvement methods for substantiating and analyzing reliability indicators based on statistical data from actual use. We have calculated reliability indicators and determined distribution laws that enable us to assess the functional reliability of top drive system elements in a qualitative manner. This will enable us to make informed decisions while ensuring the necessary level of reliability.

Keywords: top drive system, reliability, mobile drilling complexes

Funding: the study is supported by the Russian Foundation for Basic Research (object 18-07-391).

For citation: Antonyuk P. V., Ostreikovsky V. A., Sorochkin A. V. Assessment of failure free properties of elements of the top drive system based on operating data of oil fields in Western Siberia. *Proceedings in Cybernetics*. 2024;23(3):6–15. <https://doi.org/10.35266/1999-7604-2024-3-1>.

ВВЕДЕНИЕ

Ежегодно растущая потребность в нефтяных и газовых продуктах и истощение имеющихся месторождений приводит к повышению рентабельности добычи «трудно-извлекаемых запасов». Для добычи «трудно-извлекаемых запасов» нефти и газа и строительства вертикальных, наклонно-направленных и горизонтальных скважин применяют системы верхнего привода (СВП) [1–3]. СВП – механизм, обеспечивающий вращение буровой колонны (ремонтного инструмента и др.), включающий верхний привод с приводом индивидуальным или от общего энергоблока и трансмиссии.

СВП, смонтированные на стационарных или мобильных буровых комплексах, значительно повышают скорость и безопасность проведения работ, расширяют технологические возможности, ранее недоступные, в сравнении с технологией бурения с применением роторного стола, уже ставшей традиционной. При этом повышается (усложняется) технический и технологический уровень эксплуатации и обслуживания бурового оборудования.

Основным преимуществом применения СВП являются:

– высокая скорость и безопасность выполнения технологических операций;

– значительное снижение вероятности «прихвата» буровой колонны;

– расширенные технологические возможности (зарезка боковых стволов, наклонно-направленное бурение);

– возможность задавать и контролировать величину крутящего момента и частоту вращения;

– возможность тормозить и удерживать буровую колонну в заданном положении, плавно снимать реактивный момент;

– возможность производить спуск обсадных труб в зонах осложнений за счет вращения и промывки;

– возможность герметизировать внутри-трубное пространство двумя шаровыми кранами (гидравлическим и механическим), что позволяет снизить вероятность выброса из скважины через буровую колонну.

Выше представленные технологические возможности СВП позволили добраться до горизонтов, которые бы ранее оставались недоступными, а также повысить дебит скважин.

СВП – вертлюг, оснащенный гидравлическим или электрическим приводом, подвешиваемый за талевый блок и совершающий движение вдоль мачты буровой установки совместно с талевым блоком. Привод СВП позволяет вращать буровую колонну с инструментом; вертлюг, совмещенный с при-

водом, осуществляет подачу бурового раствора одновременно с вращением, а также участвует в удержании всей колонны. Вращательный момент, развиваемый двигателем СВП, принимается направляющей и передается на мачту. В некоторых СВП в качестве направляющих, помимо балки, применяют стальной талевый канат.

На рынке существует большое количество производителей СВП – США, Германия, Франция, Китай. Помимо иностранных производителей разработкой и изготовлением занимается и отечественный производитель – например, ОАО «Электромеханика» в городе Санкт-Петербург, выпускающее целую линейку СВП грузоподъемностью от 80 до 320 т [1–7].

Высокая конструктивная сложность СВП требует особого внимания как на этапе проектирования и изготовления, так и на этапе эксплуатации, технического обслуживания и ремонта. Отказ СВП при бурении означает огромные финансовые потери, включающие в себя стоимость простоя, запасных частей, ремонтно-восстановительных работ, ликвидации последствий аварии (в случае возникновения). Большие экономические потери, нефтяное загрязнение крупных участков местности, человеческие жертвы, потеря дорогостоящего нефтегазодобывающего оборудования – вот характерные результаты отказов элементов СВП.

Поэтому дальнейшее повышение надежности, безопасности и рентабельности эксплуатации СВП, входящих в состав буровых комплексов, является одной из приоритетных задач нефтегазодобывающих предприятий.

МАТЕРИАЛЫ И МЕТОДЫ

Исходя из структурной схемы (рис. 1), причины и время простоев, связанных с отказами СВП в период 2013–2018 гг., были распределены по подсистемам и представлены на рис. 2–3. Наименьшее влияние на время/количество простоев оказывают отказы элементов трансмиссии – 1 % / 1 %; переводника – 1 % / 1 %; шарового крана – 4 % / 3 %; двигателя – 8 % / 8 %; вертлюга – 11 % / 15 %;

наибольшее влияние оказывают отказы элементов гидравлической системы – 47 % / 43 % и электрической системы – 18 % / 19 %.

Большое влияние элементов гидравлической системы объяснимо тем, что СВП, по сути, представляет собой гидравлическую станцию с гидравлическим приводом, т. е. гидравлическая система составляет значительную часть и является основной работающей подсистемой; также большое количество отказов наблюдается в элементах вертлюга и электрической системы.

На основе статистических данных по простоям и отказам было вычислено среднее время восстановления подсистем СВП в работоспособное состояние (табл. 1).

Значительное время затрачивается на восстановление в работоспособное состояние подсистем шарового крана – 9,7 часа на 1 отказ, самое минимальное время затрачивается на восстановление элементов вертлюга – 5,1 часа на 1 отказ, на восстановление оставшихся подсистем затрачивается от 6,2 до 7,8 часа на 1 отказ.

Для более детального анализа причин отказов СВП отказы элементов разнесены по подсистемам (рис. 4). Наибольшее количество отказов СВП связано с отказами гидравлической системы (около 40 %) и верхнего силового привода (ВСП) (32 %). Также зафиксированы отказы гидробазиса, генератора, насоса, шарового крана (они не превышают 1 %), отказы электрической системы (9 %), ДВС и трубы грязевой (5 %), каретки (3 %) и гидротормоза (2 %).

Значительное количество отказов, наблюдаемое в гидравлической системе, связано с конструктивными особенностями и нагруженностью гидравлической системы, т. к. основная рабочая функция (вращение буровой колонны) выполняется гидравлическим приводом со значительным количеством элементов гидравлической системы (рукава высокого давления, насосы, клапаны, гидромоторы, гидрораспределители, дроссели), также по той причине, что в связи с высоким уровнем автоматизации процессов управления и обеспечения безопасности проведения работ на-

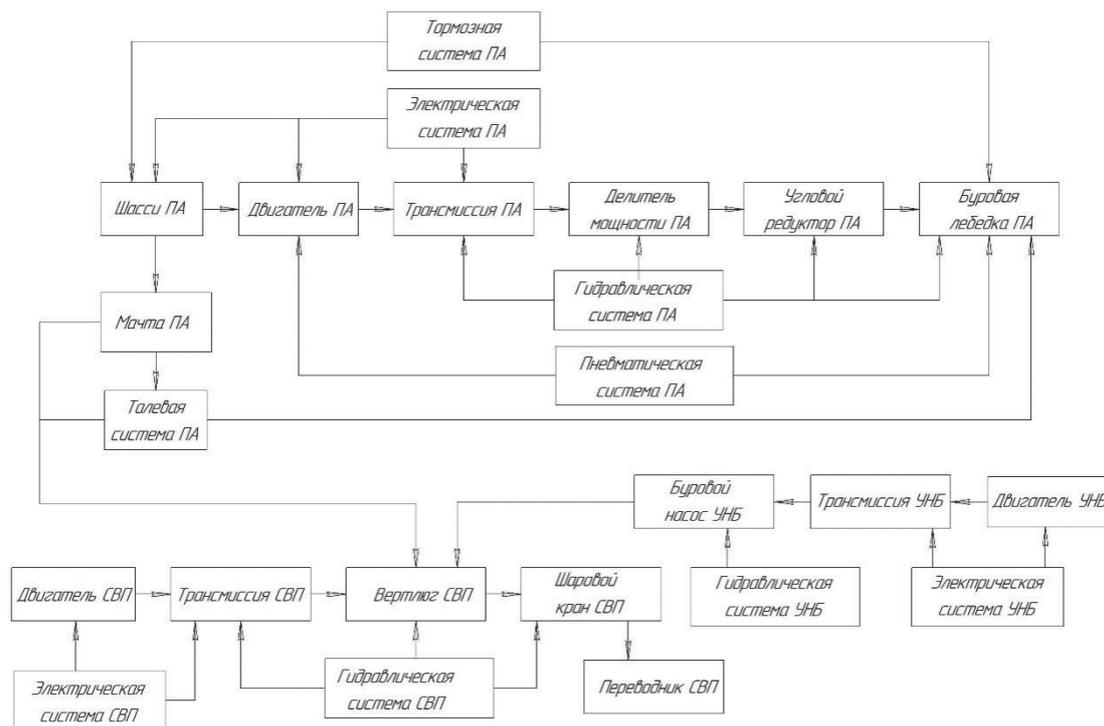


Рис. 1. Структурная схема мобильного бурового комплекса

Примечание: составлено авторами.



Рис. 2. Время простоев из-за отказов подсистем СВП

Примечание: составлено авторами.



Рис. 3. Количество простоев из-за отказов подсистем СВП

Примечание: составлено авторами.

Таблица 1

Среднее время восстановления в работоспособное состояние

Наименование	Среднее время восстановления, час
Двигатель	6,2
Трансмиссия	7,8
Гидравлическая система	7,4
Электрическая система	6,2
Вертлюг	5,1
Шаровой кран	9,7
Переводник	5,0
Прочие	6,9

Примечание: составлено авторами.

блюдается достаточно большое количество отказов электрической системы.

Из рис. 2–4 видно, что не все отказы приводят к простоям. Так, например, на шаровой кран приходится 1 % отказов, 3 % простоев и 4 % времени простоев, что объяснимо невозможностью устранения неисправности в течение короткого промежутка времени непосредственно на месте проведения работ и высокой стоимостью запасных частей; 9 % отказов электрической системы приводит к 19 % количества простоев и 18 % времени простоев, что объясняется высокой сложностью поиска и устранения неисправности; 5 % отказов ДВС приводит к 8 % количества и времени простоев; 32 % отказов ВСП приводит к 15 % количества простоев и 11 % времени простоев, что говорит о достаточно быстром устранении неисправностей непосредственно на месте проведения работ и о достаточно высокой скорости обнаружения неисправности. Также хочется отметить фактор наличия требуемого ремонтного резерва непосредственно на месте проведения работ, а также допускаемой возможности работы оборудования СВП при незначительных неисправностях до завершения операций по текущему или капитальному ремонту скважин. В дополнение хочется отметить необходимость проведения ремонтно-восстановительных работ и формирование ремонтного запаса непосредственно на месте проведения работ; так, при 5 % отказов грязевой трубы простоев по причине отказов

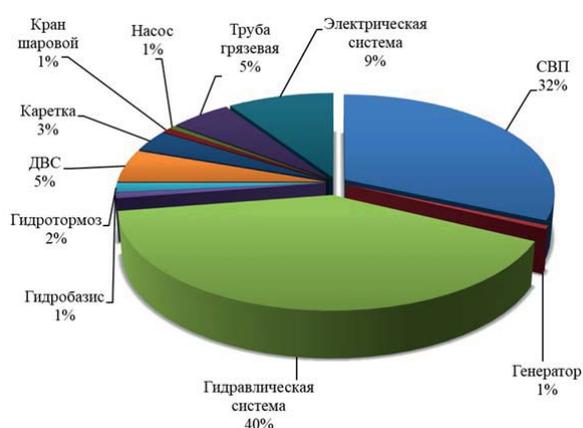


Рис. 4. Отказы СВП по подсистемам
 Примечание: составлено авторами.

грязевой трубы не зафиксировано. Таким образом, сбор, хранение и анализ статистических данных по отказам и простоям оказывают значительное влияние на эффективность и рентабельность эксплуатации СВП и бурового оборудования.

Определение законов распределения случайной величины выполнялось на основе статистических данных эксплуатации СВП. Сбор, хранение, обработка и анализ статистических данных эксплуатации СВП проводились с помощью программного обеспечения Microsoft Excel, Statistica, Mathcad [8–10]. Работоспособность узлов и механизмов СВП рассматривалась как отдельно по узлам и механизмам, так и по системам в целом.

В процессе обработки и анализа статистических данных было необходимо проверить гипотезы о виде законов распределения с определением эмпирических законов распределения [11].

При обработке статистических данных по показателям надежности для оценки значимости различий между фактическим (выявленным в результате исследования) количеством исходов или качественных характеристик выборки, попадающих в каждую категорию, и теоретическим количеством, которое можно ожидать в изучаемых группах при справедливости нулевой гипотезы [8–16], используют критерий согласия Пирсона (критерий χ^2 – хи-квадрат):

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}, \quad (1)$$

где n_i – опытная частота в i -м интервале статистического ряда;

np_i – теоретическая частота (рассчитанная по закону распределения) в i -м интервале;

k – число интервалов укрупненного статистического ряда.

Рассчитав χ^2 и количество степеней свободы ν , определяют $P(\chi^2)$ того, что χ^2 не превысит заданного значения χ_0^2 .

Одной из задач обработки статистических данных об отказах СВП для оценки показателей надежности является получение функции распределения случайной величины.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

По результатам обработки статистических данных эксплуатации СВП были получены эмпирические функции распределения и выполнен расчет показателей надежности элементов СВП. Результаты представлены в табл. 2; результаты обработки данных в программе Statistica так же представлены в графическом виде (рис. 5–8).

С целью определения законов распределения случайной величины рассмотрен вариант деления времени эксплуатации на периоды.

Первый период эксплуатации соответствует периоду «приработки», находится в интервале времени до 2 500 часов и не противоречит

экспоненциальному закону распределения случайной величины (рис. 7).

Второй период эксплуатации соответствует периоду приработки, «нормальной эксплуатации», находится в интервале времени от 2 500 до 4 500 часов и не противоречит нормальному закону распределения случайной величины (рис. 8).

Если принятие решения о справедливости гипотезы о законе распределения можно осуществить, ориентируясь на эмпирическое значение χ^2 , либо на расчетное значение вероятности $P(\chi^2)$ с принятым уровнем значимости, принимая вероятностную точку зрения на происхождение данных, то дальнейшие суждения, основанные на представленных дан-

Таблица 2

Показатели безотказности элементов СВП

Наименование	Кол-во отказов, n	Средняя наработка на отказ, ч. T_{cp}	Среднее квадр. откл., ч. σ_t	Коэффициент вариации	Закон распределения	χ^2	$P(\chi^2)$
Труба грязевая	51	3653,18	2219,92	0,61	нормальный	3,39	0,49
Гидробазис MFA	11	26 282,18	12 948,61	0,49	–	–	–
Подшипник В7032С	19	20 458,10	15 643,47	0,76	экспоненциальный	0,078	0,77
Кран шаровой	10	27 283,20	16 335,97	0,60	–	–	–
Ролик конусный ходовой	14	33 372	16 198,80	0,60	–	–	–
Гидравлический аксиально-поршневой насос 90R180	5	38 035,2	21 115,54	0,55	–	–	–
Подшипник роликовый NNU4928B/SPW33	15	22 638,4	14 705,3	0,65	–	–	–
Подшипник 32028X FAG	20	19 754,4	14 043,4	0,71	экспоненциальный	1,4	0,23
Гидротормоз ORTLINGHAUS	6	47 168	21 779,70	0,46	–	–	–
Гидрораспределитель DHA	5	48 235,2	2 101,02	0,04	–	–	–
Подшипник 29403E	8	25 887	18 206,09	0,70	–	–	–
Индуктивный датчик NJ8-18GK-N-150	10	24 096	15 999,64	0,66	–	–	–
Ролик цилиндрический упорный	11	38 203,6	13 823,65	0,36	–	–	–
Подшипник FAG 32224A	7	37 937,20	12 852,65	0,33	–	–	–
Кольца седельные, шаровой затвор шарового крана	5	40 396,80	8548,87	0,21	–	–	–
СВП в период эксплуатации от 0 до 2 500 часов	26	1316,3	813,75	0,61	экспоненциальный	1	0,6
СВП в период эксплуатации от 2 500 до 4 500 часов	30	3546,4	517,57	0,14	нормальный	0,6	0,42

Примечание: составлено авторами.

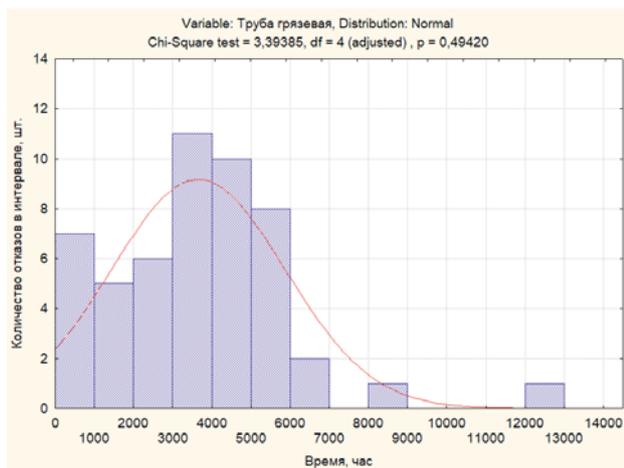


Рис. 5. Гистограмма наработок на отказ трубы грязевой
 Примечание: составлено авторами.

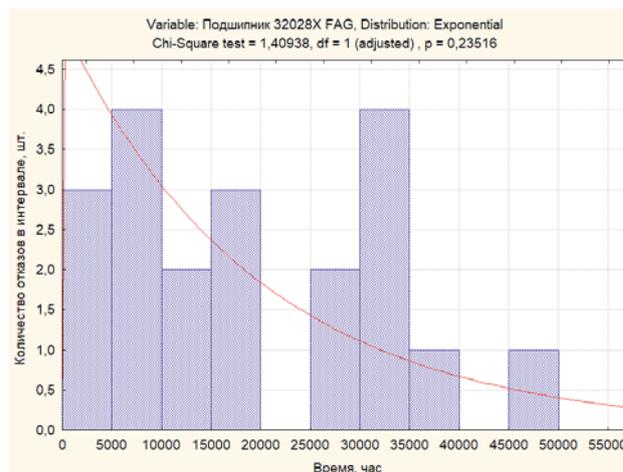


Рис. 6. Гистограмма наработок на отказ подшипника 32028X FAG
 Примечание: составлено авторами.

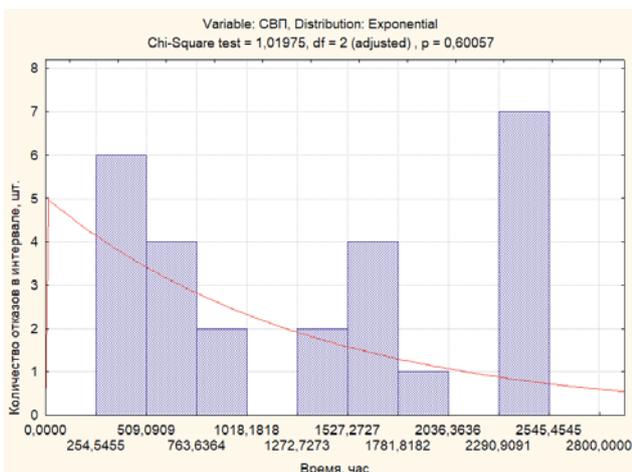


Рис. 7. Гистограмма наработок на отказ СВП до 2 500 часов эксплуатации
 Примечание: составлено авторами.

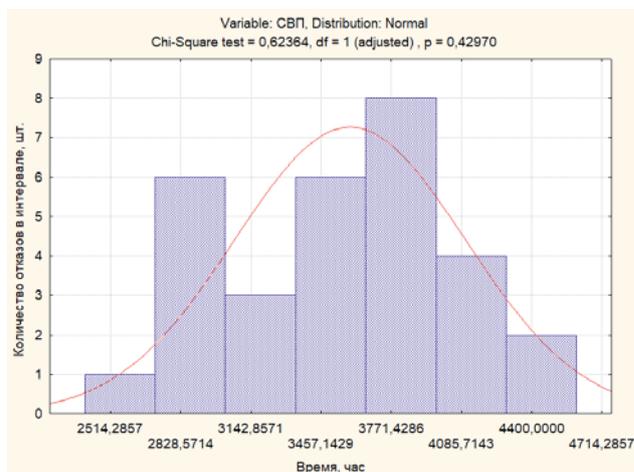


Рис. 8. Гистограмма наработок на отказ СВП от 2 500 до 4 500 часов эксплуатации
 Примечание: составлено авторами.

ных, будут иметь вероятностный характер. И с определенной вероятностью они могут оказаться неверными. Величина вероятности зависит от того, какой опасностью грозит ошибка; при проверке статистических гипотез полагают малой вероятностью, начиная с 0,05–0,01. Обычно при проверке статистических гипотез принимают три уровня значимости: 5 %-й (вероятность ошибочной оценки $\alpha = 0,05$), 1 %-й ($\alpha = 0,01$) и 0,1 %-й ($\alpha = 0,001$). В промышленной статистике считается достаточным 5 %-й уровень значимости. При этом нулевую гипотезу не отвергают, если в результате исследования окажется, что вероятность ошибочности оценки относительно правильности принятой превышает

5 %, т. е. $\alpha > 0,05$; если же $\alpha < 0,05$, то принятую гипотезу следует отвергнуть на взятом уровне значимости. Ошибка при этом возможна не более чем в 5 % случаев, т. е. она маловероятна. При более ответственных исследованиях уровень значимости может быть уменьшен до 1 % или даже до 0,1 %. В пакете Statistica значение задаваемого уровня значимости не используется, как правило, в выходных данных содержатся выборочные значения статистики критерия и вероятность того, что случайная величина превышает это выборочное значение при условии, что верна гипотеза [8–10].

При $n < 20$, формулы для обычной (большой) выборки дают значительные неточ-

ности [17]. Так, в табл. 2 представлены результаты обработки статистических данных с помощью программы Statistica из 14 элементов с $n < 20$, и эмпирические данные только двух из них не противоречат теоретическому экспоненциальному закону распределения с вероятностью $0,23 \leq P(\chi^2) \leq 0,77$ для подшипника В7032С и для подшипника 32028Х FAG в период эксплуатации от 2 500 до 4 500 часов. Эмпирические данные не противоречат теоретическому экспоненциальному закону распределения с вероятностью $P(\chi^2) = 0,6$ для СВП в период эксплуатации от 2 500 до 4 500 часов при $n = 26$.

Из табл. 2 видно, что эмпирические данные не противоречат теоретическому нормальному закону распределения с вероятностью $0,42 \leq P(\chi^2) \leq 0,49$ для трубы грязевой и СВП в период эксплуатации от 2 500 до 4 500 часов.

ЗАКЛЮЧЕНИЕ

Результаты анализа статистических данных по эксплуатации СВП, представленные на рис. 1–8 и в табл. 1 и 2, позволяют качественно оценить надежность СВП и его элементов.

Проанализировав полученные результаты возможно сделать вывод, что наиболее надежными из сравниваемых элементов являются следующие: гидрораспределитель ДНА, средняя наработка на отказ которого составляет 48 235,2 часа; кольца седельные с шаровым затвором шарового крана, средняя наработка на отказ которой составляет 40 396,80 часа; ролик цилиндрический упорный со средней наработкой на отказ 38 203,6 часа – такая высокая надежность обусловлена относительной простотой конструкции, заложенным при проектировании запасом прочности, а также высоким качеством изготовления. Наименее надежным элементом является труба грязевая, средняя наработка на отказ которой составляет 3 653,18 часа – низкая надежность

объясняется тем, что при выполнении работ данный элемент работает практически все время, имеет достаточно низкий коэффициент запаса по прочности, и нередки случаи низкого качества изготовления.

Отдельно выделяется достаточно высокая надежность гидравлического аксиально-поршневого насоса 90R180 со средней наработкой на отказ 38 035,2 часа. Данный насос обладает высокой конструктивной сложностью и является одним из основных рабочих элементов СВП, испытывающим постоянные знакопеременные динамические нагрузки. Высокая надежность гидравлического аксиально-поршневого насоса 90R180 объяснима только высоким качеством производства.

Рассмотренные выше законы распределения случайной величины могут характеризовать надежность СВП только лишь в определенные интервалы времени.

Так, в интервале времени «приработка» распределение наработки на отказ элемента или подсистемы не противоречит гамма-распределению или экспоненциальному закону, а в интервале времени «нормальная эксплуатация» – не противоречит закону распределения Вейбулла – Гнеденко, нормальному или логарифмически нормальному.

Если закон распределения случайной величины относится к функции двух случайных аргументов и эти случайные величины независимы, то появляется общий закон распределения случайной величины «композиция».

Композиционные законы распределения могут быть сочетаниями двух или более законов.

Проведенный расчет показателей надежности элементов СВП позволяет качественно оценить функциональную надежность элементов, что поможет принимать обоснованные решения для обеспечения требуемого уровня надежности при эксплуатации СВП на месторождениях Севера Сибири.

Список источников

1. Зернов В. И., Чичуа Д. А. Бурение горизонтальных скважин в Западной Сибири // Исследования молодых ученых : материалы XI Междунар. науч. конф., 20–23 июня 2020 г., г. Казань. Казань : Молодой ученый, 2020. С. 16–19.
2. Петрухин В. В., Коротков С. А., Черский М. И. Анализ систем верхнего привода для бурения горизонтальных скважин // Геология и нефтегазосносность Западно-Сибирского мегабассейна (опыт, инновации) : материалы X междунар. науч.-технич. конф., 24 ноября 2016 г., г. Тюмень. Т. 2. Тюмень : Тюменский индустриальный университет, 2016. С. 178–181.
3. Ткаченко Н. И., Башняк С. Е. Надежность технических систем и техногенный риск. Персиановский : Донской ГАУ, 2015. 60 с.
4. TDS-9SA Буровая система верхнего привода. 302 с. URL: https://electrorig.narod.ru/olderfiles/1/TOP_DRIVE_TDS_9SA_MANUAL_RUS.pdf (дата обращения: 07.09.2024).
5. Сызранцева К. В., Ильиных В. Н., Колбасин Д. С. Оценка вероятности отказа редуктора системы верхнего привода буровой установки с учетом режима его эксплуатации // Известия высших учебных заведений. Нефть и газ. 2019. № 3. С. 127–134.
6. Молчанов Г. В., Молчанов А. Г. Машины и оборудование для добычи нефти и газа. М. : Недра, 1984. 464 с.
7. Нефтегазодобывающее оборудование. URL: http://elmech.ru/product-catalog?mode=folder&folder_id=61069215 (дата обращения: 27.06.2024).
8. Куприенко Н. В., Пономарева О. А., Тихонов Д. В. Статистика. Методы анализа распределений. Выборочное наблюдение. 3-е изд. СПб. : Изд-во Политехн. ун-та, 2009. 138 с.
9. Стукач О. В. Программный комплекс Statistica в решении задач управления качеством. Томск : Изд-во Томского политехнического университета, 2011. 163 с.
10. Ходыревская С. В. Программные статистические комплексы. Курск : ЮЗГУ, 2020. 163 с.
11. Вентцель Е. С. Теория вероятностей. 10-е изд., стер. М. : Высш. шк., 2006. 575 с.
12. Ишемгузин И. Е., Атнагулов А. Р., Ишемгузин Е. И. Специальные критерии согласия для малой выборки при оценке надежности нефтегазодобывающего оборудования // Нефтегазовое дело. 2008. № 1. С. 28.
13. Острейковский В. А. Теория надежности. М. : Высш. шк., 2003. 463 с.
14. РД 50-690-89. Методические указания (надежность в технике). Методы оценки показателей надежности по экспериментальным данным. М. : Изд-во стандартов, 1990. 132 с.
15. Хэнли Э. Дж. Надежность технических систем и оценка риска / пер. с англ. В. С. Сыромятникова,

References

1. Zernov V. I., Chichua D. A. Burenie gorizontalnykh skvazhin v Zapadnoy Sibiri. In: *Proceedings of the XI International Scientific Conference "Issledovaniya molodykh uchenykh"*, June 20–23, 2020, Kazan. Kazan: Molodoy ucheny. 2020:16–19. (In Russ.).
2. Petrukhin V. V., Korotkov S. A., Cherskiy M. I. Analiz sistem verkhnego privoda dlya bureniya gorizontalnykh skvazhin. In: *Proceedings of the X international scientific and technical conference "Geologiya i neftegazonosnost Zapadno-Sibirskogo megabassey-na (opyt, innovatsii)"*, November 24, 2016, Tyumen. Tyumen: Tyumenskiy industrialnyy universitet. 2016;2:178–181. (In Russ.).
3. Tkachenko N. I., Bashnyak S. E. Nadezhnost tekhnicheskikh sistem i tekhnogenny risk. Persianovskiy: Donskoy GAU, 2015. 60 p. (In Russ.).
4. TDS-9SA Burovaya sistema verkhnego privoda. 302 p. URL: https://electrorig.narod.ru/olderfiles/1/TOP_DRIVE_TDS_9SA_MANUAL_RUS.pdf (accessed: 07.09.2024). (In Russ.).
5. Syzrantseva K. V., Il'inykh V. N., Kolbasin D. S. Assessing the failure probability of gearhead of drilling rig top drive system taking into account the mode of its operation. *Oil and gas studies*. 2019;(3):127–134. (In Russ.).
6. Molchanov G. V., Molchanov A. G. Mashiny i oborudovanie dlya dobychi nefiti i gaza. Moscow: Nedra, 1984. 464 p. (In Russ.).
7. Neftegazodobyvayushchee oborudovanie. URL: http://elmech.ru/product-catalog?mode=folder&folder_id=61069215 (accessed: 27.06.2024). (In Russ.).
8. Kuprienko N. V., Ponomareva O. A., Tikhonov D. V. Statistika. Metody analiza raspredeleniy. Vyborochnoe nablyudenie. 3d ed. Saint Petersburg: Izd-vo Politekhn. un-ta, 2009. 138 p. (In Russ.).
9. Stukach O. V. Programmnyy kompleks Statistica v reshenii zadach upravleniya kachestvom. Tomsk: Izd-vo Tomskogo politekhnicheskogo universiteta, 2011. 163 p. (In Russ.).
10. Khodyrevskaya S. V. Programmnye statisticheskie komplekсы. Kursk: YUZGU, 2020. 163 p. (In Russ.).
11. Venttsel E. S. Teoriya veroyatnostey. 10th ed. Moscow: Vyssh. shk., 2006. 575 p. (In Russ.).
12. Ishemguzhin I. E., Atmagulov A. R., Ishemguzhin E. I. Spetsialnye kriterii soglasiya dlya maloy vyborki pri otsenke nadezhnosti neftegazodobyvayushchego oborudovaniya. *Neftegazovoe delo*. 2008;(1):28. (In Russ.).
13. Ostreykovskiy V. A. Teoriya nadezhnosti. Moscow: Vyssh. shk., 2003. 463 p. (In Russ.).
14. RD 50-690-89. Metodicheskie ukazaniya (nadezhnost v tekhnike). Metody otsenki pokazateley nadezhnosti po eksperimentalnym dannym. Moscow: Izd-vo standartov, 1990. 132 p. (In Russ.).
15. Khenli E. Dzh. Nadezhnost tekhnicheskikh sistem i otsenka riska. Trans. by Syromyatnikov V. S.,

- Г. С. Деминой ; под общ. ред. В. С. Сыромятникова. М. : Машиностроение, 1984. 528 с.
16. Антонюк П. В., Острейковский В. А. Разработка и апробация методики расчета необходимого числа резервных мобильных буровых комплексов и их оборудования при эксплуатации на месторождениях Западной Сибири // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2021. № 12. С. 27–34.
17. Венецкий И. Г., Венецкая В. И. Основы математико-статистические понятия и формулы в экономическом анализе. 2-е изд., перераб. и доп. М. : Статистика, 1979. 448 с.
- Demina G. S.; Ed. Syromyatnikov V. S. Moscow: Mashinostroenie, 1984. 528 p. (In Russ.).
16. Antonyuk P. V., Ostreykovsky V. A. Development and testing of the methods for calculation of the necessary number of standby mobile drilling complexes and their equipment when operating in the fields of the Western Siberia. *Automation, telemechanization and communication in oil industry*. 2021;(12):27–34. (In Russ.).
17. Venetskiy I. G., Venetskaya V. I. Osnovy matematiko-statisticheskie ponyatiya i formuly v ekonomicheskom analize. 2nd ed., rev. Moscow: Statistika, 1979. 448 p. (In Russ.).

Информация об авторах

П. В. Антонюк – аспирант.

В. А. Острейковский – доктор технических наук, профессор.

А. В. Сорочкин – аспирант.

About the authors

P. V. Antonyuk – Postgraduate.

V. A. Ostreykovsky – Doctor of Sciences (Engineering), Professor.

A. V. Sorochkin – Postgraduate.

Научная статья

УДК 004.451

<https://doi.org/10.35266/1999-7604-2024-3-2>



Проектирование информационной системы для диагностики аппаратной части компьютера

Алексей Ильич Веревкин¹, Светлана Александровна Лысенкова²✉

^{1,2}Сургутский государственный университет, Сургут, Россия

¹alekseverevkin@yandex.ru

²lisa1108@mail.ru ✉, <https://orcid.org/0000-0003-1007-7610>

Аннотация. Рассмотрена прогнозируемая динамика продаж персональных компьютеров в период 2022–2030 гг., из чего сделан вывод о потенциальном объеме рынка. Проведен опрос среди экспертов в области обслуживания компьютерной техники. Результаты опроса были проанализированы с помощью языка программирования Python, после чего была построена показательная модель информационной системы для диагностики аппаратной части компьютера, которая будет использоваться в ходе разработки программы.

Ключевые слова: датчики персональных компьютеров, частота компонентов, анализ рынка, напряжение компонентов, температура компонентов

Для цитирования: Веревкин А. И., Лысенкова С. А. Проектирование информационной системы для диагностики аппаратной части компьютера // Вестник кибернетики. 2024. Т. 23, № 3. С. 16–21. <https://doi.org/10.35266/1999-7604-2024-3-2>.

Original article

Designing an information system for computer hardware diagnostics

Aleksey I. Verevkin¹, Svetlana A. Lysenkova²✉

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

¹alekseverevkin@yandex.ru

²lisa1108@mail.ru ✉, <https://orcid.org/0000-0003-1007-7610>

Abstract. The article examines the projected PC sales dynamics for the period 2022–2030, which allows estimating the potential market size. We conducted a survey among experts in computer hardware maintenance and analyzed the results using the Python programming language. The team analyzed the survey results and constructed an illustrative model of an information system for diagnosing computer hardware, which they will use during program development.

Keywords: PC sensors, component frequency, market analysis, component voltage, component temperature

For citation: Verevkin A. I., Lysenkova S. A. Designing an information system for computer hardware diagnostics. *Proceedings in Cybernetics*. 2024;23(3):16–21. <https://doi.org/10.35266/1999-7604-2024-3-2>.

ВВЕДЕНИЕ

В первом квартале 2023 г. в России было продано более 1,8 млн персональных компьютеров, что на 0,4 % больше, чем за аналогичный период 2022 г. [1]. Даже с учетом ухода с российского рынка множества иностранных

компаний-производителей электронной техники, наблюдается положительная динамика в количестве продаваемых персональных компьютеров [2].

Согласно прогнозам аналитиков, такой скачок в количестве продаваемых устройств свя-

зан со спросом, накопившимся за предыдущие года, поэтому, несмотря на возможный кратковременный спад в количестве продаваемых устройств, в ближайшие несколько лет динамика будет оставаться положительной (рис. 1) [2].

Как и любая другая электронная техника, персональные компьютеры (ПК) имеют свойство выходить из строя. Для предотвращения этого необходимо своевременное обслуживание оборудования, а в случае если неполадка возникла, нужно ее устранять.

Уход иностранных компаний с российского рынка, упомянутый ранее, нанес сильный удар как по сфере ремонта, так и по сфере обслуживания компьютерной техники. Возникло немалое количество проблем – недостаток компонентов для ремонта, отказ производителей операционных систем от деятельности на территории Российской Федерации [3], а также отсутствие лицензий для коммерческого использования программ, используемых в диагностике ПК [4].

В связи со сложившейся ситуацией на рынке, а также с политикой государства, направленной на импортозамещение во всех сферах жизни граждан, в частности сфере информационных технологий [5], образовалась ниша, которую необходимо заполнить.

МАТЕРИАЛЫ И МЕТОДЫ

В рамках проводимого исследования была сформирована экспертная группа из 16 человек, включавшая сотрудников различных компьютерных мастерских в Сургуте, а также сотрудников отделов технического обслуживания в крупных компаниях. Каждый из участников опроса был проинформирован об анонимности опроса и дал согласие на обработку данных в рамках научной статьи.

Всем экспертам был предложен опрос (рис. 2.1–2.2), который включал в себя вопросы касательно датчиков каждого компонента системы для того, чтобы определить, какую информацию должна предоставлять разрабатываемая система, а также вопросы, касающиеся операционных систем, с помощью которых возможно будет определить платформы, которым в первое время разработки необходимо отдать приоритет.

После того как эксперты были опрошены, результаты опроса были выгружены в Microsoft Excel. На основе этих данных, с помощью языка программирования Python и специализированной библиотеки (pandas), был проведен статистический анализ.

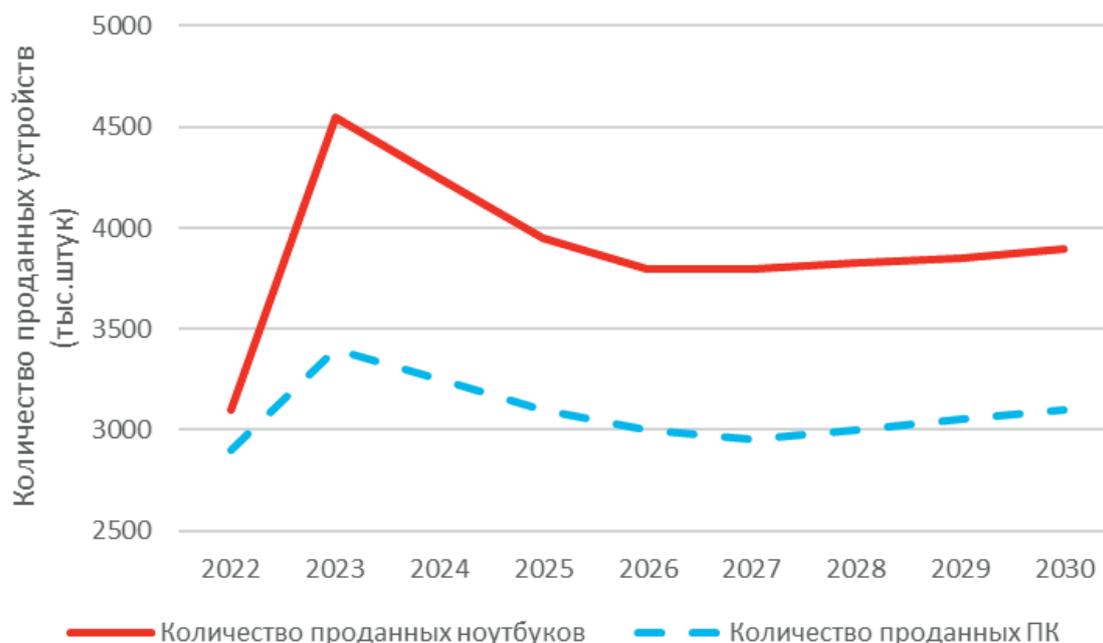


Рис. 1. Прогнозируемая динамика продаж персональных компьютеров и ноутбуков в период 2022–2030 гг.
Примечание: составлено авторами.

Введите название компании *

Мой ответ

Сколько компьютеров на ОС Windows Вы обслужили за последний месяц? *

Мой ответ

Сколько компьютеров на ОС macOS Вы обслужили за последний месяц? *

Мой ответ

Сколько компьютеров на ОС Linux Вы обслужили за последний месяц? *

Мой ответ

Рис. 2.1. Первая часть опроса
Примечание: составлено авторами.

Подсчитаны частоты выбора каждого ответа среди всех ответов на вопрос, путем нахождения отношения количества ответов, где есть выбранный пункт, к общему числу.

Код, используемый для обработки данных:
`import pandas as pd`

```
work_file = pd.read_excel("Опрос для ВКР (Ответы).xlsx", sheet_name=1)
for col_index in range(1, 5):
    col_values = work_file.iloc[:, col_index].astype(str)
    col_values = col_values.str.replace("Напряжение, ", "Напряжение ").str.replace("ядра, hot spot, видеопамяти", "ядра hot spot видеопамяти").str.replace(";", " ; ")
    unique = col_values.str.split("; ").explode().unique()
    values_counter = {value: col_values.str.contains(value).sum() for value in unique}
    sorted_values_counter = sorted(values_counter.items(), key=lambda x: x[1], reverse=True)
    for value, count in sorted_values_counter:
        print(f"{value} - {round(count/len(col_values) * 100, 2)} %")
    print(«-» * 30)
```

Сколько компьютеров на ОС, отличных от указанных выше, Вы обслужили за последний месяц? *

Мой ответ

Выберите, какие из показателей датчиков процессора Вы бы хотели видеть. *

Температура ядер

Частота работы ядер

Загруженность ядер и потоков

Вольтаж ядер

Напряжение, подающееся на процессор

Другое: _____

Выберите, какие из показателей датчиков видеокарты Вы бы хотели видеть. *

Температура ядра, hot spot, видеопамяти

Частота работы ядра и видеопамяти

Загруженность ядра и видеопамяти

Напряжение, подающееся на видеокарту

Другое: _____

Выберите, какие из показателей датчиков оперативной памяти Вы бы хотели видеть. *

Загруженность физической памяти

Объем используемой физической памяти

Загруженность виртуальной памяти

Объем используемой виртуальной памяти

Другое: _____

Выберите, какие из показателей датчиков материнской платы Вы бы хотели видеть. *

Температура

Вольтажи

Загруженность контроллеров вентиляторов

Скорость вращения вентиляторов

Другое: _____

Укажите, если хотели бы видеть еще какую-либо информацию о датчиках системы

Мой ответ

Рис. 2.2. Вторая часть опроса
Примечание: составлено авторами.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В результате проведения опроса удалось выявить наиболее предпочтительные для экспертов поля информации о персональном компьютере. Так, для процессора ранжирование интересующих числовых данных выглядит следующим образом:

1. Температура ядер – 93,33 %;
2. Загруженность ядер и потоков – 80,0 %;
3. Напряжение, подающееся на процессор – 73,33 %;
4. Частота работы ядер – 73,33 %;
5. Вольтаж ядер – 46,67 %.

Для видеокарты:

1. Температура ядра, hot spot, видеопамати – 100,0 %;
2. Напряжение, подающееся на видеокарту – 60,0 %;
3. Частота работы ядра и видеопамати – 46,67 %;
4. Загруженность ядра и видеопамати – 46,67 %.

Для оперативной памяти:

1. Загруженность физической памяти – 100,0 %;
2. Объем используемой физической памяти – 100,0 %;
3. Загруженность виртуальной памяти – 53,33 %;
4. Объем используемой виртуальной памяти – 53,33 %;
5. Вольтаж – 6,67 %.

Для материнской платы:

1. Скорость вращения вентиляторов – 86,67 %;
2. Вольтажи – 80,0 %;
3. Загруженность контроллеров вентиляторов – 73,33 %;
4. Температура – 46,67 %.

Запросов на какую-либо дополнительную информацию о системе, помимо перечисленной выше, поступило крайне мало, поэтому для получения более конкретных данных необходимо расширить экспертную группу.

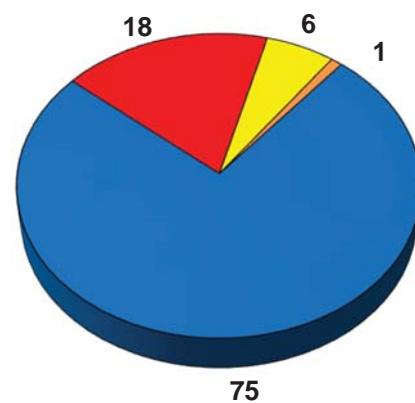
Для операционных систем также удалось выявить тенденцию – преобладающей операционной системой (ОС) является Windows – 75 %, далее идет macOS – 18 %,

Linux – 6 %, замыкают данный список прочие ОС, доля которых составила 1 % (рис. 3). Данное распределение устройств совпадает с информацией, предоставляемой разработчиками операционных систем: 72,47 % – Windows, 14,68 % – macOS, 8,79 % – прочие ОС, 4,05 % – Linux [6].

Как можно заметить, по всему миру процент прочих операционных систем составляет большую часть, чем Linux. Это можно объяснить тем, что есть продукты, которые не поставляются в РФ, однако они тоже учитываются в общей статистике [7]. Также данное явление можно объяснить массовым переходом государственных структур и крупных компаний в России на ОС Linux, в частности дистрибутив Astra Linux [8].

Такие точные выводы о результатах опроса позволяет сделать высокий коэффициент согласованности экспертной группы, который составляет 97 %.

Опираясь на данные, полученные в ходе анализа результатов опроса, можно построить показательную модель информационной системы. Данный вид моделей главным образом используется для объяснения и понимания терминов и процессов внутри темы. Она поможет в дальнейшем разработать саму программу, а также понимать сложные структуры и предсказывать их поведение. Данная модель приведена на рис. 4.



■ Windows ■ macOS ■ Linux ■ Прочие

Рис. 3. Распределение устройств по операционным системам

Примечание: составлено авторами.

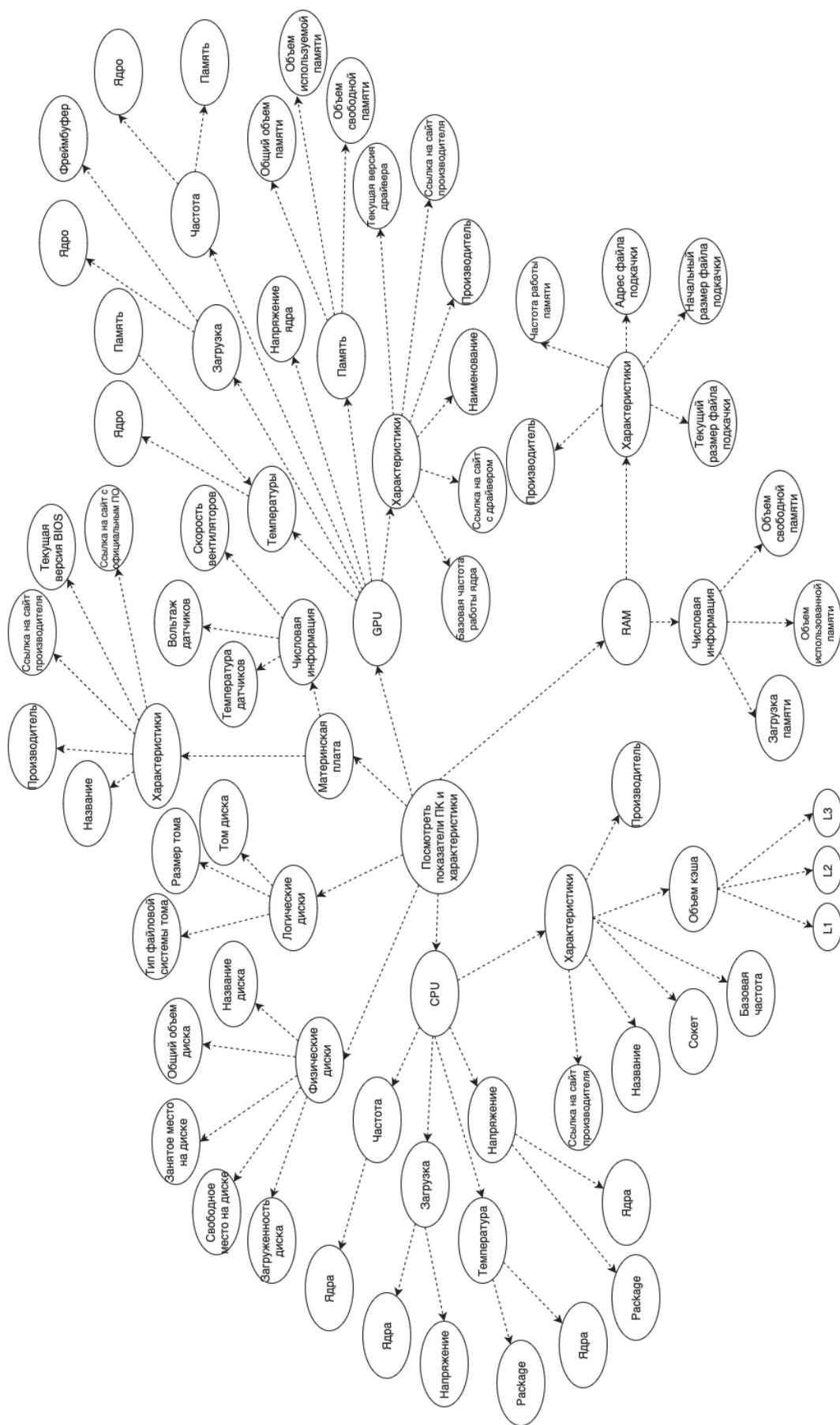


Рис. 4. Показательная модель информационной системы
 Примечание: составлено авторами.

ЗАКЛЮЧЕНИЕ

В результате проведенного исследования была проведена работа с потенциальными клиентами, от которых была получена обратная связь касательно как функционала разрабатываемой системы, так и операционных систем, с которыми

связан итоговый продукт. На основе анализа результатов опроса была построена пояснительная модель информационной системы с использованием языка универсального моделирования (UML). Данная диаграмма отражает связи в системе и служит для упрощения разработки.

Список источников

1. Компьютеры (рынок России). Рост продаж в первом квартале на 33 % до 1,8 млн ПК // TADVISER. Государство. Бизнес. Технологии : офиц. сайт. URL: <https://clck.ru/3A6YA6> (дата обращения: 10.06.2024).
2. Сколько в России Компьютеров? Знает ITResearch // Бестселлеры IT-рынка. Аналитика Российского рынка IT : офиц. сайт. URL: <https://www.itbestsellers.ru/statistics/detail.php?ID=53890> (дата обращения: 12.03.2024).
3. Кориненко Е. Монополия страны ОС: Windows на мировом рынке теснят конкуренты // Известия : офиц. сайт. URL: <https://iz.ru/1371116/ekaterina-korinenko/monopoliia-strany-os-windows-na-mirovom-rynke-tesniat-konkurenty> (дата обращения: 15.02.2024).
4. Алексеев Д. Отодвинуть за софт: как новый закон поможет использовать зарубежное ПО // Известия : офиц. сайт. URL: <https://iz.ru/1394309/dmitrii-alekseev/otodvinut-za-soft-kak-novyi-zakon-pomozhet-ispolzovat-zarubezhnoe-po> (дата обращения: 15.02.2024).
5. Бурсак А. Как бизнес учится обходиться без иностранных IT-решений // Ведомости. Технологии : офиц. сайт. URL: https://www.vedomosti.ru/technologies/innovation_policy/articles/2023/10/12/1000157-kak-biznes-uchitsya-obhoditsya-bez-inostrannih-it-reshenii (дата обращения: 17.04.2024).
6. Desktop Operating System Market Share Worldwide. StatCounter. URL: <https://gs.statcounter.com/os-market-share/desktop/worldwide> (дата обращения: 17.04.2024).
7. Козачинский А. Рынок хромбуков рухнул. В чем причина и почему это может быть хорошо для покупателей в РФ? URL: https://4pda.to/2021/11/03/392651/rynok_khrombukov_rukhnul_v_chyom_prichina_i_pochemu_eto_mozhet_byt_khorosho_dlya_pokupatelej_v_rf/ (дата обращения: 17.04.2024).
8. Плеханов А. Почему госструктуры переходят на операционную систему Astra Linux и чем она отличается от Windows. URL: <https://gol.ru/materials/19205-astra-linux> (дата обращения: 17.04.2024).

Информация об авторах

А. И. Веровкин – бакалавр.

С. А. Лысенкова – кандидат физико-математических наук, доцент.

References

1. Kompyutery (rynok Rossii). Rost prodazh v pervom kvartale na 33 % do 1,8 mln PK. TADVISER. Gosudarstvo. Biznes. Tekhnologii. URL: <https://clck.ru/3A6YA6> (accessed: 10.06.2024). (In Russ.).
2. Skolko v Rossii Kompyuterov? Znaet ITResearch. Bestsellery IT-rynka. Analitika Rossiyskogo rynka IT. URL: <https://www.itbestsellers.ru/statistics/detail.php?ID=53890> (accessed: 12.03.2024). (In Russ.).
3. Korinenko E. Monopoliya strany OS: Windows na mirovom rynke tesnyat konkurenty. Izvestiya. URL: <https://iz.ru/1371116/ekaterina-korinenko/monopoliia-strany-os-windows-na-mirovom-rynke-tesniat-konkurenty> (accessed: 15.02.2024). (In Russ.).
4. Alekseev D. Otodvinut za soft: kak novyy zakon pomozhet ispolzovat zarubezhnoe PO. Izvestiya. URL: <https://iz.ru/1394309/dmitrii-alekseev/otodvinut-za-soft-kak-novyi-zakon-pomozhet-ispolzovat-zarubezhnoe-po> (accessed: 15.02.2024). (In Russ.).
5. Bursak A. Kak biznes uchitsya obkhoditsya bez inostrannykh IT-resheniy. Vedomosti. Tekhnologii. URL: https://www.vedomosti.ru/technologies/innovation_policy/articles/2023/10/12/1000157-kak-biznes-uchitsya-obhoditsya-bez-inostrannih-it-reshenii (accessed: 17.04.2024). (In Russ.).
6. Desktop Operating System Market Share Worldwide. StatCounter. URL: <https://gs.statcounter.com/os-market-share/desktop/worldwide> (accessed: 17.04.2024).
7. Kozachinskiy A. Rynok khrombukov rukhnul. V chem prichina i pochemu eto mozhet byt khorosho dlya pokupatelej v RF? URL: https://4pda.to/2021/11/03/392651/rynok_khrombukov_rukhnul_v_chyom_prichina_i_pochemu_eto_mozhet_byt_khorosho_dlya_pokupatelej_v_rf/ (accessed: 17.04.2024). (In Russ.).
8. Plekhanov A. Pochemu gosstruktury perekhodyat na operatsionnyu sistemu Astra Linux i chem ona otlichaetsya ot Windows. URL: <https://gol.ru/materials/19205-astra-linux> (accessed: 17.04.2024). (In Russ.).

About the authors

A. I. Verevkin – Bachelor's Degree Student.

S. A. Lysenkova – Candidate of Sciences (Physics and Mathematics), Associate Professor.

Научная статья
УДК 004.412
<https://doi.org/10.35266/1999-7604-2024-3-3>



Сравнительный анализ метрик векторного расстояния растровых изображений

*Александр Александрович Детков*¹✉, *Валерия Алексеевна Воронина*²,
*Юлия Владимировна Гарифуллина*³, *Андрей Максимович Корепанов*⁴,
*Алина Юрьевна Вишнякова*⁵

^{1, 2, 3, 4, 5}Уральский федеральный университет имени первого Президента России Б. Н. Ельцина,
Екатеринбург, Россия

¹samik1982@mail.ru✉

³u.garifullina@yandex.ru

⁵Alina.vishniakova@urfu.ru

Аннотация. В статье описаны метрики и алгоритмы, позволяющие оценить различие или сходство следующих исследуемых объектов: расстояние Хэмминга, расстояние Левенштейна, сходство Джаро – Винклера, коэффициент Серенсена, SIFT, перцептивный хэш-алгоритм. С целью выявления алгоритма, оптимально подходящего для практического применения, а именно для оценки уникальности логотипов путем сравнения их друг с другом, метрики и алгоритмы были реализованы в виде программ на языке программирования Python и протестированы на предмет скорости и качества их работы. Наиболее быстро работающий алгоритм сравнения логотипов основан на расчете перцептивной хэш-суммы с последующим вычислением расстояния Хэмминга, наиболее точные результаты сравнения логотипов были получены в процессе работы программы, основанной на расчете расстояния Хэмминга. В качестве оптимального алгоритма, объединяющего в себе высокую скорость работы и качество полученных результатов, был выбран программно-реализованный алгоритм, основанный на расчете расстояния Хэмминга.

Ключевые слова: метрика, расстояние, изображение, сходство, различие, сравнение изображений, алгоритм, Python

Для цитирования: Детков А. А., Воронина В. А., Гарифуллина Ю. В., Корепанов А. М., Вишнякова А. Ю. Сравнительный анализ метрик векторного расстояния растровых изображений // Вестник кибернетики. 2024. Т. 23, № 3. С. 22–30. <https://doi.org/10.35266/1999-7604-2024-3-3>.

Original article

Comparative analysis of vector distance metrics for raster images

*Aleksandr A. Detkov*¹✉, *Valeriya A. Voronina*², *Yuliya V. Garifullina*³,
*Andrey M. Korepanov*⁴, *Alina Yu. Vishnyakova*⁵

^{1, 2, 3, 4, 5}Ural Federal University named after the First President of Russia B. N. Yeltsin,
Yekaterinburg, Russia

¹samik1982@mail.ru✉

³u.garifullina@yandex.ru

⁵Alina.vishniakova@urfu.ru

Abstract. The article describes metrics and algorithms that allow evaluating the difference or similarity of the objects under study: Hamming distance, Levenshtein distance, Jaro-Winkler similarity, Dice-Sørensen coefficient, SIFT, Perceptual Hash algorithm. To identify an algorithm that is optimally suitable for evaluating the uniqueness of logos by comparing them with each other, we implemented metrics and algorithms as programs in the Python

programming language. Then, we tested their speed and quality of work. The fastest working algorithm for comparing logos is based on the calculation of a perceptual hash sum, followed by the calculation of the Hamming distance. The most accurate results of comparing logos were obtained during the operation of the program based on the calculation of the Hamming distance. We chose a software-implemented algorithm based on Hamming distance calculation as the optimal algorithm that combines high speed of operation and the quality of the obtained results.

Keywords: metric, distance, image, similarity, difference, image comparison, algorithm, Python

For citation: Detkov A. A., Voronina V. A., Garifullina Yu. V., Korepanov A. M., Vishnyakova A. Yu. Comparative analysis of vector distance metrics for raster images. *Proceedings in Cybernetics*. 2024;23(3):22–30. <https://doi.org/10.35266/1999-7604-2024-3-3>.

ВВЕДЕНИЕ

Метрика – это функция, определяющая расстояние в метрическом пространстве; мера, определяющая расстояние между элементами множества [1]. Метрика определена свойствами: больше или равна нулю; равна нулю при условии, что элементы множества совпадают; симметричная; подчиняется закону треугольника. Введение метрики на множестве определяет метрическое пространство [2]. Метрическое пространство – это множество, в котором определено расстояние между любой парой элементов [3].

Понятие расстояния введено для количественного измерения сходства и различия двух последовательностей. Мера расстояния – функция, присваивающая численное значение паре последовательностей: чем больше расстояние, тем меньше подобие [4].

В настоящее время известно множество метрик, позволяющих оценить различие или сходство исследуемых объектов, например, расстояние Хэмминга, расстояние Левенштейна, сходство Джаро – Винклера, коэффициент Серенсена, каскад Хаар и др.

Метрики сходства или различия широко используются и применяются для решения различных задач – распознавания образов, классификации документов, исправления ошибок, обнаружения вирусов, обнаружения дубликатов и т. д.

Так, метрики сходства или различия применимы для такой предметной области как «интеллектуальная собственность», в узком смысле «правовая охрана средств индивидуализации». Средства индивидуализации – обозначения, при помощи которых коммерческие фирмы и их продукция, а именно

товары и услуги, идентифицируются среди множества других подобных, приобретают уникальность. Правовая охрана интеллектуальной собственности, согласно Статье 1232 ГК РФ, предоставляется на основе государственной регистрации, которую осуществляет федеральный орган исполнительной власти Роспатент [5]. По статистике, около 30 % поданных заявок на регистрацию обозначения получают отказ, на основании наличия тождественных или схожих до степени смешения товарных знаков, зарегистрированных ранее. Проведение предварительной проверки обозначения на предмет наличия схожих или тождественных товарных знаков, зарегистрированных ранее, позволит снизить риски отказа в его государственной регистрации. Использование метрик позволит автоматизировать процесс проверки изобразительного обозначения: уменьшить сроки, повысить качество проверки и минимизировать человеческое влияние на результат.

Целью данного исследования является разработка и оценка эффективности и скорости работы программ, позволяющих выявить различие или сходство двух изобразительных обозначений между собой, путем расчета метрик измерения расстояния между двумя последовательностями.

Задачи исследования:

1. Поиск и анализ научных статей на тему расчета метрик измерения расстояний между последовательностями, их применения для решения различных задач и возможности их программной реализации.

2. Поиск и изучение существующих метрик, измеряющих расстояния между последовательностями.

3. Анализ метрик с точки зрения их применимости для сравнения изобразительных обозначений.

4. Программная реализация алгоритма расчета метрик различия или схожести изобразительных обозначений.

5. Проведение испытаний и оценка эффективности и скорости работы программ при сравнении изобразительных обозначений, основанных на расчете метрик различия или схожести.

МАТЕРИАЛЫ И МЕТОДЫ

Источниками информации в области товарных знаков и их государственной регистрации выступили Руководство по осуществлению административных процедур и действий в рамках предоставления государственной услуги по регистрации товарного знака [6] и четвертая часть Гражданского кодекса Российской Федерации [5]. В качестве источника информации о метриках измерения расстояния между последовательностями были использованы научные труды российских и зарубежных ученых. Для поиска и проведения литературного обзора научных статей были использованы ресурсы поисковой системы eLIBRARY, поиск проведен по вышеуказанным ключевым словам. Для поиска и анализа программных кодов, реализующих расчет метрик расстояния, были использованы веб-сервисы, позволяющие просматривать исходные коды программ [7]. Для оценки применимости и эффективности использования различных метрик расстояния между последовательностями были рассчитаны следующие величины: расстояние Хэмминга, расстояние Левенштейна, расстояние Джаро – Винклера, мера Серенсена.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В ходе проведения исследования была найдена и рассмотрена статья: «Актуальные задачи анализа изображений», подтверждающая актуальность выбранной темы. В статье авторы (А. И. Либерман и Т. Э. Шульга) представляют обзор современного состояния интеллектуального анализа изобра-

жений, рассматривают сферы применения анализа изображений и анализируют методы искусственного интеллекта для решения различных практических задач. Авторы приходят к выводу о возможности комбинации методов анализа изображений и онтологического моделирования с целью дальнейшего развития области анализа изображений [8].

Е. Ю. Коломийцева в статье «Анализ изображений в интернете: возможности современных сервисов» рассматривает анализ изображений с помощью искусственного интеллекта с точки зрения применимости и ценности в области журналистской работы. Автор приводит примеры параметров изображений, которые в настоящее время можно анализировать с использованием современных сервисов, например, такие параметры как определение шрифта, определение того, что изображено на фото, распознавание эмоций человека, определение личности человека и другое [9].

В статье «Анализ изображений. Определение дубликатов и степени размытости изображений», автором которой является Д. В. Шестоперов, было предложено решение проблемы затрудненного поиска фотографий в большом файловом хранилище, не имеющем фильтрации – анализ изображений и удаление неподходящих фотографий с использованием алгоритма разностного хеширования и сравнения хэшей при помощи расстояния Хэмминга для поиска дубликатов изображений [10].

Л. К. Ситниковой, М. О. Еланцевым и Р. О. Султановым в статье «Двухэтапный метод поиска изображений в большой базе изображений» были зафиксированы результаты анализа способов поиска в базе данных большого размера. Авторами был сделан вывод о том, что совместное применение метода сравнения по перцептивному хэшу и метода классификации при помощи нейронной сети является наиболее эффективным [11].

После рассмотрения научных статей и публикаций были сделаны выводы об актуальности исследуемой темы и определены

наиболее часто используемые для анализа изображений метрики измерения расстояний между последовательностями.

Метрики и алгоритмы, позволяющие оценить различие или сходство исследуемых объектов:

1. Расстояние Хэмминга – метрика различия объектов одинаковой размерности; метрика измерения расстояния между двумя последовательностями. Метрикой для определения меры различия между двоичными векторами в векторном пространстве кодовых последовательностей – расстоянием Хэмминга $d(x; y)$ между двумя двоичными последовательностями – векторами: x и y , длины n – называется число позиций, в которых они различны. В общем виде расстояние Хэмминга для объектов и размерностей задается функцией [12]:

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}|. \quad (1)$$

2. Расстояние Левенштейна – метрика сходства между двумя последовательностями; метрика измеряющая по модулю разность между двумя последовательностями символов. Расстояние Левенштейна определяется как минимальное количество односимвольных операций – вставки, удаления, замены, необходимых для превращения одной последовательности элементов в другую. Пусть S_1 и S_2 – две строки некоторого алфавита, длины которых M и N соответственно, тогда расстояние Левенштейна $d(S_1, S_2)$, можно рассчитать по формуле:

$$D(i, j) = \begin{cases} d(S_1, S_2) = D(M, N), \text{ где} \\ 0, & i = 0, j = 0 \\ i, & j = 0, i > 0 \\ \min\{ \\ D(i, j - 1) + 1, \\ D(i - 1, j) + 1, & j = 0, i > 0 \\ D(i - 1, j - 1) + m(S_1[i], S_2[j]) \\ \} \end{cases} \quad (2)$$

где $m(a, b)$ равна нулю, если $a = b$ и единице в противном случае $\min\{a, b, c\}$ возвращает наименьшее из аргументов. Шаг по i символизирует удаление из первой строки, по

j – вставку в первую строку, а шаг по обоим индексам символизирует замену символа или отсутствие изменений [13].

3. Сходство Джаро – Винклера – мера схожести для измерения расстояния между двумя последовательностями символов.

Расстояние Джаро – Винклера – это минимальное число односимвольных преобразований, которое необходимо для того, чтобы изменить одно слово в другое. Чем меньше расстояние Джаро – Винклера d_w для двух последовательностей, тем больше сходства имеют эти последовательности друг с другом. Результат нормируется так, что $d_w = 0$ – отсутствие сходства, $d_w = 1$ – точное сходство. Сходство Джаро – Винклера равно $1 - d_w$.

При расчете расстояния Джаро – Винклера используется коэффициент масштабирования p , что позволяет присвоить более благоприятные рейтинги строкам, которые совпадают друг с другом от начала до определенной длины l , которая называется префиксом.

Пусть S_1 и S_2 – две строки, расстояние Джаро – Винклера d_w можно рассчитать по формуле:

$$d_w = d_j + (l_p (1 - d_j)), \quad (3)$$

где d_j – расстояние Джаро для строк S_1 и S_2 ;

l – длина общего префикса от начала строки до максимума четырех символов;

p – постоянный коэффициент масштабирования, использующийся для того, чтобы скорректировать оценку в сторону повышения для выявления наличия общих префиксов: $p = const = 0,1$ [14].

4. Коэффициент Серенсена – бинарная мера сходства. Коэффициент Серенсена создан для определения схожести любых множеств. Коэффициент Серенсена K_s рассчитывается по формуле:

$$K_s = \frac{2c}{a + b}, \quad (4)$$

где a – количество видов первой экосистемы;

b – количество видов второй экосистемы;

c – количество общих для 1-й и 2-й экосистемы видов [15].

5. Алгоритм SIFT (масштабно-инвариантная трансформация признаков) – алгоритм выявления признаков в компьютерном зрении для выявления и описания локальных признаков в изображениях.

Первостепенно в алгоритме SIFT извлекаются ключевые точки объектов из набора изображений и запоминаются в базе данных. Заданный объект в новом изображении распознается путем сравнения признаков из нового изображения с признаками из базы данных и нахождения признаков-кандидатов на основе евклидова расстояния между векторами признаков. Ключевые точки отбираются из набора соответствий в новом изображении, на основании наиболее близкого сходства с объектом по его местоположению, масштабу и ориентации. Подходящие блоки признаков определяются с помощью реализации хэш-таблицы обобщенного преобразования Хафа, каждый согласующийся с объектом блок из трех или более признаков подлежит дальнейшей подробной проверке. Резко отклоняющиеся блоки признаков отбрасываются. Определяется набор признаков, который с высокой вероятностью подтвердит наличие искомого объекта на рассматриваемом изображении [16].

6. Перцептивный хэш-алгоритм. Хэш-алгоритмы основаны на генерации сравнимых хэшей. Данные, используемые для генерации хэша, выполняют роль источника случайных чисел – одинаковые данные дадут одинаковый результат, разные данные – разный результат. Характеристики изображений используются для генерации индивидуальных отпечатков, которые можно сравнивать друг с другом. Выводы, которые можно сделать на основании сравнения хэшей: хэши отличаются – данные разные; хэши совпадают – данные одинаковые. Изменение размера, соотношения сторон и цветовых характеристик (яркость, контраст и т. д.) сравниваемых изображений не повлияет на результат вычисления перцептивного хэша.

Этапы получения перцептивного хэша:

– уменьшение размера рассматриваемого изображения;

– перевод изображения в градации серого;
– вычисление среднего значения яркости получившегося изображения;

– бинаризация изображения, выборка пикселей, значение которых больше среднего;

– построение хэша, перевод полученных отдельных значений в одно 64-битное значение [17].

На основании рассмотренных выше методов были разработаны алгоритмы для сравнения логотипов с целью проверки их уникальности. Основой для алгоритмов выступили шесть методов: расстояние Хэмминга, расстояние Левенштейна, сходство Джаро – Винклера, коэффициент Серенсена, SIFT и метод перцептивной хэш-суммы с последующим вычислением расстояния Хэмминга для сравнения. Методы реализованы на языке программирования Python с использованием библиотек: OpenCV, NumPy, Editdistance, Strsimpy и ImageHash.

Для реализации алгоритма, основанного на расчете расстояния Хэмминга, использована формула расчета количества несовпадающих битов между двумя бинарными строками с применением библиотек: OpenCV, NumPy.

Расстояние Левенштейна было реализовано с помощью алгоритма Дамерау – Левенштейна, который позволяет находить минимальное количество операций редактирования, необходимых для превращения одной строки в другую, реализован с применением библиотек: Editdistance, NumPy.

Для сравнения строк на основе сходства Джаро – Винклера использована формула, которая учитывает совпадение символов в строках и их расстояние друг от друга, реализован с помощью библиотек: Strsimpy, NumPy.

Алгоритм, основанный на расчете коэффициента Серенсена, реализован как отношение числа общих символов в строках к общему числу символов, с применением библиотек: Strsimpy, NumPy.

Метод SIFT реализован с помощью построения гауссовых пирамид и поиска экстремумов на разных уровнях пирамиды, а также определения градиентов и ориентаций угло-

вых точек на изображении с использованием библиотеки OpenCV.

Перцептивный хэш-алгоритм был реализован путем вычисления хэш-суммы на основе яркости каждого пикселя и дальнейшего вычисления расстояния Хэмминга между хэш-суммами двух изображений, с использованием библиотеки ImageHash.

Разработанные алгоритмы протестированы путем сравнения заранее подготовленных тестовых логотипов между собой. Проверка работы алгоритмов была нацелена на получение результатов вывода программ, а также скорости их обработки.

Разработанными программами были обработаны четыре пары логотипов, размером 1024×1024 пикселя:

- 1) полностью черный/полностью белый (далее – черно-белые);
- 2) существующий логотип/его инвертированная версия (далее – инвертированные);
- 3) существующий логотип/точно такой же существующий логотип (далее – идентичные);
- 4) существующий логотип/другой существующий логотип (далее – разные).

Результаты тестирования представлены в табл. 1. Сходство логотипов изменяет-

ся в процентах, где 100 % – полное сходство, 0 % – отсутствие сходства.

Результаты работы программ подтверждают корректную работу алгоритмов: черно-белые логотипы – 0 % сходства, идентичные алгоритмы – 100 % сходства. Результаты сравнения инвестированной пары логотипов: расстояние Хэмминга – 0 % сходства, SIFT – 4 %, расстояние Левенштейна – 17 %, коэффициент Серенсена – 23 %, сходство Джаро – Винклера – 33 %, перцептивный хэш-алгоритм – 38 %. Результаты сравнения разных логотипов существенно отличаются: расстояние Хэмминга, расстояние Левенштейна и хэш-сумма – 60–70 % сходства, сходство Джаро – Винклера – 87 %, коэффициент Серенсена – 13 %, SIFT – 1 % сходства.

Результаты тестирования программ на предмет скорости их работы представлены в табл. 2.

Из результатов тестирования программ на скорость их работы видно, что в основе самого быстроработающего алгоритма лежит перцептивный хэш-алгоритм с последующим вычислением расстояния Хэмминга, далее по скорости работы идут метод SIFT и коэффициент Серенсена, после – расстояние Хэм-

Таблица 1

Результаты тестирования

Типы изображений	Метрики векторного расстояния					
	Расстояние Хэмминга	Расстояние Левенштейна	Сходство Джаро – Винклера	Коэффициент Серенсена	SIFT	Хэш-сумма
Черный/белый	0 %	0 %	0 %	0 %	0 %	0 %
Инвертированный	0 %	17 %	33 %	23 %	4 %	38 %
Идентичный	100 %	100 %	100 %	100 %	100 %	100 %
Разный	60 %	62 %	87 %	13 %	1 %	70 %

Примечание: составлено авторами.

Таблица 2

Результаты тестирования

Время выполнения метрик	Метрики векторного расстояния					
	Расстояние Хэмминга	Расстояние Левенштейна	Сходство Джаро – Винклера	Коэффициент Серенсена	SIFT	Хэш-сумма
Скорость обработки данных, с	1,3	1 281	45 108	1,08	0,249	0,212

Примечание: составлено авторами.

минга, расстояние Левенштейна и сходство Джаро – Винклера.

На основании проделанной работы и полученных результатов можно сделать вывод о том, что алгоритм, основанный на расчете расстояния Хэмминга, является оптимальным для сравнения логотипов с целью оценки их уникальности: высокая скорость обработки данных и точность результата сравнения. Неприменимой на практике является программа, основанная на расчете сходства Джаро – Винклера, скорость работы которой низкая, а качество результата работ невысокое.

ЗАКЛЮЧЕНИЕ

В ходе проведения исследования были проанализированы научные статьи, тематика которых связана с математическими метриками, позволяющими измерить расстояние между последовательностями, с применением данных метрик для решения различных задач, а также с возможностью программной реализации алгоритмов расчета расстояния между последовательностями. Обзор научных статей подтвердил актуальность темы исследования, позволил определить и изучить часто используемые в ходе анализа математиче-

ские метрики: расстояние Хэмминга, расстояние Левенштейна, сходство Джаро – Винклера, коэффициент Серенсена, алгоритм SIFT и перцептивный хэш-алгоритм.

На основании изученного материала был реализован программный алгоритм расчета метрик различия или сходства изобразительных обозначений. Для каждой метрики были применены собственные алгоритмы и формулы в целях их реализации.

Проведены испытания и оценка эффективности и скорости работы программ при сравнении изобразительных обозначений, основанных на расчете метрик различия или схожести путем сравнения заранее подготовленных тестовых логотипов между собой. В ходе проверки были получены результаты, проанализирована скорость их обработки, на основе чего была разработана программа, позволяющая выявить различие или сходство двух изобразительных обозначений между собой путем расчета метрик измерения расстояния между двумя последовательностями с помощью алгоритма, основанного на расчете расстояния Хэмминга, позволившего достичь целевого результата соотношения скорости и сходства логотипов.

Список источников

1. Большой энциклопедический словарь. 2-е изд., перераб. и доп. М. ; СПб. : Большая российская энциклопедия, Норинт, 2000. 1434 с.
2. Определение метрики и ее свойства. Шкалы: номинальная, порядковая, интервальная, относительная. URL: <https://studfile.net/preview/8985938/page:7/> (дата обращения: 04.05.2024).
3. Битюцков В. И., Войтсеховский М. И., Иванов А. Б. и др. Математическая энциклопедия. М. : Советская Энциклопедия, 1984. 1246 с.
4. Метрика (математика). URL: [https://ru.wikibrief.org/wiki/Metric_\(mathematics\)](https://ru.wikibrief.org/wiki/Metric_(mathematics)) (дата обращения: 04.05.2024).
5. Гражданский кодекс Российской Федерации (часть четвертая) от 18.12.2006 № 230-ФЗ. Доступ из СПС «КонсультантПлюс».
6. Об утверждении Руководства по осуществлению административных процедур и действий в рамках предоставления государственной услуги по государственной регистрации товарного знака, знака обслуживания, коллективного знака и выдаче свидетельств на товарный знак, знак обслуживания,

References

1. Bolshoy entsiklopedicheskiy slovar. 2nd ed., revised. Moscow; Saint Petersburg: Bolshaya rossiyskaya entsiklopediya, Norint, 2000. 1434 p. (In Russ.).
2. Opredelenie metriki i ee svoystva. Shkaly: nominalnaya, poryadkovaya, intervalnaya, otnositelnaya. URL: <https://studfile.net/preview/8985938/page:7/> (accessed: 04.05.2024). (In Russ.).
3. Bityutskov V. I., Voytsekhovskiy M. I., Ivanov A. B. et al. Matematicheskaya entsiklopediya. Moscow: Sovetskaya Entsiklopediya, 1984. 1246 p. (In Russ.).
4. Metric (mathematics). URL: [https://ru.wikibrief.org/wiki/Metric_\(mathematics\)](https://ru.wikibrief.org/wiki/Metric_(mathematics)) (accessed: 04.05.2024). (In Russ.).
5. Civil Code of the Russian Federation (Pt. 4) of 18.12.2006 No. 230 Federal Law. Accessed through Law assistance system “Consultant Plus”. (In Russ.).
6. Ob utverzhdenii Rukovodstva po osushchestvleniyu administrativnykh protsedur i deystviy v ramkakh predostavleniya gosudarstvennoy uslugi po gosudarstvennoy registratsii tovarnogo znaka, znaka obsluzhivaniya, kollektivnogo znaka i vydache svidetelstv na tovarnyy znak, znak obsluzhivaniya, kollektivnyy znak,

- коллективный знак, их дубликатов : приказ ФГБУ ФИПС 20.01.2020 № 12 (ред. от 25.03.2022). URL: <https://legalacts.ru/doc/prikaz-fgbu-fips-ot-20012020-n-12-ob-utverzhenii/> (дата обращения: 04.05.2024).
7. GitHub. URL: <https://github.com/?ysclid=lhjgyylis6v237864613> (дата обращения: 04.05.2024).
 8. Либерман А. И., Шульга Т. Э. Актуальные задачи анализа изображений // Проблемы управления в социально-экономических и технических системах : сб. науч. ст. XVIII Междунар. науч.-практич. конф., 14–15 апреля 2022 г., г. Саратов. Саратов : Наука, 2022. С. 41–46.
 9. Коломийцева Е. Ю. Анализ изображений в Интернете: возможности современных сервисов // Медиа в современном мире. 61-е Петербургские чтения : ст. участников ежегодного апрельского научного форума, 21–22 апреля 2022 г., Санкт-Петербург. Т. 2. СПб. : Медиапир, 2022. С. 66–67.
 10. Шестоперов Д. В. Анализ изображений. Определение дубликатов и степени размытости изображений // Вопросы устойчивого развития общества. 2022. № 5. С. 1089–1096.
 11. Ситникова Л. К., Еланцев М. О., Султанов Р. О. Двухэтапный метод поиска изображений в большой базе изображений // Выставка Инноваций – 2022 (весенняя сессия) : сб. материалов XXXIII Республиканской выставки-сессии студенческих инновационных проектов, 29 апреля 2022 г., г. Ижевск. Ижевск : Ижевский государственный технический университет имени М. Т. Калашникова, 2022. С. 304–309.
 12. Питерсон У. У., Уэлдон Э. Коды, исправляющие ошибки / пер. с англ., под ред. Р. Л. Добрушина, С. И. Самойленко. М. : Мир, 1976. 594 с.
 13. Метод Левенштейна. URL: https://spravochnick.ru/informatika/metod_levenshteyna/?ysclid=li96moyh99224030240 (дата обращения: 10.05.2024).
 14. Brinardi L., Seng H. Text documents plagiarism detection using Rabin Karp and Jaro-Winkler Distance algorithms // Indonesian Journal of Electrical Engineering and Computer Science. 2017. Vol. 5, no. 2. P. 462–471.
 15. Трасс Х. Х. Геоботаника: история и современные тенденции развития. Л. : Наука, 1976. 252 с.
 16. Построение SIFT дескрипторов и задача сопоставления изображений. URL: <https://habr.com/ru/articles/106302/> (дата обращения: 10.05.2024).
 17. Как бороться с репостами или пара слов о перцептивных хешах. Перцептивные хэш-алгоритмы. URL: <https://habr.com/ru/articles/237307/> (дата обращения: 10.05.2024).
 - ikh dublikatov: prikaz FGBU FIPS 20.01.2020 No. 12 (ed. 25.03.2022). URL: <https://legalacts.ru/doc/prikaz-fgbu-fips-ot-20012020-n-12-ob-utverzhenii/> (accessed: 04.05.2024). (In Russ.).
 7. GitHub. URL: <https://github.com/?ysclid=lhjgyylis6v237864613> (accessed: 04.05.2024).
 8. Liberman A. I., Shulga T. E. Aktualnye zadachi analiza izobrazheniy. In: *Collection of scientific Articles of the XVIII International scientific-practical conference “Problemy upravleniya v sotsialno-ekonomicheskikh i tekhnicheskikh sistemakh”*, April 14–15, 2022, Saratov. Saratov: Nauka, 2022. P. 41–46. (In Russ.).
 9. Kolomiytseva E. Yu. Analiz izobrazheniy v Internetе: vozmozhnosti sovremennykh servisov. In: *Articles of the participants of the annual April scientific forum “Media v sovremennom mire. 61-e Peterburgskie chteniya”*, April 21–22, 2022, Saint Petersburg. Vol. 2. Saint Petersburg: Mediapir, 2022. P. 66–67. (In Russ.).
 10. Shestoperov D. V. Image analysis. Identification of blur and duplicates of images. *Voprosi ustoychivogo razvitiya obshchestva*. 2022;(5):1089–1096. (In Russ.).
 11. Sitnikova L. K., Elantcev M. O., Sultanov R. O. A two-step method of searching for similar images in a large image database. In: *materials of the XXXIII Republican exhibition-session of student innovation projects “Vystavka Innovatsiy – 2022 (vesenniyaya sessiya)”*, April 29, 2022, Izhevsk. Izhevsk: Izhevskiy gosudarstvennyy tekhnicheskii universitet imeni M. T. Kalashnikova, 2022. P. 304–309. (In Russ.).
 12. Piterson U. U., Ueldon E. Kody, ispravlyayushchie oshibki. Trans. Eds. R. L. Dobrushin, S. I. Samoylenko. Moscow: Mir, 1976. 594 p. (In Russ.).
 13. Metod Levenshteyna. URL: https://spravochnick.ru/informatika/metod_levenshteyna/?ysclid=li96moyh99224030240 (accessed: 10.05.2024). (In Russ.).
 14. Brinardi L., Seng H. Text documents plagiarism detection using Rabin Karp and Jaro-Winkler Distance algorithms. *Indonesian Journal of Electrical Engineering and Computer Science*. 2017;5(2):462–471.
 15. Trass H. Vegetation science. History and contemporary trends of development. Leningrad: Nauka, 1976. 252 p. (In Russ.).
 16. Postroenie SIFT deskriptorov i zadacha sopostavleniya izobrazheniy. URL: <https://habr.com/ru/articles/106302/> (accessed: 10.05.2024). (In Russ.).
 17. Kak borotsya s repostami ili para slov o pertseptivnykh khashakh. Pertseptivnye khash-algoritmy. URL: <https://habr.com/ru/articles/237307/> (accessed: 10.05.2024). (In Russ.).

Информация об авторах

А. А. Детков – кандидат экономических наук, доцент.

В. А. Воронина – студент.

Ю. В. Гарифуллина – студент.

А. М. Корепанов – студент.

А. Ю. Вишнякова – старший преподаватель.

About the authors

A. A. Detkov – Candidate of Sciences (Economics), Associate Professor.

V. A. Voronina – Student.

Yu. V. Garifullina – Student.

A. M. Korepanov – Student.

A. Yu. Vishnyakova – Senior Lecturer.

Научная статья
УДК 004.738.52:004.434
<https://doi.org/10.35266/1999-7604-2024-3-4>



Краткое аннотирование новостей из области информационных технологий

Кирилл Владимирович Кузнецов¹, Светлана Александровна Лысенкова²✉

^{1,2}Сургутский государственный университет, Сургут, Россия

¹kirill.kuznetsov.v@icloud.com, <https://orcid.org/0009-0007-3401-1962>

²lysenkova_sa@surgu.ru✉, <https://orcid.org/0000-0003-1007-7610>

Аннотация. В статье рассматривается концепция веб-платформы, созданной для сбора и аннотирования новостей в сфере информационных технологий. Платформа собирает данные из различных источников, предлагая пользователям удобный интерфейс для поиска и чтения новостных материалов. Одной из ключевых особенностей системы является возможность классификации статей, написание аннотаций и присвоение тегов при использовании больших языковых моделей (LLM), а также автоматического перевода статей с оригинального языка на другие.

Ключевые слова: веб-платформа, классификация информации, большие языковые модели, аннотирование новостей, сбор данных, контейнеризация

Для цитирования: Кузнецов К. В., Лысенкова С. А. Краткое аннотирование новостей из области информационных технологий // Вестник кибернетики. 2024. Т. 23, № 3. С. 31–39. <https://doi.org/10.35266/1999-7604-2024-3-4>.

Original article

Summary of news in information technology

Kirill V. Kuznetsov¹, Svetlana A. Lysenkova²✉

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

¹kirill.kuznetsov.v@icloud.com, <https://orcid.org/0009-0007-3401-1962>

²lysenkova_sa@surgu.ru✉, <https://orcid.org/0000-0003-1007-7610>

Abstract. The article discusses the concept of a web platform designed for collecting and summarizing news in information technology. The platform gathers data from various sources, offering users a convenient interface for searching and reading news materials. The key features of the system are the ability to classify articles, make summaries, and assign tags using large language models (LLM), as well as the automatic translation of articles into other languages.

Keywords: web platform, information classification, large language models, news summary, data collection, OS-level virtualization

For citation: Kuznetsov K. V., Lysenkova S. A. Summary of news in information technology. *Proceedings in Cybernetics*. 2024;23(3):31–39. <https://doi.org/10.35266/1999-7604-2024-3-4>.

ВВЕДЕНИЕ

Ежедневно пользователи сталкиваются с тысячами новостных материалов, связанных с информационными технологиями. В таких условиях стремительного развития сложно оставаться в курсе последних новостей и тенденций.

Прогнозируется, что к 2025 г. объем данных, генерируемых, потребляемых, копируемых и хранимых, превысит 180 зеттабайт. В то же время в 2020 г. этот объем составил 64,2 зеттабайта. С 2012 по 2020 гг. процент полезной информации для анализа вырос с 22 до 37 %. Эти

данные включают информацию из различных областей: социальные сети, развлечения, мониторинг и многое другое [1–2].

Для решения проблем, связанных с увеличивающимся объемом информации, возникает необходимость в разработке платформы, способной автоматизировать процессы сбора, фильтрации и анализа данных. Она должна обеспечивать сбор новостей из различных источников, фильтрацию по тематике, написание аннотаций по каждой собранной статье, иметь возможность, основываясь на заголовке и контексте статьи, соотносить ее с одной из категорий сферы информационных технологий, определять теги, а также в случае, если язык оригинальной статьи отличается от русского языка, переводить заголовок и контекст.

МАТЕРИАЛЫ И МЕТОДЫ

Решение о выборе конкретной архитектуры для проектируемой платформы, принимаемое перед началом ее разработки, может зависеть от некоторых факторов.

В первую очередь платформа должна отвечать требованиям высокой скорости обработки большого количества источников. Это особенно важно, так как в современных условиях объемы данных растут экспоненциально [1, 2]. Не менее важным условием, которым система должна отвечать, является масштабируемость – архитектура должна позволять системе легко масштабироваться вместе с увеличением числа пользователей и количества источников новостных статей. Без надлежащей адаптивности система может быстро стать устаревшей. Надежность также играет важную роль – архитектура должна обеспечивать стабильную работу платформы в сценарии высоких нагрузок, при необходимости распределяя задачи между другими узлами, элементами системы.

Одним из основных требований для разрабатываемой платформы является возможность ее легкого развертывания и обеспечения наилучшей совместимости на различных операционных системах. В данном контексте оптимальным решением является применение технологии контейнеризации с исполь-

зованием Docker и инструмента управления контейнерами Docker Compose.

Docker – это средство, позволяющее паковать всевозможные приложения прямо вместе со всеми включенными в них зависимостями в стандартизированные контейнеры. Эти самые контейнеры легко и просто могут быть перемещены между разными средами разработки и продакшна, обеспечивая консистентность и надежность выполнения приложений. Docker гарантирует изоляцию отличных процессов и ресурсов, что делает его мощным инструментом для создания, тестирования, а также развертывания всевозможных приложений в разнообразных сценариях [3].

Среди наиболее распространенных типов архитектуры программного обеспечения (ПО) выделяются следующие: монолитная, микросервисная, многоуровневая, сервис-ориентированная и клиент-серверная.

Монолитная архитектура предполагает организацию платформы в виде единого модуля, состоящего из тесно связанных между собой структурных компонентов. Этот метод является традиционным и обычно не требует значительных затрат для внедрения. Все программные компоненты монолитной системы взаимозависимы из-за использования встроенных механизмов обмена данными внутри системы. Модификация монолитной архитектуры возможна лишь частично и занимает много времени, поскольку даже небольшие изменения затрагивают большие области базы кода [4].

Идея микросервисной архитектуры основана на сервис-ориентированном подходе и предполагает создание независимых сервисов для обеспечения функциональности платформы. При таком подходе приложение разрабатывается как набор небольших сервисов; каждая служба является автономной и должна реализовывать одну бизнес-задачу в ограниченном контексте [5–8].

Многоуровневая архитектура, подобно клиент-серверной, делит платформу на уровни для управления данными и получения информации [5]. Архитектура данного типа может содержать два, три уровня или больше, в зависимости от потребностей проекта.

Сервис-ориентированная архитектура (SOA) – это метод разработки программного обеспечения, который использует программные компоненты, называемые сервисами, для создания бизнес-приложений. Каждый сервис предоставляет бизнес-возможности, сервисы также могут взаимодействовать друг с другом на разных платформах и языках. Разработчики применяют SOA для многократного использования сервисов в различных системах или объединения нескольких независимых сервисов для выполнения сложных задач [9].

Клиент-серверная архитектура ПО разделяет систему на два основных приложения, в которых клиент отправляет запросы на сервер, получая от него ответы [10]. Такой подход позволяет переложить задачу обработки данных на серверную часть.

При выборе архитектурного подхода для создания программного обеспечения важно учитывать особенности и ограничения каждого подхода. Рассмотрим, почему использование микросервисной архитектуры предпочтительно в проекте, где предусмотрено использование контейнеризации и Docker Compose, по сравнению с другими архитектурными стилями.

При рассмотрении клиент-серверной архитектуры возникают ограничения в масштабируемости и гибкости, однако, несмотря на это, она остается довольно хорошим выбором. При использовании клиент-серверного подхода проектирования ПО могут возникнуть проблемы с расширением серверной части при большом количестве запросов, что может привести к невозможности использования веб-платформы. Следует отметить, что некоторые из требований будет значительно сложнее реализовать при соблюдении данной архитектуры. При рассмотрении сценария одновременной обработки большого количества источников может возникнуть проблема с отображением аннотаций в реальном времени и в общей работе сервера.

Несмотря на свою простоту, монолитная архитектура имеет свои недостатки, которые могут затруднить ее применение в платформе.

В таких системах все части приложения тесно взаимосвязаны, что усложняет их поддержку и масштабирование. Существует вероятность, что для внесения изменений придется значительно модифицировать участки кода программы. Один из весомых недостатков использования монолитной архитектуры заключается в ограниченной возможности гибкого управления каждого из компонентов; сбой в одном из компонентов повлияет на всю систему.

Многослойная структура платформы позволяет разделить ее на различные уровни для управления информацией и обработки данных, что способствует повышению гибкости и некоторой масштабируемости приложения. Однако такая структура может усложнить внесение изменений из-за своей монолитной природы. В рамках такого подхода к разработке программного обеспечения данные должны проходить через каждый слой даже в случаях, когда это не требуется.

Сервис-ориентированная архитектура (SOA) позволяет эффективно использовать сервисы и масштабировать их, однако для этого часто требуются сложные методы взаимодействия, такие как применение ESB (Enterprise Service Bus) – «сервисной шины предприятия» (связующего ПО, обеспечивающего централизованный и унифицированный событийно-ориентированный обмен сообщениями между различными информационными системами).

При выборе такого архитектурного подхода обычно придерживаются более унифицированного технологического стека для всех сервисов. Несмотря на то что такая архитектура ПО предусматривает разделение системы на отдельные сервисы, эти сервисы часто тесно связаны друг с другом.

Исследовав все вышеперечисленные подходы к созданию программного обеспечения, можно сделать вывод, что микросервисная архитектура является наиболее предпочтительной. Несмотря на некоторые недостатки (более длительное время разработки по сравнению с другими архитектурными решениями, необходимость проведения

комплексного тестирования для отладки всех компонентов), преимуществ у этого типа архитектуры больше. Микросервисная архитектура обеспечивает гибкость за счет использования небольших автономных сервисов, которые легко модифицировать или масштабировать при помощи контейнеризации. Развитие функциональности платформы через подход разделения на микросервисы позволяет создавать, внедрять и масштабировать каждый модуль индивидуально. Этот метод является оптимальным для проектов, где требуется быстрое развитие и способность оперативно реагировать на изменения рынка.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Веб-платформа для обработки новостной информации объединяет технологии анализа текста с использованием больших языковых моделей (large language models, LLM) и информационных технологий с целью автоматизации сбора, анализа и представления актуальных новостей. Ниже представлена концептуальная модель, демонстрирующая

ключевые компоненты и взаимосвязи между ними.

На диаграмме (рис. 1), представлена модель платформы. Модель включает несколько ключевых компонентов и модулей, которые взаимодействуют между собой для обеспечения работы системы.

На диаграмме представлены следующие ключевые компоненты:

1. Парсер:

Атрибуты: Ресурс

Методы: Запуск(), Собрать статьи(), Следующая страница(), Отправить на сервер()

Функции: Парсер отвечает за извлечение информации из источников. Он собирает ссылки на новостные статьи, переходя по страницам информационного ресурса, после чего отправляет их сервису. Сервис создает очередь ссылок на статьи для дальнейшего получения их контекста и отправки на сервер базы данных.

2. Источник:

Атрибуты: Название, Ссылка, Язык

Функции: Источник предоставляет информацию, из которой парсер будет извлекать статьи.

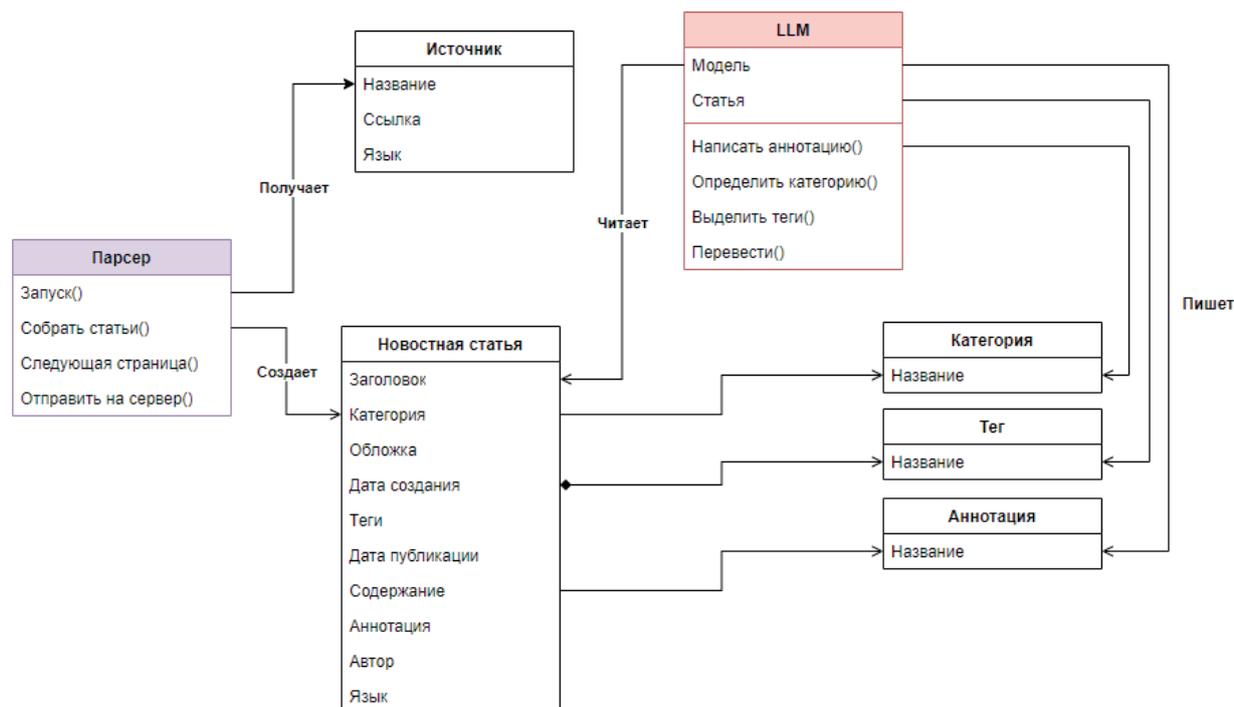


Рис. 1. Концептуальная диаграмма веб-платформы
Примечание: составлено авторами.

3. LLM:

Атрибуты: Модель, Статья

Методы: Написать аннотацию(), Определить категорию(), Выделить теги(), Перевести()

Функции: OpenChat обрабатывает статьи собранные парсером, анализирует их наполнение, пишет аннотации, определяет категории, присваивает теги и переводит.

4. Новостная статья:

Атрибуты: Заголовок, Категория, Обложка, Дата создания, Теги, Дата публикации, Содержание, Аннотация, Автор, Язык.

Функции: Этот объект содержит информацию о новостной статье, включая заголовок, дату создания, ссылку на источник, обложку, содержание, категорию, теги, аннотацию и язык.

5. Категория:

Атрибуты: Название

Функции: Включает в себя категорию статьи.

6. Тег:

Атрибуты: Название

Функции: Содержит теги, связанные с содержанием статьи.

7. Аннотация:

Атрибуты: Контекст

Функции: Хранит краткое описание содержания статьи.

Для более глубокого понимания архитектуры платформы и ее элементов, представим диаграмму, на которой отображены ключевые компоненты системы и их связи.

На рис. 2 представлены основные компоненты системы для анализа и классификации новостных статей, запущенные в контейнерах Docker. Каждый контейнер специализируется на выполнении своей конкретной задачи. Для некоторых групп контейнеров были настроены сетевые параметры для полной изоляции одних групп от других. Например, контейнер с работающим сервером базы данных (БД) находится в одной сети только с контейнером «golang-server», так что другие контейнеры, работающие с данными из БД, должны обращаться к этому контейнеру через соответствующие запросы к API.

Веб-интерфейс позволяет пользователям взаимодействовать с системой через браузер и HTTP-запросы. Клиентские запросы направляются на фронтенд, расположенный

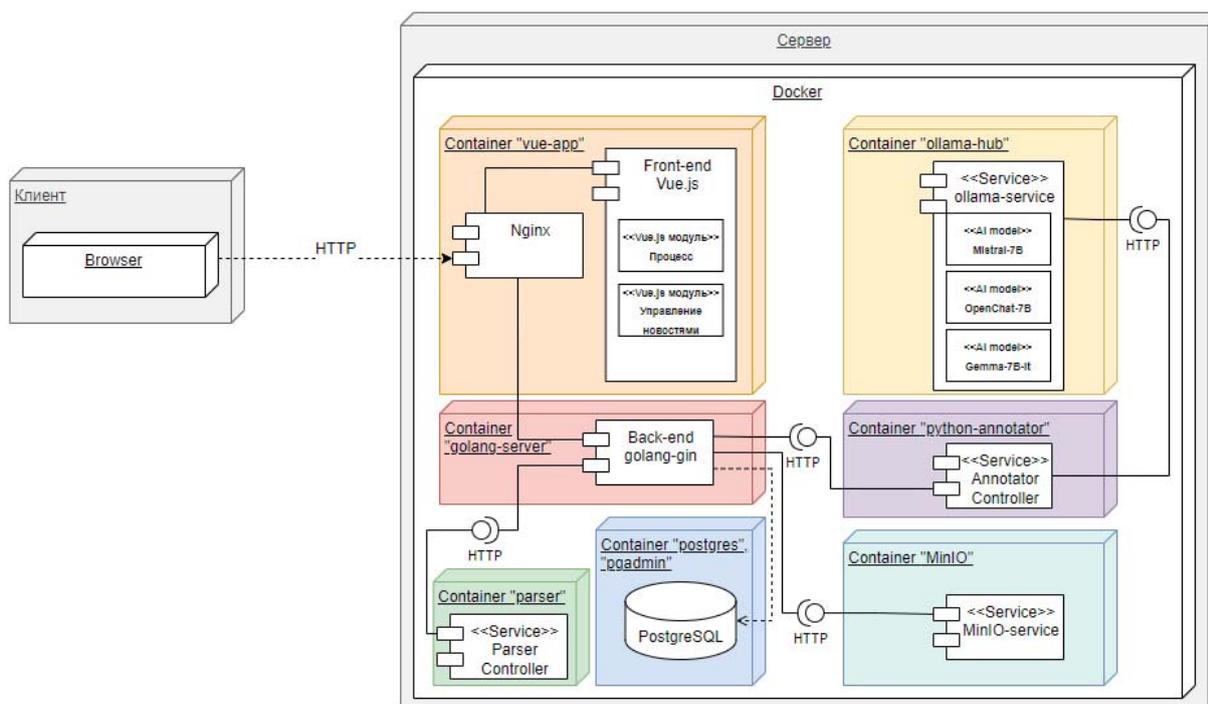


Рис. 2. Диаграмма компонентов
Примечание: составлено авторами.

в контейнере «vue-app». Nginx обрабатывает HTTP-запросы, а приложение на Vue.js предоставляет возможность просмотра новостных статей пользователю.

Используя контейнеризацию, можно упаковать каждый сервис со всеми его зависимостями в отдельное изолированное окружение для обеспечения единообразия среды как на этапе разработки, так и при развертывании и поддержке платформы. Кроме того, это позволяет избежать конфликтов между зависимостями и различными версиями программного обеспечения.

Рассмотрим классификацию новостных статей с использованием больших языковых моделей. Статья может быть отнесена к одной из следующих категорий: криптовалюта, конфиденциальность, безопасность и технологии. Эти категории хранятся в базе данных, в рис. 3. Значение «theme_id» используется как внешний ключ в соответствующем поле таблицы «articles».

	theme_id [PK] bigint	theme_name character varying (255)
1	1	technology
2	2	crypto
3	3	privacy
4	4	security

Рис. 3. Таблица «themes»

Примечание: составлено авторами.

После того как парсер соберет данные с веб-страницы, он проверяет все поля на заполненность, в случае успешной проверки, передаст их сервису. Сервис отправляет информацию на сервер для добавления в базу данных (табл. 1). После ответа от сервера, ссылка на статью удаляется из очереди сервиса.

Сервис аннотации запрашивает необработанные статьи с сервера каждые 4 часа. Сервер, взаимодействующий с базой данных, выполняет запрос (рис. 4) для поиска статей без аннотаций на определенных языках. Полученная информация о статьях (их исходном и отсутствующем языке аннотации) сортируется по дате публикации; затем сервер возвращает результат (рис. 5).

Сервис аннотаций отправляет запрос ollama-hub, фреймворк предназначен для запуска и управления большими языковыми моделями (LLM) на локальных вычислительных ресурсах, обеспечивающий возможность загрузки и развертывания выбранной LLM, а также доступ к ней через API, используя имя и параметры модели (рис. 6), запрос с заголовком статьи, содержанием и тематиками, к одной из которых можно отнести наполнение статьи.

После получения ответа требуется валидация результата нейронной сети (рис. 7). Для этого необходимо проверить наличие каждой тематики в ответе. В случае если совпадение было найдено, устанавливается тематика статьи, в противном случае, если совпадения не были найдены, то есть в ответе большой

Таблица 1

Некоторые данные из собранных новостных статей

Заголовок	Автор	Тематика	Язык	Дата публикации	Дата добавления
Spotify is once again hiking prices in the US	Paulius Grinkevičius	null	English	2024-06-03 13:35:00	2024-06-08 18:55:52.001527
Microsoft Edge introduces AI feature to improve...	Konstancija Gasaitytė	null	English	2024-05-22 11:54:00	2024-06-08 18:58:30.764772
Intel unveils its Lunar Lake chip for AI PCs	Paulius Grinkevičius	null	English	2024-06-04 08:56:00	2024-06-08 18:55:43.879135
Muslim Tinder exposes secrets, risks...	Vilius Petkauskas	null	English	2024-05-31 06:32:00	2024-06-08 18:56:36.649774
OneCoin fraudster gets ten years for laundering \$400M	Damien Black	null	English	2024-01-26 09:40:00	2024-06-08 13:00:09.248526

Примечание: составлено авторами.

```
query := `
SELECT
  a.id AS article_id,
  ln.language_code AS native_language,
  l.language_code,
  l.language_name
FROM
  articles a
CROSS JOIN
  languages l
LEFT JOIN
  languages ln ON a.language_id = ln.language_id
LEFT JOIN
  annotations an ON a.id = an.article_id AND l.language_id = an.language_id
WHERE
  an.article_id IS NULL
ORDER BY a.post_date DESC;`
```

Рис. 4. Запрос нахождение статей без аннотации
Примечание: составлено авторами.

```
JSON
1  {
2  "data": [
3  {
4    "article_id": "a484f6d1-110e-4569-997e-4e33ef3e93e8",
5    "native_language": "en-US",
6    "language_code": "en-US",
7    "language_name": "English"
8  },
9  {
10   "article_id": "a484f6d1-110e-4569-997e-4e33ef3e93e8",
11   "native_language": "en-US",
12   "language_code": "ru-RU",
13   "language_name": "Russian"
14 },
15 {
16   "article_id": "a484f6d1-110e-4569-997e-4e33ef3e93e8",
17   "native_language": "en-US",
18   "language_code": "fr-FR",
19   "language_name": "French"
20 },
21 {
22   "article_id": "a484f6d1-110e-4569-997e-4e33ef3e93e8",
23   "native_language": "en-US",
24   "language_code": "de-DE",
25   "language_name": "German"
26 },
27 {
28   "article_id": "b3b7c39b-90e5-4642-8c31-ec7d15d561bc",
29   "native_language": "en-US",
30   "language_code": "en-US",
31   "language_name": "English"
32 },
33 {
34   "article_id": "b476e0db-1136-41c5-972f-f42c5b9cd3b2",
35   "native_language": "en-US",
36   "language_code": "fr-FR",
37   "language_name": "French"
38 },
39 ]
40 }
```

Рис. 5. Ответ сервера, список необработанных статей
Примечание: составлено авторами.

```
def __init__(self):
    super().__init__()
    self.model_name = "openchat"
    self.data = {
        "stream": False, # Placeholder for stream option
        "options": {
            "temperature": 0.8, # Placeholder for temperature option
            "top_p": 0.9
        }
    }
```

Рис. 6. Название модели и параметры большой языковой модели
Примечание: составлено авторами.

```
words = ["technology", "crypto", "privacy", "security"]
prompt = """
Article:
...

Title: %s
Content: %s
...

Classify the article to one of the topics:
%s

Use template to answer:
Answer template: "Topic: topic 1"
"""
prompt = prompt % (article.title, article.body, ", ".join(words))
```

Рис. 7. Шаблон запроса для большой языковой модели
Примечание: составлено авторами.

языковой модели нет ни одной из указанной тематик, устанавливается значение «None». В дальнейшем перед отправкой на сервер для добавления в базу данных это позволит определить, корректно ли была обработана новостная статья, избегая некорректных записей.

Определение темы новостной статьи – один из обязательных этапов ее обработки (рис. 8). После загрузки статьи в систему необходимо выполнить ряд ключевых процедур, таких как аннотирование, перевод (в случае если язык статьи отличается от целевого языка, получаемого из ответа), категоризация и присвоение тегов (рис. 5). Эти действия позволяют группировать статьи по ряду параметров и повышают доступность для пользователей (табл. 2).

```
answer = self.generate(prompt=prompt, stream=stream, options=options)
answer = answer.message["content"].lower()

# Check if any of the words are present in the string
for word in words:
    if word in answer:
        article.theme_name = word
        break
else:
    article.theme_name = None
```

Рис. 8. Вызов метода генерации и валидация ответа нейронной сети

Примечание: составлено авторами.

ЗАКЛЮЧЕНИЕ

Применение микросервисной архитектуры кажется достаточно обоснованным и перспективным решением, поскольку это позволяет эффективно управлять различными компонентами системы, обеспечивая их изолированность и гибкость. Каждый микросервис может быть создан, внедрен и масштабирован индивидуально, что упрощает процесс разработки, обновления и поддержки всей платформы. На основе анализа популярных подходов к построению ПО, можно сделать вывод о целесообразности использования микросервисной архитектуры для разрабатываемой платформы.

Дальнейшее развитие веб-платформы может включать оптимизацию производительности и масштабируемости системы, параллельное выполнение процессов парсинга и аннотирования данных, а также использование кэширования для ускорения доступа к часто запрашиваемой информации. Также имеется возможность расширения числа источников новостных статей за счет добавления дополнительных информационных ресурсов, специализирующихся на области информационных технологий.

Таблица 2

Некоторые данные из аннотированных новостных статей

Заголовок	Автор	Тематика	Язык	Дата публикации	Дата добавления
Spotify is once again hiking prices in the US	Paulius Grinkevičius	technology	English	2024-06-03 13:35:00	2024-06-08 18:55:52.001527
Microsoft Edge introduces AI feature to improve...	Konstancija Gasaitytė	security	English	2024-05-22 11:54:00	2024-06-08 18:58:30.764772
Intel unveils its Lunar Lake chip for AI PCs	Paulius Grinkevičius	technology	English	2024-06-04 08:56:00	2024-06-08 18:55:43.879135
Muslim Tinder exposes secrets, risks...	Vilius Petkauskas	privacy	English	2024-05-31 06:32:00	2024-06-08 18:56:36.649774
OneCoin fraudster gets ten years for laundering \$400M	Damien Black	crypto	English	2024-01-26 09:40:00	2024-04-29 13:00:09.248526

Примечание: составлено авторами.

Список источников

1. Big Data Statistics 2023: How Much Data is in The World? URL: <https://firstsiteguide.com/big-data-stats> (дата обращения: 25.06.2024).
2. Big data – statistics & facts. URL: <https://www.statista.com/topics/1464/big-data> (дата обращения: 25.06.2024).
3. Docker: как создавать образы контейнеров и развертывать приложения. URL: <https://habr.com/ru/articles/776188> (дата обращения: 25.06.2024).
4. В чем разница между монолитной архитектурой и архитектурой микросервисов? URL: <https://web.archive.org/web/20240622123247/https://aws.amazon.com/ru/compare/the-difference-between-monolithic-and-microservices-architecture> (дата обращения: 25.06.2024).
5. 4 типа архитектуры программного обеспечения. URL: <https://nuancesprog.ru/p/12019> (дата обращения: 26.06.2024).
6. Microservices architecture design. URL: <https://learn.microsoft.com/en-us/azure/architecture/microservices> (дата обращения: 25.06.2024).
7. Артамонов Ю. С., Востокин С. В. Разработка распределенных приложений сбора и анализа данных на базе микросервисной архитектуры // Известия Самарского научного центра Российской академии наук. 2016. Т. 18, № 4–4. С. 688–693.
8. Вальдивия Х. А., Лора-Гонсалес А., Лимон К. и др. Паттерны микросервисной архитектуры: многопрофильный обзор литературы // Труды Института системного программирования РАН. 2021. Т. 33, № 1. С. 81–96.
9. Что такое сервис-ориентированная архитектура (SOA)? URL: <https://web.archive.org/web/20240622115612/https://aws.amazon.com/ru/what-is/service-oriented-architecture> (дата обращения: 25.06.2024).
10. Щекочихин О. В., Черкасова Н. В. Анализ шаблонов проектирования информационных систем клиент-серверной архитектуры // Информационно-экономические аспекты стандартизации и технического регулирования. 2019. № 5. С. 26–29.

Информация об авторах

К. В. Кузнецов – бакалавр.

С. А. Лысенкова – кандидат физико-математических наук, доцент.

References

1. Big Data Statistics 2023: How Much Data is in The World? URL: <https://firstsiteguide.com/big-data-stats> (accessed: 25.06.2024).
2. Big data – statistics & facts. URL: <https://www.statista.com/topics/1464/big-data> (accessed: 25.06.2024).
3. Docker: kak sozdavat obrazy konteynerov i razvertyvat prilozheniya. URL: <https://habr.com/ru/articles/776188> (accessed: 25.06.2024). (In Russ.).
4. V chem raznitsa mezhdru monolitnoy arkhitekturoy i arkhitekturoy mikroservisov? URL: <https://web.archive.org/web/20240622123247/https://aws.amazon.com/ru/compare/the-difference-between-monolithic-and-microservices-architecture> (accessed: 25.06.2024). (In Russ.).
5. 4 tipa arkhitektury programmno obespечeniya. URL: <https://nuancesprog.ru/p/12019> (accessed: 26.06.2024). (In Russ.).
6. Microservices architecture design. URL: <https://learn.microsoft.com/en-us/azure/architecture/microservices> (accessed: 25.06.2024).
7. Artamonov Y. S., Vostokin S. V. Development of distributed applications for data collection and analysis on the basis of a microservice architecture. *Izvestia of Samara Scientific Center of the Russian Academy of Sciences*. 2016;18(4–4):688–693. (In Russ.).
8. Valdivia J. A., Lora-González A., Limón X. et al. Patterns Related to Microservice Architecture: A Multivocal Literature Review. *Trudy ISP RAN/Proc. ISP RAS*. 2021;33(1):81–96. (In Russ.).
9. Chto takoe servis-orientirovannaya arkhitektura (SOA)? URL: <https://web.archive.org/web/20240622115612/https://aws.amazon.com/ru/what-is/service-oriented-architecture> (accessed: 25.06.2024). (In Russ.).
10. Shchekochikhin O. V., Cherkasova N. V. Analysis of templates of designing information systems of client-server architecture. *Informatsionno-ekonomicheskiye aspekty standartizatsii i tekhnicheskogo regulirovaniya*. 2019;(5):26–29. (In Russ.).

About the authors

K. V. Kuznetsov – Bachelor's Degree Student.

S. A. Lysenkova – Candidate of Sciences (Physics and Mathematics), Associate Professor.

Научная статья
УДК 519.866:369



<https://doi.org/10.35266/1999-7604-2024-3-5>

Анализ срабатываний предикторов в функции Леонтьева для объема финансирования отделения социального фонда

Сергей Иванович Носков¹✉, Александр Петрович Медведев²,
Иван Владимирович Овсянников³

^{1, 2, 3}Иркутский государственный университет путей сообщения, Иркутск, Россия

¹sergey.noskov.57@mail.ru✉

²medvedeff.a.p@yandex.ru

³ivani29030609@gmail.com

Аннотация. В работе дан краткий обзор публикаций по применению методов математического моделирования в области пенсионного обеспечения. В частности, рассмотрены: применение модели Хестона – Халла – Уайта в управлении пенсионным фондом; описание доходов пенсионных фондов в Испании и Великобритании с помощью авторегрессионной модели и нейронной сети; модель участников пенсионного плана в Словацкой Республике; проблема оптимального управления пенсионными фондами с установленными взносами; задача оптимального инвестирования участника пенсионной схемы, столкнувшегося со стохастической инфляцией; модель нелинейной оптимизации для итальянской пенсионной системы с фиксированными взносами. В статье разработаны две версии кусочно-линейной функции Леонтьева для суммарного годового объема финансирования отделения социального фонда России по Иркутской области – обычная и с оптимальными преобразованиями независимых переменных. Обе эти версии характеризуются допустимой точностью. Их векторы срабатываний указывают на то, что фактором, в основном определяющим динамику выходной переменной, является количество самозанятых (плательщиков взносов за себя).

Ключевые слова: кусочно-линейная функция Леонтьева, пенсионный фонд, задача линейно-булева программирования, страховые пенсии, самозанятые, получатели пособий, застрахованные лица, количество пенсионеров

Для цитирования: Носков С. И., Медведев А. П., Овсянников И. В. Анализ срабатываний предикторов в функции Леонтьева для объема финансирования отделения социального фонда // Вестник кибернетики. 2024. Т. 23, № 3. С. 40–45. <https://doi.org/10.35266/1999-7604-2024-3-5>.

Original article

Analysis of predictor responses in the Leontief function for financing amount of the social fund department

Sergey I. Noskov¹✉, Aleksandr P. Medvedev², Ivan V. Ovsyannikov³

^{1, 2, 3}Irkutsk State Transport University, Irkutsk, Russia

¹sergey.noskov.57@mail.ru✉

²medvedeff.a.p@yandex.ru

³ivani29030609@gmail.com

Abstract. The paper gives a brief review of publications on the application of mathematical modeling methods in pension provision. Regarding the application of the Heston-Hull-White model in pension fund management, specifically, we examine: the description of pension fund returns in Spain and Great Britain using an autoregressive model and neural network; the model of pension plan participants in the Slovak Re-

public; the problem of optimal management of defined contribution pension funds; the problem of optimal investment of a pension scheme participant facing stochastic inflation; the nonlinear optimization model for the Italian defined contribution pension system. The paper develops two versions of the Leontief piecewise linear function for the total annual funding of the Russian Social Fund's Irkutsk Oblast branch: the ordinary version and the version with optimal transformations of the independent variables. Both versions are characterized by acceptable accuracy. Their triggering vectors indicate that the factor determining the dynamics of the output variable is the number of self-employed (payers of contributions for themselves).

Keywords: Leontief piecewise linear function, pension fund, linear Boolean programming problem, insurance pensions, self-employed, benefit recipients, insured persons, number of pensioners

For citation: Noskov S. I., Medvedev A. P., Ovsyannikov I. V. Analysis of predictor responses in the Leontief function for financing amount of the social fund department. *Proceedings in Cybernetics*. 2024;23(3):40–45. <https://doi.org/10.35266/1999-7604-2024-3-5>.

ВВЕДЕНИЕ

Пенсионное обеспечение населения является важнейшей социальной функцией любого государства. Вопросам анализа устойчивых закономерностей этого процесса и поиска путей повышения его эффективности в мире уделяется большое внимание. Традиционным средством проведения исследований в этой области является математическое моделирование. При этом используются как традиционные линейные, так и существенно нелинейные модели.

Так, в работе [1] представлена модель Хестона – Халла – Уайта в управлении пенсионным фондом. В явном виде представлены оптимальная политика инвестирования и выплаты пособий с возможностью прощадки доходов. Управляющему пенсионным фондом разрешено инвестировать средства фонда в безрисковые и рискованные активы. Динамика цен рискованных активов развивается в соответствии с моделью гибридной стохастической волатильности. В источнике [2] показано, как доходы пенсионных фондов в Испании и Великобритании адекватно описываются соответственно с помощью авторегрессионной модели и нейронной сети. Статья [3] посвящена описанию многопериодной стохастической оптимизационной модели для решения задачи оптимального выбора пенсионного фонда участником пенсионного плана. В модели участникам пенсионного плана предоставляется возможность периодически переключаться между J-типами фондов с разными профилями риска и таким образом ак-

тивно управлять своей подверженностью риску и ожидаемой доходностью. Минимизация показателя отклонения среднего значения риска за несколько периодов при ограничении ожидаемой доходности приводит к крупномасштабной линейной программе для пенсионной системы Словацкой Республики. В исследовании [4] изучается проблема оптимального управления пенсионными фондами с установленными взносами на этапе распределения в условиях инфляции, смертности и неопределенности модели. При этом применяются методы робастного управления и динамического программирования. В источнике [5] рассматривается задача оптимального инвестирования участника пенсионной схемы, столкнувшегося со стохастической инфляцией, с использованием критерия средней дисперсии Марковица, методов Лагранжа и стохастического динамического программирования.

В работе [6] исследуется эффективность испанских пенсионных фондов с использованием ряда линейных и нелинейных моделей эффективности. Поскольку в выборке представлены характеристики моментов высокого порядка, традиционные показатели эффективности искажаются. Поэтому формируются альтернативные модели производительности, которые включают факторы риска, моделирующие асимметрию и эксцесс. При этом учитываются факторы, которые отражают нелинейность, присущую некоторым базовым активам, используемым в пенсионных фондах. В исследовании [7]

показано, как критерий средней дисперсии может быть применен к многопериодной ситуации, чтобы получить эффективные портфели в контексте активов и пассивов. Модель оптимизации позволяет оценить транзакционные издержки и стохастические волатильности как активов, так и обязательств. Кроме того, дается общая основа для прогнозирования обязательств пенсионного фонда, а также для формирования доходности активов. В источнике [8] представлена новая нелинейная стохастическая модель процентных ставок, которая, как особый случай, содержит стандартную линейную модель Паркера. Получено множество практических применений для движения финансов и пенсий. В статье [9] предлагается надежная основа для управления пенсионным фондом с установленными выплатами, основанная на оптимизации. Показывается, что этот подход позволяет гибко моделировать многие особенности проблемы управления пенсионным фондом. Он является вычислительно приемлемой альтернативой подходам, основанным на стохастическом программировании. Наконец, в источнике [10] представлена модель нелинейной оптимизации для итальянской пенсионной системы с фиксированными взносами, основанной на трех контрольных переменных: индексации пенсий, условной ставке и ставке взносов. Целевая функция учитывает как социальную адекватность, так и устойчивость ставок взносов в условиях ограничений на ликвидность и устойчивость.

МАТЕРИАЛЫ И МЕТОДЫ

В публикации [11] представлена разработанная на основе статистической информации за 2012–2022 гг. линейная по параметрам, но нелинейная по независимым переменным регрессионная модель объема финансирования отделения пенсионного фонда в России:

$$y = -4927260 + 0,35293x_1 + 115801 \frac{x_2 + x_3}{x_1} + \frac{59811000}{\ln x_4}, \quad (1)$$

$R = 0,995371$; $F = 501,69$; $E = 0,85$; $T = (-31,11; 4,487; 3,475; 28,1)$.

Приведем весь полный набор зависимой и потенциальных независимых переменных, использованных при проведении конкурса альтернативных вариантов для модели (1):

y – суммарный годовой объем финансирования отделения социального фонда России по Иркутской области, тыс. руб.;

x_1 – количество страховых пенсий по старости;

x_2 – количество страхователей-работодателей, чел.;

x_3 – количество самозанятых (плательщиков взносов за себя), чел.;

x_4 – количество получателей пособий единовременной выплаты, чел.;

x_5 – количество застрахованных лиц, чел.;

x_6 – количество пенсионеров, чел.;

x_7 – количество работающих пенсионеров, чел.

Для оценки адекватности модели использованы критерии адекватности:

R – критерий множественной детерминации;

F – критерий Фишера;

E – средняя относительная ошибка аппроксимации;

T – вектор значений критерия Стьюдента для параметров модели.

Таким образом, в модель (1) оказались включены лишь четыре независимых переменных из семи возможных. Она является весьма адекватной исследуемому объекту, на что указывают значения перечисленных выше критериев.

Поставим задачу построения на основе той же информации кусочно-линейной функции Леонтьева:

$$y = \min(a_1x_1, a_2x_2, \dots, a_7x_7). \quad (2)$$

Ее отличительной особенностью является то, что для любого наблюдения выборки минимум в модели (2) реализуется на одной независимой переменной, значение которой лимитирует значение y . При этом любое значение других переменных не изменит выходную переменную. Это свойство модели (2) позволяет сформировать вектор срабатываний $\lambda = (\lambda_2, \dots, \lambda_{11})$, k -я компонента которого

λ_k равна номеру независимой переменной, на которой реализовался минимум в (2).

Оценивание неизвестных параметров кусочно-линейной модели (2) сводится к решению задачи линейно-булева программирования (ЛБП) (см., например, [12]).

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

При разработке кусочно-линейной зависимости (2) будем иметь в виду, что так же, как и при построении регрессии (1), финальный вариант модели может не содержать все переменные x_i , $i = \overline{1,7}$. В результате проведения соответствующего конкурса альтернатив он принял вид:

$$y = \min(3,5784x_2, 2,0353x_3, 0,0579x_4), \quad (3)$$

$$E = 8,1211; \lambda = (3,3,3,3,3,3,3,4,2).$$

По точности модель (3), если судить по критерию E , существенно уступает модели (1), что совершенно естественно ввиду кусочного характера первой. Тем не менее значение средней относительной ошибки аппроксимации (8,1211 %) для нее вполне приемлемо. На рисунке приведены значения фак-

тических, заданных в выборке, и расчетных, вычисленных по модели, значений выходной переменной y .

В модель (3) удалось включить лишь три независимые переменные из семи возможных вследствие статистической незначимости остальных. При этом в девяти годах из одиннадцати (в 2012–2020 гг.) лимитирующим фактором является количество самозанятых (плательщиков взносов за себя). Таким образом, именно этот показатель главным образом и определяет характер динамики суммарного годового объема финансирования отделения Социального фонда России по Иркутской области. Сделаем некоторые пояснения к этому результату.

К самозанятым гражданам относятся их категории, уплачивающие налог на профессиональный доход 4 или 6 %. Налоговые ставки зависят от источника дохода налогоплательщика. Ставка 4 % используется, если доход за товар, работы или услуги поступил от физического лица. Ставка 6 % применяется, если это поступление от юридического лица или индивидуального предпринимателя. Ограничение накладывает и лимит на доходы самоза-

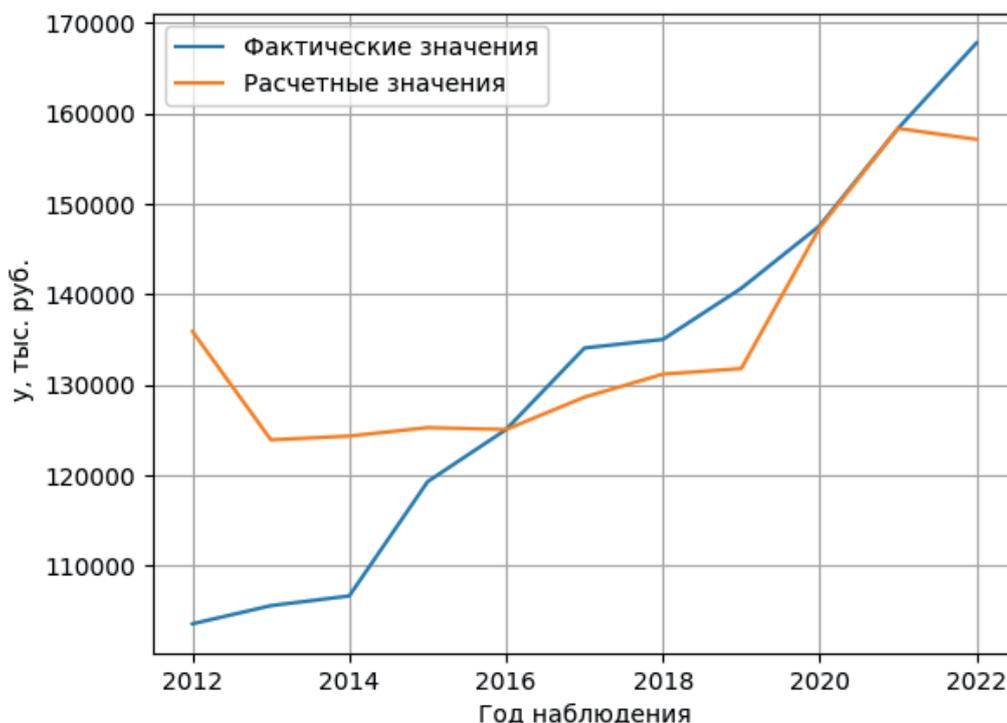


Рисунок. Расчетные и фактические значения зависимой переменной

Примечание: составлено авторами.

нятого – суммарный годовой доход не должен превышать 2,4 млн руб. В случае превышения лимита гражданин теряет право применять этот режим налогообложения.

В общем случае самозанятые категории граждан не платят обязательные страховые взносы, однако с целью формирования пенсии такие категории могут добровольно вступать в правоотношения по обязательному пенсионному страхованию, направив заявление в территориальный орган Социального фонда России и уплачивать страховые взносы самостоятельно. Вместе с этим утратится их право на индексацию пенсии; страховую пенсию по случаю потери кормильца в связи с уходом за детьми до 14 лет; на пенсию, назначенную по предложению службы занятости; на компенсационные выплаты по уходу за нетрудоспособными лицами и часть других выплат. Таким образом, категория самозанятых граждан (в случае добровольной оплаты взносов на обязательное пенсионное страхование) должна рассматриваться как с точки зрения количественной характеристики, непосредственно определяющей размеры страховых взносов, так и с точки зрения качественной величины, одновременно накладывающей ограничения на индексацию пенсий и часть социальных выплат.

Список источников

1. Mwanakatwe P., Song L., Hagenimana E. et al. Management strategies for a defined contribution pension fund under the hybrid stochastic volatility model // *Computational and Applied Mathematics*. 2019. Vol. 38, no. 2. <https://doi.org/10.1007/s40314-019-0790-z>.
2. Alda M., Agudo L. Linear and nonlinear financial time series: evidence in a sample of pension funds in Spain and the United Kingdom // *Applied Economics Letters*. 2012. Vol. 19, no. 18. P. 1933–1937. <https://doi.org/10.1080/13504851.2012.669454>.
3. Kilianová S., Pflug G. C. Optimal pension fund management under multi-period risk minimization // *Annals of Operations Research*. 2008. Vol. 166, no. 1. P. 261–270. <https://doi.org/10.1007/s10479-008-0405-3>.
4. Baltas I., Dopierala L., Kolodziejczyk K. et al. Optimal management of defined contribution pension funds under the effect of inflation, mortality and uncertainty // *European Journal of Operational Research*. 2022. Vol. 298, no. 3. P. 1162–1174. <https://doi.org/10.1016/j.ejor.2021.08.038>.

Построим теперь модифицированный аналог кусочно-линейной модели Леонтьева (3), выбрав в соответствии с алгоритмом из источника [12] оптимальные преобразования для каждой независимой переменной из некоего набора элементарных функций – логарифмических, степенных и показательных. В результате получим модель:

$$y = \min(0,0155x_2^{1,5}; 0,0078x_3^{1,5}; 332,6545\sqrt{x_4}), \quad (4)$$

$$E = 7,921; \lambda = (3,3,3,3,3,3,3,3,4,2).$$

Модель (4) характеризуется несколько большей точностью, чем (3), однако их векторы срабатываний совпадают.

ЗАКЛЮЧЕНИЕ

В работе построены два варианта кусочно-линейной функции Леонтьева для суммарного годового объема финансирования отделения Социального фонда России по Иркутской области – обычный и с оптимальными преобразованиями независимых переменных. Оба варианта обладают приемлемой точностью, а анализ их векторов срабатывания указывает на то, что показателем, главным образом и определяющим характер динамики выходной переменной, является количество самозанятых (плательщиков взносов за себя).

References

1. Mwanakatwe P., Song L., Hagenimana E. et al. Management strategies for a defined contribution pension fund under the hybrid stochastic volatility model. *Computational and Applied Mathematics*. 2019;38(2). <https://doi.org/10.1007/s40314-019-0790-z>.
2. Alda M., Agudo L. Linear and nonlinear financial time series: evidence in a sample of pension funds in Spain and the United Kingdom. *Applied Economics Letters*. 2012;19(18):1933–1937. <https://doi.org/10.1080/13504851.2012.669454>.
3. Kilianová S., Pflug G. C. Optimal pension fund management under multi-period risk minimization. *Annals of Operations Research*. 2008;166(1):261–270. <https://doi.org/10.1007/s10479-008-0405-3>.
4. Baltas I., Dopierala L., Kolodziejczyk K. et al. Optimal management of defined contribution pension funds under the effect of inflation, mortality and uncertainty. *European Journal of Operational Research*. 2022;298(3):1162–1174. <https://doi.org/10.1016/j.ejor.2021.08.038>.

5. Yao H., Yang Z., Chen P. Markowitz's mean-variance defined contribution pension fund management under inflation: a continuous-time model // *Insurance: Mathematics and Economics*. 2013. Vol. 53, no. 3. P. 851–863. <https://doi.org/10.1016/j.insmatheco.2013.10.002>.
6. Alda M., Agudo L., Gallagher L. Performance of Spanish pension funds: robust evidence from alternative models // *Applied Financial Economics*. 2013. Vol. 23, no. 4. P. 297–314. <https://doi.org/10.1080/09603107.2012.720011>.
7. Frauendorfer K., Jacoby U., Schwendener A. Regime switching based portfolio selection for pension funds // *Journal of Banking & Finance*. 2007. Vol. 31, no. 8. P. 2265–2280. <https://doi.org/10.1016/j.jbankfin.2007.02.003>.
8. Stehlik M., Kiseľák J., Potocký R., Jordanova P. Generalized interest rate dynamics and its impacts on finance and pensions // *Stochastic Analysis and Applications*. 2016. Vol. 35, no. 1. P. 178–190. <https://doi.org/10.1080/07362994.2016.1224975>.
9. Iyengar G., Ka A., Chun M. A robust optimization approach to pension fund management // *Asset and Liability Management Handbook*. 2011. P. 308–330. https://doi.org/10.1057/9780230307230_12.
10. Devolder P., Levantesi S., Menziatti M. Automatic balance mechanisms for notional defined contribution pension systems guaranteeing social adequacy and financial sustainability: an application to the Italian pension system // *Annals of Operations Research*. 2020. Vol. 299, no. 1–2. P. 765–795. <https://doi.org/10.1007/s10479-020-03819-x>.
11. Носков С. И. Медведев А. П. Реализация конкурса регрессионных моделей при оценке объема финансирования социального и пенсионного обеспечения // *Инженерный вестник Дона*. 2024. № 4. URL: <http://www.ivdon.ru/ru/magazine/archive/n4y2024/9155> (дата обращения: 17.05.2024).
12. Носков С. И., Попов Е. С. Подход к построению функции Леонтьева с нелинейными предикторами // *Известия Тульского государственного университета. Технические науки*. 2024. № 2. С. 105–108.
5. Yao H., Yang Z., Chen P. Markowitz's mean-variance defined contribution pension fund management under inflation: a continuous-time model. *Insurance: Mathematics and Economics*. 2013;53(3):851–863. <https://doi.org/10.1016/j.insmatheco.2013.10.002>.
6. Alda M., Agudo L., Gallagher L. Performance of Spanish pension funds: robust evidence from alternative models. *Applied Financial Economics*. 2013;23(4):297–314. <https://doi.org/10.1080/09603107.2012.720011>.
7. Frauendorfer K., Jacoby U., Schwendener A. Regime switching based portfolio selection for pension funds. *Journal of Banking & Finance*. 2007;31(8):2265–2280. <https://doi.org/10.1016/j.jbankfin.2007.02.003>.
8. Stehlik M., Kiseľák J., Potocký R., Jordanova P. Generalized interest rate dynamics and its impacts on finance and pensions. *Stochastic Analysis and Applications*. 2016;35(1):178–190. <https://doi.org/10.1080/07362994.2016.1224975>.
9. Iyengar G., Ka A., Chun M. A robust optimization approach to pension fund management. *Asset and Liability Management Handbook*. 2011:308–330. https://doi.org/10.1057/9780230307230_12.
10. Devolder P., Levantesi S., Menziatti M. Automatic balance mechanisms for notional defined contribution pension systems guaranteeing social adequacy and financial sustainability: an application to the Italian pension system. *Annals of Operations Research*. 2020;299(1–2):765–795. <https://doi.org/10.1007/s10479-020-03819-x>.
11. Noskov S. I., Medvedev A. P. Implementation of a competition for regression models in assessing the amount of social and pension funding. *Engineering Journal of Don*. 2024;(4). URL: <http://www.ivdon.ru/ru/magazine/archive/n4y2024/9155> (accessed: 17.05.2024). (In Russ.).
12. Noskov S. I., Popov E. S. Approach to constructing the Leontief function with nonlinear predictors. *News of the Tula state university. Technical sciences*. 2024;(2):105–108. (In Russ.).

Информация об авторах

С. И. Носков – доктор технических наук, профессор.
А. П. Медведев – ассистент.
И. В. Овсянников – студент.

About the authors

S. I. Noskov – Doctor of Sciences (Engineering), Professor.
A. P. Medvedev – Assistant Professor.
I. V. Ovsyannikov – Student.

Научная статья

УДК 004.032.26:004.312.26

<https://doi.org/10.35266/1999-7604-2024-3-6>



Разработка архитектуры нейросетевого декодирования, основанной на системах стробирования и распределения весовых коэффициентов

Александр Александрович Пирогов¹, Марина Владимировна Хорошайлова²,
Элеонора Викторовна Сёмка³✉

^{1, 2}Воронежский государственный технический университет, Воронеж, Россия

³Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина» Минобороны России, Воронеж, Россия

¹Pirogov.alx@gmail.com, <https://orcid.org/0000-0002-5105-0505>

²pmv2205@mail.ru, <https://orcid.org/0000-0001-9167-9538>

³semka_elya@mail.ru✉, <https://orcid.org/0000-0002-0194-6979>

Аннотация. Проведен анализ рабочей архитектуры для алгоритма нейросетевого декодирования, основанного на распространении доверия, в которой контролируется количество обучаемых параметров и вычислений в нейронной сети благодаря распределению весов и эффективному вычислительному графу. Операция распределения весовых коэффициентов включает в себя вычисление взвешенной суммы выходных сигналов нейронов уровня слоя, умноженной на соответствующие веса, и сложение смещений. Метод выделения участка данных предполагает применение нелинейной функции активации к выходным сигналам нейронов. После нескольких итераций локального декодирования значение потерь рассчитывается с использованием функции потерь среднеквадратичной ошибки.

Результаты моделирования показывали, что производительность улучшается по сравнению со стандартным декодером, построенного с использованием стандартного алгоритма распространения доверия (belief propagation, BP), благодаря применению подхода, подобного нейросетевому алгоритму BP.

Предложена надежная схема декодирования на основе нейронной сети, предназначенная для систем беспроводной связи. Данная архитектура рекуррентной нейронной сети, основанная на алгоритмах стробирования и распределения весовых коэффициентов, предназначена для выполнения декодирования распространения доверия без предварительного знания схемы кодирования.

Ключевые слова: нейросетевой декодер, рекуррентная нейронная сеть, блочные коды, алгоритм распространения доверия, функция весовых коэффициентов

Финансирование: работа выполнена при поддержке Минобрнауки России в рамках Федерального проекта «Подготовка кадров и научного фундамента для электронной промышленности» Государственной программы Российской Федерации «Научно-технологическое развитие Российской Федерации» при реализации Программы развития учебного дизайн-центра электроники «Силовая электроника» (соглашение № 075-02-2024-1517 от 7.03.2024 г. и № 075-02-2024-1520 от 7.03.2024 г.) в ФГБОУ ВО «Воронежский государственный технический университет».

Для цитирования: Пирогов А. А., Хорошайлова М. В., Сёмка Э. В. Разработка архитектуры нейросетевого декодирования, основанной на системах стробирования и распределения весовых коэффициентов // Вестник кибернетики. 2024. Т. 23, № 3. С. 46–55. <https://doi.org/10.35266/1999-7604-2024-3-6>.

Original article

Development of a neural network decoding architecture based on gating and weight distribution systems

Aleksandr A. Pirogov¹, Marina V. Khoroshailova², Eleonora V. Syomka³✉

^{1,2}Voronezh State Technical University, Voronezh, Russia

³Military Educational and Scientific Centre of the Air Force N. E. Zhukovsky and Y. A. Gagarin Air Force Academy the Ministry of Defence of the Russian Federation, Voronezh, Russia

¹pirogov.alx@gmail.com, <https://orcid.org/0000-0002-5105-0505>

²pmv2205@mail.ru, <https://orcid.org/0000-0001-9167-9538>

³semka_elya@mail.ru✉, <https://orcid.org/0000-0002-0194-6979>

Abstract. The research analyzes the working architecture for a neural network decoding algorithm based on belief propagation. It is found that the weight distribution and an efficient computational graph determine the number of trainable parameters and computations in the neural network. The weight distribution involves calculating the weighted sum of the output signals of the neurons of a layer, multiplied by the corresponding weights, and adding the biases. The data region extraction method involves applying a nonlinear activation function to the output signals of neurons. After several iterations of local decoding, the algorithm calculates the loss value using the mean square error loss function.

The simulation results indicated that using an approach similar to the neural network Belief Propagation (BP) algorithm improved the performance compared to the standard decoder built using the standard BP algorithm.

A robust neural network-based decoding scheme for wireless communication systems is proposed. This recurrent neural network architecture, based on gating and weight distribution algorithms, is designed to perform belief propagation decoding without prior knowledge of the coding scheme.

Keywords: neural network decoder, recurrent neural network, block codes, belief propagation, weight function

Funding: the work is supported by the Ministry of Education and Science of the Russian Federation within the Federal Project “Podgotovka kadrov i nauchnogo fundamenta dlya elektronnoy promyshlennosti” of the State Program of the Russian Federation “Nauchno-tekhnologicheskoe razvitie Rossiyskoy Federatsii” in the implementation of the Program for the Development of the Educational Design Center for Electronics “Silovaya elektronika” (agreement No. 075-02-2024-1517 dated March 7, 2024, and No. 075-02-2024-1520 dated March 7, 2024) at the Voronezh State Technical University.

For citation: Pirogov A. A., Khoroshailova M. V., Syomka E. V. Development of a neural network decoding architecture based on gating and weight distribution systems. *Proceedings of Cybernetics*. 2024;23(3):46–55. <https://doi.org/10.35266/1999-7604-2024-3-6>.

ВВЕДЕНИЕ

Современная система связи должна быть оперативна, достоверна, помехоустойчива, надежна и скрытна. Опираясь на данные требования к связи, можно утверждать, что помехоустойчивое кодирование является неотъемлемой частью цифровой обработки информационных данных. Рассмотрены архитектура каскадного кодера для программируемой логической интегральной схемы (ПЛИС) с использованием внутреннего низкоплотностного кода стандарта DVB-S2 и внешнего кода Рида – Соломона [1], архитектура ПЛИС для квазициклических кодов

четности с низкой плотностью (QC-LDPC), основанных на построении идентичной матрицы circulant-1 [2]. Алгоритм декодирования с использованием метода низкой плотности проверки на четность (LDPC) на основе ПЛИС для реализации полностью параллельных LDPC-декодеров предназначен для оптимизации использования логики ПЛИС и уменьшения времени задержки декодирования [3]. Каскадная сетевая структура для распознавания LDPC кодирования вслепую рассматривалась авторами в работе [4]. Декодирование LDPC-алгоритма побитовым адаптивным порогом инверсии бита (adaptive

threshold bit flipping, ATBF) исследовали авторы в работе [5].

Коды с исправлением ошибок, имеющие широкое применение при передаче данных, хранении данных и отказоустойчивых вычислениях, предназначены для защиты информации от случайных ошибок. Основным классом помехоустойчивых кодов является класс линейных блочных кодов (LBC). В LBC к строке информационных символов добавляются некоторые дополнительные символы, предназначенные для формирования кодового слова, таким образом, чтобы отделить каждый информационный вектор дальше друг от друга в пространстве кодовых слов. Следовательно, когда на кодовое слово влияет некоторый шум во время передачи или обработки, декодер может исправить ошибку, связав зашумленную строку с ближайшим допустимым кодовым словом.

Современные инструменты машинного обучения, универсальные и простые в использовании могут значительно улучшить производительность декодирования [6]. Нейронные декодеры распространения доверия были представлены как способ улучшения производительности декодирования итерационного алгоритма распространения доверия для линейных блочных кодов короткой и средней длины. Основная идея, которая лежит в основе этих декодеров, заключается в представлении распространения убеждений в виде нейронной сети, позволяющей адаптивно взвешивать процесс декодирования.

Цель: разработка упорядоченной архитектуры, обеспечивающей применение алгоритма нейросетевого распространения доверия для декодирования линейных блочных кодов с улучшенной гибкостью декодера благодаря его возможностям обучения и повышения производительности по сравнению со стандартным декодером BP.

МАТЕРИАЛЫ И МЕТОДЫ

Коды проверки четности низкой плотности (LDPC) были выбраны в стандарте 5G наряду с полярными кодами. Низкоплотностные коды представляют собой класс линейных блочных

кодов и являются одними из самых эффективных помехоустойчивых кодов, выполняющими итеративные алгоритмы декодирования. Тем не менее в некоторых случаях использования, таких как IoT (Internet of Things – интернет вещей) с низким энергопотреблением, нельзя применять большие коды из-за сложности их декодирования и их несоответствия относительно небольшим полезным нагрузкам, с которыми приходится сталкиваться. При меньшей длине кода матрица проверки четности (ПЧ) кодов LDPC не становится более разреженной; свойство «низкой плотности» теряется, и становится неизбежным наличие коротких циклов в ПЧ-матрице. Как следствие, производительность алгоритма декодирования, основанного на распространении доверия (BP), также известного как sum-product (SP), становится низкоэффективной. В такой ситуации предлагается представить алгоритм BP в виде нейронной сети (НС), что позволяет адаптивно взвешивать сообщения, которыми обмениваются в процессе декодирования.

Рассматривается линейный блочный код C длиной n и разрядом k . Связанная матрица генератора размером $k \times n$ и соответствующая матрица проверки четности размером $(n - k) \times n$ обозначаются как G и H соответственно. Матрица проверки четности может быть представлена в виде двудольной графовой модели, называемой графом Таннера, определенного типа факторного графа (ФГ). Исследование способа идентификации циклов в графах Таннера LDPC кодов на основе пересечений коротких замкнутых структур в протографах было рассмотрено в работе [7]. Такое представление обеспечивает эффективное итеративное декодирование на основе классических алгоритмов передачи сообщений, таких как BP. Этот алгоритм направлен на то, чтобы сходиться к передаваемому кодовому слову путем итеративного обмена «предположениями» между узлами графа о вероятных значениях полученных битов кодового слова. Обычно обмениваемые сообщения связаны с логарифмическими коэффициентами правдоподобия (LLR) полученных

битов, и правило обновления sum-product может быть применено к различным узлам, используя следующие уравнения:

– правило обновления SP, примененное к переменному узлу i для вычисления сообщения для контрольного узла j :

$$\mu_{v_i c_j} = \lambda_i + \sum_{k \neq j} \mu_{c_k v_i}, \quad (1)$$

где λ_i – предшествующее значение LLR, полученное переменным узлом i ,

$\mu_{c_k v_i}$ – сообщения, полученные переменным узлом от соседних контрольных узлов k .

– правило обновления SP, применяемое к контрольному узлу j для вычисления сообщения для переменного узла i :

$$\mu_{c_j v_i} = 2 \times \arctan h \left[\prod_{k \neq i} \tanh \left(\frac{\mu_{v_k c_j}}{2} \right) \right], \quad (2)$$

где $\mu_{v_k c_j}$ – сообщения, полученные контрольным узлом j от соседних переменных узлов k ;

– правило обновления SP, применяемое к переменному узлу i для вычисления бита i для последующего LLR:

$$\tilde{\lambda}_i = \lambda_i + \sum_k \mu_{c_k v_i}. \quad (3)$$

Поскольку ФГ не является циклическим, необходимо применить итеративное декодирование путем передачи сообщений туда и обратно между переменным и контрольным узлами, используя уравнения (1) и (2), прежде чем прийти к удовлетворительному решению. Алгоритм нейросетевого распространения доверия (NBP) предлагает научиться взвешивать сообщения и вводить LLR уравнений (1), (2) и (3), чтобы уменьшить негативное влияние коротких циклов на конечную производительность декодирования.

По мере увеличения количества слоев в нейронной сети улучшаются возможности обработки данных моделей. Искусственная нейронная сеть состоит из нескольких нейронов, соединяющих друг друга с другими ребрами. Нейронные сети обеспечивают входной уровень, несколько скрытых слоев

и выходной слой, что позволяет выполнять более точные и сложные задачи определения и классификации образов. Поскольку нейронная сеть содержит несколько скрытых слоев, она называется глубокой нейронной сетью. В моделях глубокой нейронной сети (ГНС) каждый нейрон и связанные с ним ребра имеют свои характерные особенности. На каждом нейроне есть функция активации и параметр смещения, а также параметр веса на ребре, соединенном с нейроном. Функция активации представляет собой нелинейное преобразование выходного сигнала нейрона. Входной сигнал умножается на вес, прибавляется к смещению, а затем производится для получения активного выходного сигнала. В ГНС вес и смещение силы используются для контроля связей между нейронами и смещения нейронов. Функции активации, вес и смещения в ГНС взаимодействуют друг с другом, и, регулируя их, можно контролировать силу связей, смещений и выходных результатов между нейронами, тем самым достигая обучения и прогнозирования нейронной сети.

Предлагается архитектура НС для выполнения алгоритма NBP, применяемого для декодирования линейных блочных кодов. В предлагаемом способе принимающему устройству больше не требуется предварительное знание используемой схемы кодирования. Как следствие, ему необходимо изучить как топологию ФГ, так и вес NBP. Предлагаемый декодер основан на пользовательской архитектуре ячеек рекуррентной нейронной сети (РНС), которая использует механизмы распределения веса и стробирования. Сначала правила SP будут описаны как эффективные матричные операции для универсальных линейных блочных кодов. Затем предлагаемые операции будут использоваться для РНС.

На каждой итерации NBP декодер запускается путем обновления сообщений из переменных в контрольные узлы $\mu_{v_i c_j}$, следуя рассчитанному по весу выражению (1) (входные сообщения, полученные от контрольных узлов $\mu_{c_j v_i}$, инициализируются значением 0 на первой итерации). Это вычисление может

быть реализовано с использованием плотного слоя НС, как показано на рис. 1, где $N_{\text{переменных}}$ и $N_{\text{проверочных}}$ обозначают количество переменных и контрольных узлов в ФГ, $\omega(A_j)$ – нелинейная функция, применяемая к параметрам факторного графа для представления как механизмов стробирования (т. е. бинарного выбора), так и механизмов взвешивания. Например, если $\omega(A_j)$ определено как ступенчатая функция, то вычислительный график будет воспроизводить стандартный алгоритм ВР с двоичными весами. Коэффициенты β_i не должны проходить через функцию ω , потому что нет необходимости в сокращении каких-либо входных данных.

Поскольку декодирование ВР является итеративным алгоритмом, вышеупомянутые операции должны повторяться несколько раз для достижения хорошей производительности. РНС может использоваться для выполнения такого итеративного декодирования в рамках НС.

Активация нейрона в слое, отличном от входного, представляет собой сумму произведений его входных данных и параметров веса,

соответствующих соединениям, которые вводят эти входные данные. Рассмотрим j -й нейрон в скрытом слое и примем $j = 2$. Если входной сигнал равен (1,1; 2,4; 3,2; 5,1; 3,9), а конечный выходной параметр равен (0,52; 0,75; 0,97), если веса, заданные для нейрона второго скрытого слоя, задаются вектором (-0,33; 0,07; -0,45; 0,13; 0,37), то активация будет произведена следующим образом:

$$(-0,33; 1,1) + (0,07; 2,4) + (-0,45; 3,2) + (0,13; 5,1) + (0,37; 3,9) = 0,471. \quad (4)$$

Теперь добавим к этому необязательное смещение или пороговое значение, например, 0,679, чтобы получить 1,15, и используем сигмовидную функцию, заданную $I(1 + \exp(-x))$, при $x = 1,15$ выходной сигнал этого нейрона скрытого слоя будет равен 0,7595. Если вычисленный выходной параметр также оказывается равным (0,61; 0,41; 0,57; 0,53), в то время как желаемое значение равно (0,52; 0,25; 0,75; 0,97). Очевидно, существует несоответствие между желаемым и вычисленным. Различия по компонентам приведены в векторе (-0,09; -0,16; 0,18; 0,44). Этот вектор используется для формиро-

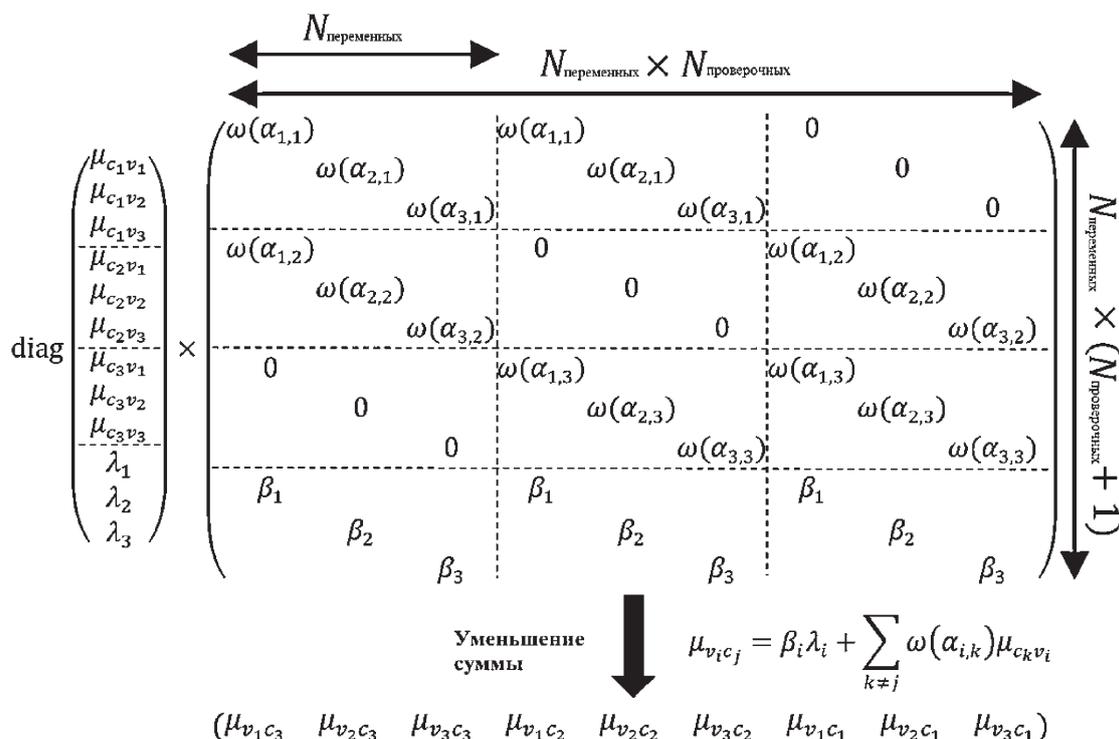


Рис. 1. Реализация взвешенной функции уравнения (1) с использованием плотного слоя
 Примечание: составлено авторами по источнику [8].

вания другого вектора, где каждый компонент является произведением компонента ошибки, соответствующего вычисленного компонента шаблона и дополнения последнего относительно 1. Для первого элемента ошибка равна $-0,09$, вычисленный элемент эталонного значения равен $0,61$, а его дополнение равно $0,39$. Умножая их вместе, получаем $-0,02$. Аналогично вычисляя другие компоненты, получаем вектор, равный $(-0,02; -0,04; 0,04; 0,11)$. Теперь необходимы веса соединений между вторым нейроном в скрытом слое и различными выходными нейронами. Если эти веса заданы вектором $(0,85; 0,62; -0,10; 0,21)$, то ошибка второго нейрона в скрытом слое может быть вычислена с использованием его выходных данных. Ошибка равна:

$$0,7595((1 - 0,7595)(0,85; -0,02) + (0,62; -0,04) + (-0,045)(-0,10; 0,4) + (0,21; 0,11)) = -0,0041. \quad (5)$$

Далее необходим параметр скорости обучения для этого уровня. Теперь оно установлено равным $0,2$. Умножив этот выходной сигнал второго нейрона на параметр скорости обучения, равный $0,2$ в скрытом слое, получим $0,1519$. Каждый из компонентов вектора $(-0,02; -0,04; 0,04; 0,11)$ теперь умножается на $0,1519$, что является последним достигнутым вычислением. Результатом является вектор, который корректирует веса соединений,

идущих от второго нейрона в скрытом слое к выходным нейронам. Эти значения заданы в векторе $(-0,003; -0,006; 0,006; 0,017)$. После добавления этих корректировок веса, которые будут использоваться в следующем цикле для связей между вторым нейроном в скрытом слое и выходными нейронами, становятся весами в векторе $(0,847; 0,614; \dots 0,094; 0,227)$.

На рис. 2 предлагается структурированная ячейка РНС, адаптированная для выполнения нейросетевого ВР-декодирования линейных блочных кодов.

Ячейка рекуррентной нейронной сети построена на основе двух типов обучаемых параметров веса:

- стробирующие веса: w_G представляют топологию факторного графа и используются для соответствующего выбора сообщений на различных этапах процесса декодирования. Чтобы представить такое поведение при бинарном выборе, к этим весам применяется сигмоидальная функция активации σ ;

- веса NBP: веса w_Σ и w_{OUT} используются для улучшения производительности схемы декодирования аналогично механизму NBP.

Все веса распределяются между различными итерациями РНС. Веса стробирования также распределяются внутри любой заданной итерации между различными операциями алгоритма ВР. Структурированная архитекту-

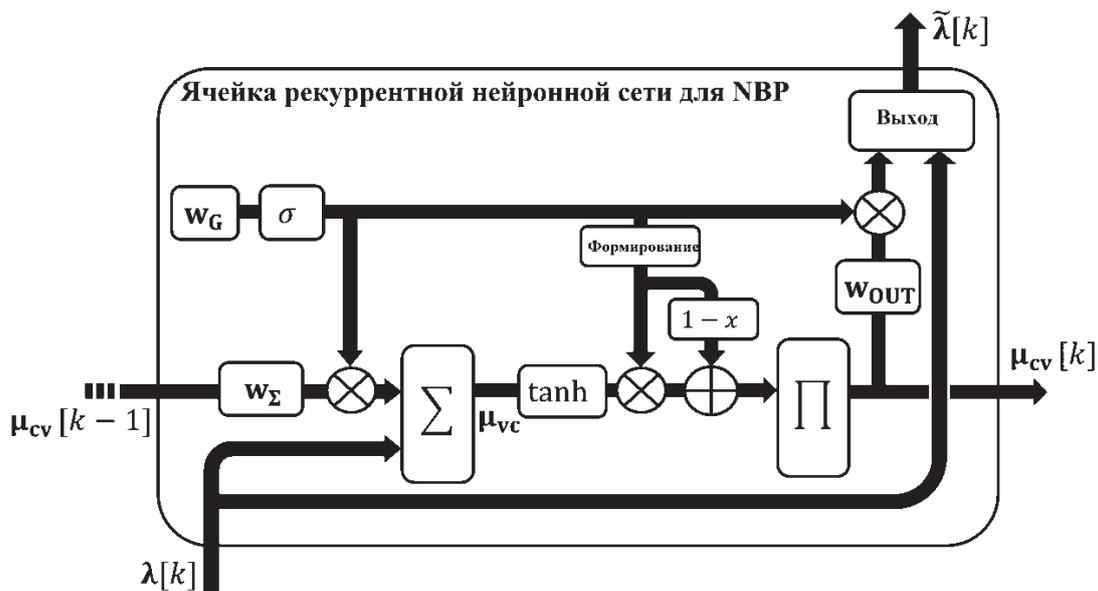


Рис. 2. Предлагаемая закрытая ячейка РНС для выполнения NBP-декодирования линейных блочных кодов
Примечание: составлено авторами по источнику [8].

ра НС гарантирует, что изученный алгоритм декодирования аналогичен NBP, но позволяет изучать топологию ФГ. В предлагаемой архитектуре вышеупомянутые функции ω , θ , γ и ψ могут быть определены следующим образом:

$$\begin{aligned}\omega(\alpha) &= w_{\Sigma} \otimes \sigma(w_G), \psi(\alpha) = w_{OUT} \otimes \sigma(w_G), \\ \theta(\alpha) &= \sigma(w_G), \gamma(\alpha) = 1 - \sigma(w_G).\end{aligned}\quad (6)$$

Система обработки сообщений выбирает сообщения, используя модифицированную функцию обработки $\sigma(w_G)$. Эта функция в значительной степени действует как сигмовидная функция, но обеспечивает лучшее обратное распространение градиента в режиме насыщения сигмовидной функции, тем самым улучшая обучаемость вентиля (за счет дополнительных обучаемых параметров). Можно отметить, что выбранная архитектура взвешивает сообщения только на определенных этапах. Использование разных параметров веса перед операциями и *OUT* кажется эффективным.

Чтобы оценить производительность предлагаемой ячейки РНС, в TensorFlow определяется сквозная модель НС со следующими настраиваемыми слоями. Системная модель показана на рис. 3, где ячейки, выделенные серым цветом, не поддаются обучению, а белым цветом – поддаются.

В данной системной модели рекуррентной нейронной сети подаваемые на вход информационные слова x кодируются с использованием систематических версий кодов Боуза – Чоудхури – Хоквингема БЧХ (15,11) или БЧХ (15,7), применяется двоичная фазовая модуляция (BPSK), затем символы кодового слова проходят через канал с аддитивным белым гауссовским шумом (АБГШ). С учетом используемой модуляции и канала вычисляется LLR принятых выборов.

Предварительная обработка LLR: для облегчения обучения РНС значения LLR нормализуются до диапазона $[-1, +1]$ в соответствии с максимальным абсолютным значением LLR в каждом кодовом слове. Затем LLR передается в соответствии с заданным числом итераций. К нормализованным кодовым словам применяется итерационный алгоритм расчета

весовых коэффициентов. Это могло бы позволить НС компенсировать недостаточный диапазон нормализации кодовых слов, с одной стороны, а также внести разнообразие в итерации РНС, изменив баланс между входными данными LLR и сообщениями из предыдущей итерации, с другой стороны. Предлагаемая закрытая ячейка NBP РНС используется для декодирования принятых кодовых слов. Нейросетевой алгоритм BP должен изучить схему кодирования, используемую в источнике.

При обработке выходных данных вместо извлечения только результата последней итерации BP выходные данные всех итераций объединяются повторно с использованием взвешенной суммы. Таким образом, уровень РНС используется в конфигурации «многие ко многим». Такой подход упрощает обратное распространение градиента, вводя его обратно на каждой итерации BP. Это также позволяет системе придавать большее значение результатам итерации, которые являются наиболее надежными с точки зрения опыта обучения. Зная положение систематических битов кода, затем применяется уровень определяемой выборки для подбора этих битов, и применяется сигмовидная функция для сжатия декодированного логарифмического коэффициента правдоподобия либо до 0, либо до 1.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Структура автоматического кодирования обеспечивает простой процесс обучения, поскольку функция потерь может быть вычислена как двоичная перекрестная энтропия между исходными информационными словами x и декодированными словами \tilde{x} . Для обоих кодов обучение выполняется с использованием случайно выбранных информационных битов, разделенных на слова размером k . Количество слов, используемых для обучения, в 10 раз превышает общее количество возможных слов, 2^k . Уровень шума, используемый на этапе обучения, зафиксирован на уровне 4 дБ. Механизм регуляризации со смещением ℓ_2 применяется для ограничения полученных весовых коэффициентов NBP, которые находятся слишком далеко от 1.

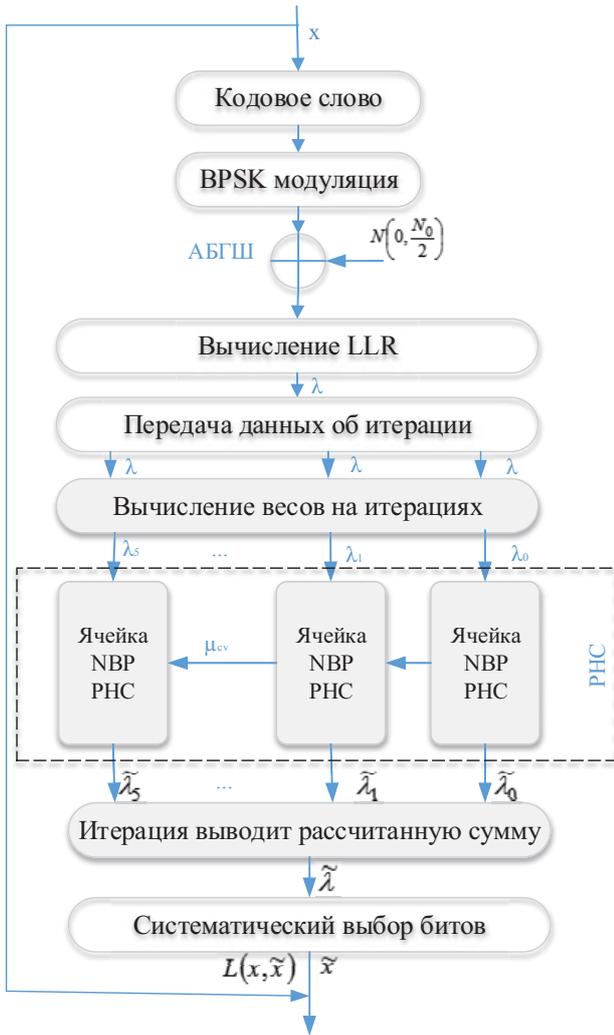


Рис. 3. Системная модель рекуррентной нейронной сети

Примечание: составлено авторами на основании данных, полученных в исследовании.

Модель оценивалась на наборе данных, состоящем из случайно выбранных слов. Для получения надежных результатов вместо фиксированного количества тестируемых слов было исправлено количество ошибок, которых необходимо достичь. Производительность лучшей модели среди 50 тренировок, обозначенной как «NBP PNC», сравнивается с базовыми показателями максимального правдоподобия (ML) и стандартного декодирования BP для обоих кодов, как показано на рис. 4.

Все кривые производительности отображаются в виде частоты ошибок блока (количество ошибочно декодированных кодовых слов среди всех обработанных кодовых слов).

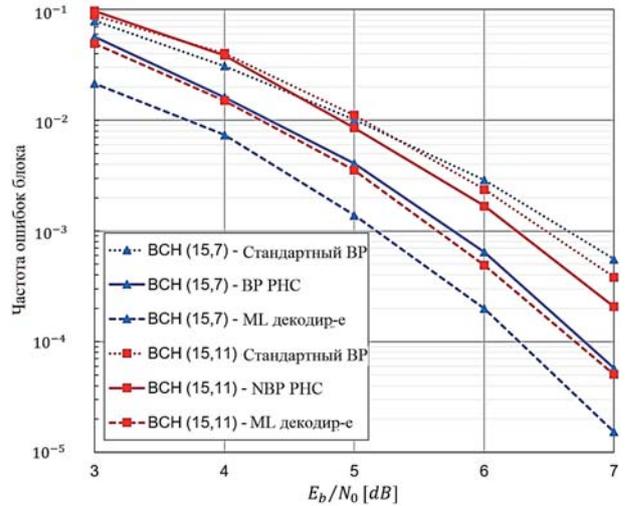


Рис. 4. Частота ошибок блока предложенной модели по сравнению со стандартными декодерами BP и максимального правдоподобия для кодов BCH (15,11) и BCH (15,7)

Примечание: составлено авторами на основании данных, полученных в исследовании.

Предлагаемая модель способна обучаться декодированию обоих кодов и превосходит стандартный алгоритм BP, особенно для кода BCH (15,7), где прирост производительности до 1 дБ значительно сокращает разрыв при декодировании с максимальным правдоподобием.

На рис. 5 приведена частота битовых ошибок для кода BCH с $N = 15$, а для BCH (15,7) и BCH (15,11) получено улучшение до 0,9 и 1,0 дБ соответственно.

ЗАКЛЮЧЕНИЕ

Представлена эффективная архитектура PNC со стробированием для декодирования линейных блочных кодов. Эта архитектура улучшает процесс итеративного декодирования BP, используя подход нейросетевого BP, при котором на приемной стороне не требуется предварительных знаний об используемой схеме кодирования. Вычислительный граф разработан несложным, универсальным и масштабируемым для больших кодов. Основными преимуществами разработанной архитектуры являются:

- контролируемое количество обучаемых параметров и вычислений в нейронной сети благодаря распределению веса и эффективно-му вычислительному графу;

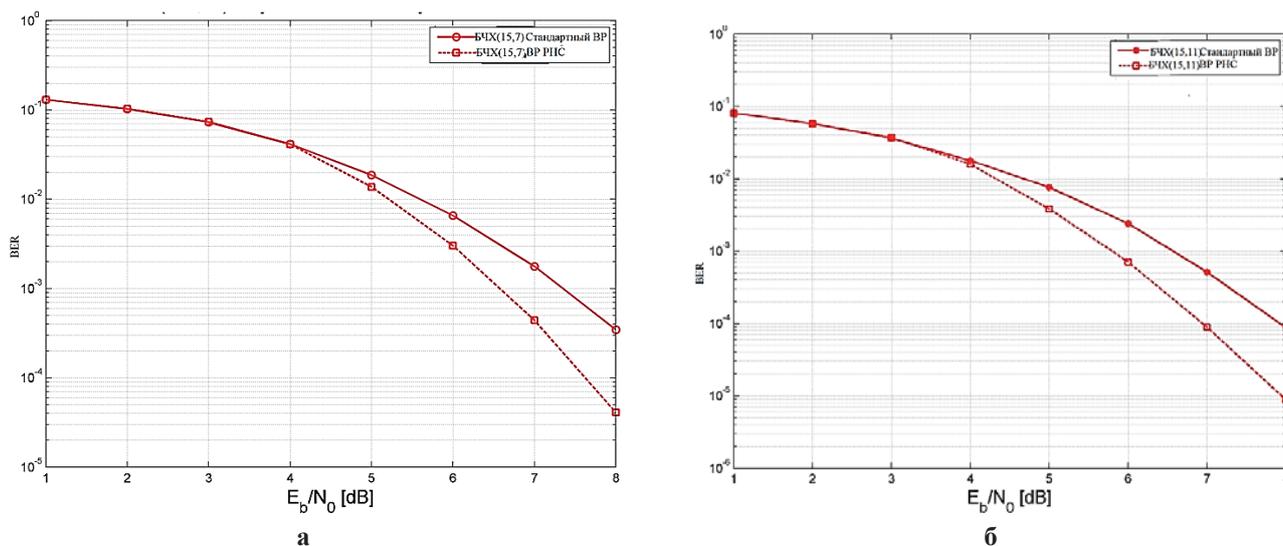


Рис. 5. Результаты BER для кода БЧХ: а – обученного с использованием матрицы проверки для алгоритмов стандартного ВР; б – нейросетевого ВР

Примечание: составлено авторами на основании данных, полученных в исследовании.

– структурированная архитектура, обеспечивающая применение алгоритма NBP для декодирования линейных блочных кодов;

– улучшенная производительность по сравнению со стандартным декодером ВР благодаря применению нейросетевого алгоритма ВР;

– повышенная гибкость декодера благодаря его возможностям обучения.

Декодер РНС может быть использован для повышения производительности и альтер-

нативного снижения вычислительной сложности алгоритма декодирования коротких и средней длины кодов в беспроводных системах передачи информации.

Планируется изучение возможностей такой структурированной архитектуры для схемы декодирования больших кодов, где невозможно предоставить все кодовые слова в обучающем наборе данных.

Список источников

1. Зинченко М. Ю., Левадний А. М., Гребенко Ю. А. Реализация LDPC декодера на ПЛИС и оптимизация потребляемой мощности // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14, № 3. С. 4–10.
2. Хорошайлова М. В. Архитектура канального кодирования на основе ПЛИС для 5G беспроводной сети с использованием высокоуровневого синтеза // Вестник Воронежского государственного технического университета. 2018. Т. 14, № 2. С. 99–105.
3. Хорошайлова М. В. Архитектура для стохастических LDPC-декодеров с использованием эффективной площади кристалла на основе ПЛИС // Вестник Воронежского государственного технического университета. 2018. Т. 14, № 1. С. 95–100.
4. Башкиров А. В., Хорошайлова М. В., Демикова А. С. Разработка архитектуры слепого распознавания линейного блочного кодирования с использованием каскадной сети // Вестник Воронежского государственного технического университета. 2023. Т. 19, № 6. С. 130–137.

References

1. Zinchenko M. Y., Levadnii A. M., Grebenko Y. A. FPGA LDPC decoder implementation and optimization its power consumption. *T-Comm*. 2020;14(3):4–10. (In Russ.).
2. Horoshaylova M. V. Architecture of FPGA based channel coding for 5G wireless network using high-level synthesis. *Bulletin of Voronezh State Technical University*. 2018;14(2):99–105. (In Russ.).
3. Horoshaylova M. V. Architecture for stochastic LDPC-decoders с using the effective chip area based on EPLD. *Bulletin of Voronezh State Technical University*. 2018;14(1):95–100. (In Russ.).
4. Bashkirov A. V., Khoroshailova M. V., Demikhova A. S. Development of a blind recognition architecture for linear block coding using a cascade network. *Bulletin of Voronezh State Technical University*. 2023;19(6):130–137. (In Russ.).
5. Bashkirov A. V., Muratov A. V., Khoroshaylova M. V. et al. Low-density codes low power decode. *Radioengineering*. 2016;5(5):32–37. (In Russ.).

5. Башкиров А. В., Муратов А. В., Хорошайлова М. В. и др. Низкоплотностные коды малой мощности декодирования // Радиотехника. 2016. № 5. С. 32–37.
6. Silver D., Huang A., Maddison C. J. et al. Mastering the game of Go with deep neural networks and tree search // Nature. 2016. Vol. 529, no. 7587. P. 484–489.
7. Овинников А. А. Способ идентификации циклов в графах Таннера LDPC кодов на основе пересечений коротких замкнутых структур в протографах // Цифровая обработка сигналов. 2016. № 4. С. 26–30.
8. Mandic D. P., Chambers J. A. Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability. London : Wiley, 2001. 308 p.
6. Silver D., Huang A., Maddison C. J. et al. Mastering the game of Go with deep neural networks and tree search. *Nature*. 2016;529(7587):484–489.
7. Ovinnikov A. A. Images of cycles of cyclic liftings in the base graph of protograph LDPC codes. *Digital Signal Processing*. 2016;(4):26–30. (In Russ.).
8. Mandic D. P., Chambers J. A. Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability. London: Wiley, 2001. 308 p.

Информация об авторах

А. А. Пирогов – кандидат технических наук, доцент.

М. В. Хорошайлова – кандидат технических наук, доцент.

Э. В. Сёмка – кандидат физико-математических наук, доцент.

About the authors

A. A. Pirogov – Candidate of Sciences (Engineering), Docent.

M. V. Khoroshailova – Candidate of Sciences (Engineering), Docent.

E. V. Syomka – Candidate of Sciences (Physics and Mathematics), Docent.

Научная статья
УДК 004.94:008.2
<https://doi.org/10.35266/1999-7604-2024-3-7>



Оптимизация при вероятностном тематическом моделировании технологической прогностической информации

Олег Русланович Попов¹✉, Сергей Олегович Крамаров²

¹Академия информатизации образования, Ростов-на-Дону, Россия

²Сургутский государственный университет, Сургут, Россия

¹cs41825@aaanet.ru✉, <https://orcid.org/0000-0001-6209-3554>

²kramarov_so@surgu.ru, <https://orcid.org/0000-0003-3743-6513>

Аннотация. На основе анализа методов мягкой кластеризации документов и вероятностных распределений терминов и тем рассмотрены вычислительные методы и инструменты моделирования динамики политематических потоков в многомерном информационном пространстве. Предложена оптимизированная стохастическая модель динамики мягкой кластеризации сетей знаний в информационном пространстве, структурированном на основе семантических связей в текстах определенной предметной области, извлеченных из наукометрических и библиографических баз данных. На основе теоретической модели разработан алгоритм и методика его применения, с помощью которой возможно применение расширенной текстовой аналитики, включая выявление скрытых тем и прогнозирование трендов. Разработанная методика позволяет с определенным уровнем научной объективности осуществлять прогнозирование новых технологий и актуальных научных направлений в заданной определенной предметной исследовательской области, в том числе для решения теоретических, прикладных и управленческих задач. На основании практических результатов, полученных в работе, разработан глоссарий прогностических терминов «Информационные технологии и коммуникации», который рекомендован к применению в учебном процессе системы общего и профессионального образования.

Ключевые слова: тематическая модель, вероятность, алгоритм, стохастическая модель, информация, семантика, текст, прогностический термин, информационно-коммуникационные технологии

Для цитирования: Попов О. Р., Крамаров С. О. Оптимизация при вероятностном тематическом моделировании технологической прогностической информации // Вестник кибернетики. 2024. Т. 23, № 3. С. 56–69. <https://doi.org/10.35266/1999-7604-2024-3-7>.

Original article

Optimization in probabilistic topic modeling of technological predictive information

Oleg R. Popov¹✉, Sergey O. Kramarov²

¹Academy of Informatization of Education, Rostov-on-Don, Russia

²Surgut State University, Surgut, Russia

¹cs41825@aaanet.ru✉, <https://orcid.org/0000-0001-6209-3554>

²kramarov_so@surgu.ru, <https://orcid.org/0000-0003-3743-6513>

Abstract. The analysis of soft clustering methods of documents and probabilistic distributions of terms and topics leads us to consider computational methods and tools for modeling the dynamics of polytopic flows in a multidimensional information space. We propose an optimized stochastic model that captures the dynamics of soft clustering of knowledge networks in an information space. This model is based on semantic connections in texts of a specific subject area, which are extracted from scientometric and bibliographic databases. Using the theoretical model, we developed an algorithm and methodology for applying advanced text analytics, which includes the identification of hidden topics and the prediction of trends. The developed methodology

allows, with a certain level of scientific objectivity, to predict new technologies and current scientific directions in a given specific research area, including for solving theoretical, applied and management problems. The practical results led to the development of a glossary of predictive terms “Information Technologies and Communications”. We recommend using this glossary in the educational process of the general and vocational education system.

Keywords: topic model, probability, algorithm, stochastic model, information, semantics, text, predictive term, information and communication technologies

For citation: Popov O. R., Kramarov S. O. Optimization in probabilistic topic modeling of technological predictive information. *Proceedings of Cybernetics*. 2024;23(3):56–69. <https://doi.org/10.35266/1999-7604-2024-3-7>.

ВВЕДЕНИЕ

Развитие стохастического анализа доказало важность подхода, основанного на стохастическом описании природы многих сложных физических и технологических систем. Наблюдается расширение использования вероятностных методов и для прогнозирования развития научных областей. В настоящее время сформировался целый блок научно-технических дисциплин, имеющих общую системную ориентацию, задающую относительно них особую плоскость существования искусственно создаваемых сложных систем. Особый интерес представляют междисциплинарные области взаимных пересечений наук и технологий. На этих стыках используются инструменты одной области для продвижения другой.

Наиболее развитыми в динамике процесса представляются информационно-коммуникационные технологии (ИКТ). Чаще всего именно эти технологии поставляют инструменты для развития других технологий через возможности имитационного компьютерного моделирования различных процессов.

Для описания процессов, учитывающих сложные нелинейные механизмы, необходимо использование специального подхода, связанного с вероятностным тематическим моделированием. Вероятностные тематические модели осуществляют «мягкую кластеризацию» (soft clustering), относя документ к нескольким кластерам-темам с некоторыми вероятностями.

Проблема вероятностного моделирования динамики мягкой кластеризации тематических информационных потоков, особенно с учетом внутри- и межструктурных взаимодействий, а также многих метаданных (мо-

дальностей), признаются открытыми научными проблемами [1–3].

Актуальной практической задачей является проблема моделирования стохастических процессов взаимодействия политематических информационных потоков, структурированных на основе семантических связей определенной предметной области, для прогнозирования новых технологий и научных направлений. В качестве предметной области выбраны прогнозируемые ИКТ, которые находятся в кластерах пересечения научных инноваций и дают рекомендации для исследований по другим зрелым или появляющимся технологиям [4].

МАТЕРИАЛЫ И МЕТОДЫ

Основные вычислительные методы и инструменты, приводящие к автоматическому режиму обнаружения знаний и скрытых ассоциаций в публикациях, включая тематическое моделирование, подробно классифицируются в источнике [5]. Анализ на уровне темы распространения дополняет процесс поиска и фильтрации информации, помимо анализа на уровне термов.

Тематические модели используются в различных предметных областях и для различных задач, включая, например, менеджмент знаний [6], анализ кластеризации сообществ [7], автоматическое выявление новых новостных тем [8]. Достаточно редки публикации, отражающие использование данных методов для анализа научно-технологических инноваций.

Аналитический метод, позволяющий сгруппировать связанные термины и фразы, определить меняющиеся тематические акценты

в сфере информатики и больших данных, составить технологическую дорожную карту (technology roadmap, TRM) предложен в источнике [9]. TRM прогнозирует будущие изменения в технологических темах и получает информацию для планирования научно-исследовательских и опытно-конструкторских работ, а также для стратегического управления. Однако данная работа представляет собой гибридный набор различных методик при отсутствии явной вычислительной модели. Для прогнозирования применяется сложный в реализации экспертный подход без автоматизации данных. Двумерное отображение семантической близости тем в виде «дорожной карты» является весьма упрощенным инструментом визуализации [10].

В источнике [11] на основе анализа обширной коллекции статей, опубликованных с 2000 по 2021 гг. на конференциях по машинному обучению и искусственному интеллекту (ИИ), реализуется задача ранней детекции трендовых научных тем. Обосновывается преимущество предлагаемого инкрементального метода вероятностного тематического моделирования, реализуемого на основе модели ARTM, в сравнении с популярными байесовскими и нейросетевыми подходами. В качестве примеров выявленных трендовых тем в машинном обучении приводятся «LSTM», «deep learning», «word2vec», «BERT», «fake news detection». Однако очевидные прикладные результаты работы модели не приводятся.

Потенциальной возможностью вероятностных подходов для решения актуальных прикладных научных задач в сфере технологий ИИ является информация о том, что в лаборатории машинного обучения и семантического анализа Института искусственного интеллекта МГУ обучили нейронную сеть для получения семантических векторных представлений (эмбеддингов) научных текстов на русском языке SciRus-tiny [12]. Это реализовано с помощью данных, предоставленных для обучения порталом eLIBRARY.RU, который содержит большое количество документов по множеству разнообразных научных тематик.

Полученные результаты показывают, что сжатое описание документа в виде вектора вероятностей тем содержит важнейшую информацию о семантике документа и может использоваться для решения многих нетривиальных задач текстовой аналитики, включая обнаружение латентной кластерной структуры в документах и выявление трендов в библиографических и патентных базах данных.

Вероятностная тематическая модель (probabilistic topic model) выявляет тематику коллекции документов, представляя каждую тему дискретным распределением вероятностей терминов, а каждый документ – дискретным распределением вероятностей тем.

Исходными данными для тематического моделирования является множество (коллекция) текстовых документов D и множество (словарь) терминов (термов) W . Под терминами понимаются слова, нормальные формы слов, словосочетания или термины, в зависимости от того, какие виды предварительной обработки текстов были выполнены.

Каждый документ $d \in D$ представляется последовательностью термов (w_p, \dots, w_n^d) из W , где n_d – длина документа. Через n_{dw} обозначается число вхождений термина w в документ d .

Существует конечное множество тем T и коллекция порождается дискретным распределением $p(d, w, t)$ на $D \times W \times T$. Документы d и термины w являются наблюдаемыми переменными, тема t – латентной переменной, т. е. каждое вхождение термина w в документ d связано с некоторой неизвестной темой t из заданного конечного множества T .

Для каждой темы t и документа d зададим вероятность темы в документе $p(t|d)$. То же самое сделаем для слов и тем: $p(w|t)$ – вероятность встретить слово w в теме t . Распределение вероятностей термов $p(w|d, t)$ зависит только от темы t , но не от документа d (гипотеза условной независимости):

$$p(w|d, t) = p(w|t). \quad (1)$$

Распределение терминов в документе $p(w|d)$ описывается вероятностной смесью распределений терминов в темах $\varphi_{ot} = p(w|t)$ с весами $\theta_{td} = p(t|d)$:

$$p(w|d) = \sum_{t \in T} p(w|d,t)p(t|d) = \sum_{t \in T} p(w|t)p(t|d) = \sum_{t \in T} \varphi_{wt} \theta_{td} \quad (2)$$

Вероятностная порождающая или генеративная модель (2) описывает процесс порождения коллекции по известным распределениям $p(w|t)$ и $p(t|d)$.

Построить тематическую модель коллекции является обратной задачей [13], что означает найти по заданной коллекции D множество тем T , условные распределения термов $\varphi_{wt} = p(w|t)$ для каждой темы $t \in T$ с весами $\theta_{td} = p(t|d)$ для каждого документа $d \in D$.

Равенство (2) можно представить в виде матрицы $P = (p_{dw})_{W \times D}$ частот слов документа (как часто каждое слово встречается в каждом документе), которая затем записывается как произведение двух матриц меньших размеров $\Phi = (\varphi_{wt})_{W \times T}$ – матрицы термов тем и $\Theta = (\theta_{td})_{T \times D}$ – матрицы тем документов. Матрица P известна, так как это исходные данные, и имеет в общем случае полный ранг, поэтому не может быть в точности равна $\Phi\Theta$. Правая часть равенства представляет собой произведение двух неизвестных матриц. Построение вероятностной тематической модели является решением задачи поиска приближенного низкорангового стохастического матричного разложения $P \approx \Phi\Theta$ (рис. 1).

Столбцы этих матриц в целом можно рассматривать как нормализованные, то есть задача декомпозиции матрицы может быть приведена к максимизации логарифма правдоподобия [14]:

$$\sum_{d \in D} \sum_{w \in d} n_{dw} \ln \sum_{t \in T} \varphi_{wt} \theta_{td} \rightarrow \max_{\Phi, \Theta} \quad (3)$$

при ограничениях неотрицательности и нормировки всех столбцов $\varphi_{wt}, \theta_{td}$:

$$\sum_{w \in W} \varphi_{wt} = 1; \quad \varphi_{wt} \geq 0;$$

$$\sum_{t \in T} \theta_{td} = 1; \quad \theta_{td} \geq 0. \quad (4)$$

Решение задачи матричного разложения будет неуникальным. Согласно теории регуляризации А. Н. Тихонова [15], решение некорректно поставленной операторной задачи возможно доопределить и сделать устойчивым. Для этого следует наложить некоторые дополнительные ограничения на $R(\Phi, \Theta)$, называемые регуляризаторами. Функция регуляризатора должна быть непрерывно дифференцируемой. К задаче максимизации логарифмического правдоподобия будет до-

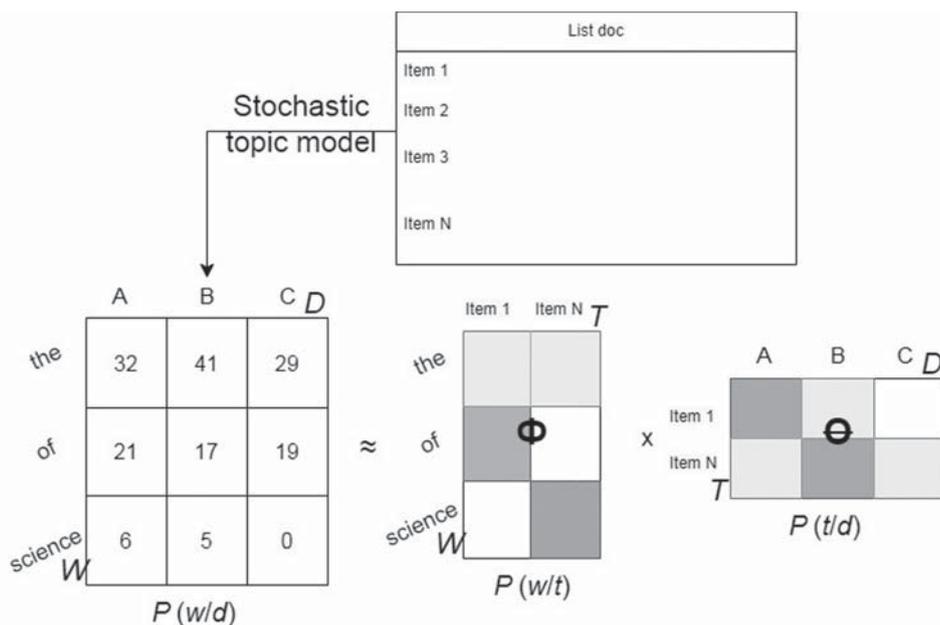


Рис. 1. Построение тематической модели: решение задачи приближенного стохастического матричного разложения $P \approx \Phi\Theta$

Примечание: составлено авторами на основании данных, полученных в исследовании.

бавлен новый компонент, и задача оптимизации примет вид:

$$\sum_{d \in D} \sum_{\omega \in d} n_{d\omega} \ln \sum_{t \in T} \phi_{\omega t} \theta_{td} + R(\Phi, \Theta) \rightarrow \max_{\Phi, \Theta}, \quad (5)$$

Первая модель, называемая pLSA (вероятностный латентный семантический анализ), предложенная в источнике [16], предполагала, что $R(\Phi, \Theta) = 0$.

В дальнейшем широкое применение получили два подхода к регуляризации вероятностных генеративных моделей: скрытое распределение Дирихле (latent Dirichlet allocation, LDA) [17], основанный на байесовском выводе, и, более классическая, аддитивная регуляризация тематических моделей, которая, как следует из названия, лежит в основе ARTM.

Модель LDA предполагает байесовскую регуляризацию таким образом, что:

$$\sum_{d \in D} \sum_{t \in T} (\alpha_t - 1) \ln \theta_{td} + \sum_{d \in D} \sum_{t \in T} (\alpha_t - 1) \ln \theta_{td}. \quad (6)$$

Здесь β_ω и α_t являются гиперпараметрами с положительной настройкой. Модель LDA предполагает, что векторы документа θ_{td} генерируются одним и тем же распределением вероятностей на нормализованных векторах, а распределение берется из семейства распределений Дирихле с параметром α [7]. Аналогично, векторы темы $\phi_{\omega t}$ генерируются распределением Дирихле с параметром β .

Распределение Дирихле существенно упрощает байесовский вывод, и большинство моделей строятся с его использованием.

Однако следует обратить внимание на ряд проблем, связанных с байесовской регуляризацией: необходимо оптимизировать гиперпараметры, инициализировать параметр β_ω для новых терминов и обеспечивать разреженность, при том, что обнулять параметры $\phi_{\omega t}$ и θ_{td} невозможно.

Введение других инструментов приближенного байесовского вывода (вариационный вывод, семплирование Гиббса, распространение ожидания) не позволяет легко комбинировать

модели и снимать ограничения, связанные с выбором распределений Дирихле. Для каждой новой модели приходится заново выполнять математические выкладки и программную реализацию [13].

Альтернативой байесовскому подходу является метод аддитивной регуляризации тематических моделей (ARTM) [18]. Это приложение классической теории регуляризации некорректно поставленных задач [15] к тематическому моделированию.

Аддитивная регуляризация тематических моделей (ARTM) основана на максимизации линейной комбинации логарифма правдоподобия и нескольких регуляризаторов $R_i(\Phi, \Theta)$, $i = 1, \dots, k$:

$$\sum_{d \in D} \sum_{w \in d} n_{dw} \ln \sum_{t \in T} \varphi_{wt} \theta_{td} + \underbrace{\sum_{i=1}^k \tau_i R_i(\Phi, \Theta)}_{R(\Phi, \Theta)} \rightarrow \max_{\Phi, \Theta}; \quad (7)$$

при прежних ограничениях (4), где τ_i – неотрицательные коэффициенты регуляризации.

Построение многофункциональных тематических моделей существенно упрощается благодаря аддитивности регуляризаторов [18]. Функционал максимизации правдоподобия (3) позволяет добавить не один метод регуляризации, а несколько, придавая вес каждому. Этот подход выражает суть аддитивности регуляризации и, как следует из названия, лежит в основе ARTM.

Таким образом, ARTM это не инкрементное улучшение одной тематической модели, а общий подход к тематическому моделированию как к задаче многокритериальной оптимизации.

В подходе ARTM распределение Дирихле является не универсальным, а одним из возможных регуляризаторов. В качестве базовой модели логичнее брать pLSA, не имеющую собственных регуляризаторов, которые можно добавлять из модульной расширяемой библиотеки в зависимости от поставленной проблемы.

С точки зрения поставленных в работе прикладных исследовательских задач, такой подход дает преимущества. В первую очередь это

касается оптимизации стратегий построения иерархических тематических моделей, включая динамические модели, которые выявляют закономерности развития кластеров не только внутри мультидисциплинарных корпусов, но и с течением времени.

При формировании базовых принципов построения многофункциональной тематической модели (далее «Модель») необходимо учитывать максимум дополнительной информации, особенности семантики и предмета текстовой коллекции. С точки зрения стратегий оптимизации при вероятностном тематическом моделировании выделяются следующие факторы:

1. Автоматическая генерация графов.

2. Ранжирование при оценке результатов поиска.

3. Учет отсутствующей в классических алгоритмах ранжирования силы связей между словами и предложениями и других семантических метрик.

4. Учет при моделировании сетевой топологии показателя центральности, содержащего необходимый и достаточный объем информации, и показателя промежуточности, определяющего степень передачи информации без потерь и искажений.

5. Привязка документа ко времени создания для оценки динамики развития событий, выявления закономерностей и прогнозирования.

6. Учет при построении модели точек сближения автоматической суммаризации и тематического подхода к обработке коллекций документов.

Общая схема многофункциональной модели обработки информационных процессов, рассматриваемой как большой научный проект science data, представлена на рис. 2. В Модели выделяются три блока, характерные также для разработки и проектирования проекта big data с использованием технологий глубокого анализа текстов (Text Mining) и нахождения при этом закономерностей и трендов:

1. Подготовка и упорядочение данных.

2. Импорт данных, их обработка и генерация новой информации.

3. Валидация и оценка качества результатов моделирования.

В рамках данных блоков сформированы пять взаимосвязанных этапов алгоритма Модели:

1. Предварительная информационная экспертиза.

2. Формирование понятийного ядра предметной области и концептуальных тематических запросов.

3. Автономный сбор и накопление метаданных коллекций документов.

4. Многофункциональная обработка данных, генерация и визуализация информации.

5. Контроль, оценка и валидация данных моделирования.

Заданный алгоритмический подход преследует цель обеспечить максимальную объ-



Рис. 2. Общая схема моделирования обработки информационных процессов как научного проекта science data

Примечание: составлено авторами на основании данных, полученных в исследовании.

ективность поиска и скорость, с которой он позволяет углубиться в выбранную предметную область исследований.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

На первом подготовительном этапе практической реализации Модели используется авторская методика расчета показателя уровня зрелости самоорганизующихся интеллектуальных систем «SML», интегрирующего в себе показатели «технологические уровни готовности» (TPL) и «уровни социоэкономической адаптированности» (SPL) прогнозируемых технологий, входящих в исследуемую социотехническую систему [19].

В целях определения базовых структур предметной области экспертным методом выделены категории перспективных направлений развития ИКТ, соответствующих определенному уровню SML-матрицы. Например, в таблице приведены выделенные в рамках 4 категорий 22 перспективных направления

развития информационно-коммуникационных технологий, соответствующих четвертому уровню:

1. Человеко-машинные интерфейсы (human-computer interfaces);
2. Инженерия вычислений (computing engineering);
3. Технологии памяти и хранения данных (memory and data storage technologies);
4. Электроника и коммуникации (electronics and communications) [20].

Далее, исходя из предложенного подхода к формированию понятийного ядра предметной области, разработан алгоритм, применимый к задаче формирования «словаря прогностических терминов» [21].

Данный алгоритм позволяет на базе внешних сервисов (например, «Википедии») формировать словарь контекстно-близких терминов по отношению к изначально заданному термину. При этом задается набор изначально терминов, и по результатам их обработки

Таблица

Категории перспективных направлений развития ИКТ, соответствующих четвертому уровню SML-матрицы

I	Human – computer interfaces	
1	ambient intelligence	интеллектуальная среда
2	brain – computer interface	интерфейс «мозг– компьютер»
3	city brain	городской мозг
4	semantic web	семантический веб
5	smart city	умный город
II	Computing engineering	
1	exascale computing	масштабные вычисления
2	neuromorphic engineering	нейроморфная инженерия
3	optical computing	оптические (фотонные) вычисления
4	quantum computing	квантовые вычисления
III	Memory and data storage technologies	
1	3D optical data storage	3D оптическое хранение данных
2	DNA digital data storage	цифровое хранилище данных ДНК
3	holographic data storage	голографическое хранилище данных
4	patterned media	узорчатые носители
5	phase-change memory	память с фазовым переходом
6	quantum memory	квантовая память
IV	Electronics and communications	
1	atomtronics	атомтроника
2	carbon nanotube field-effect transistor	полевой транзистор из углеродных нанотрубок
3	Li-Fi (Light Fidelity)	Li-Fi
4	memristor	мемристор, мемтранзистор, мемистор, транситор
6	software-defined radio	программно-определяемое радио
7	spintronics	спинтроника, твистроника, валлейтроника

Примечание: составлено по источнику [21].

информацию. Фрагмент основных информационных полей (Global Number, Time, Title, Authors, Abstract) библиографических данных по запросу «quantum computing» в БД Google Scholar показан на рис. 4.

Сбор документов осуществляется в автоматическом режиме, с обработкой полученных исходных данных в соответствии с определенными правилами и сохранением метаданных. Выходные данные препроцессора документа представляют собой набор «чистого» текстового контента со связанными метаданными.

Обработанные текстовые документы передаются на вход многофункциональной системы обработки данных, генерации и визуализации информации.

Модуль тематического моделирования (BigARTM, Gensim-LDA) автоматически выводит набор скрытых тем. Результаты анализа темы, включая ключевое слово темы, дистрибутивы и распределения по темам документов, также сохраняются и индексируются во время индексации тем.

В результате работы алгоритма на данном этапе темпоральная (датированная) коллекция текстовых документов систематизируется по двум уровням:

- в соответствии с базовыми концепциями тематического запроса;
- в соответствии с проведенным тематическим анализом документов, сформированных внешней БД в ответ на запрос.

Автономный и сетевой режимы ориентированы на решение как фундаментальных исследовательских задач, так и прикладных интересов пользователей информационного ресурса в процессе активного взаимодействия с системой. В данном основном блоке реализованы базовые процессы обработки big data, генерации новой информации и ее визуального представления с целью нахождения при этом закономерностей и прогностических трендов.

Чтобы отслеживать появление новых тем, для каждого временного шага в БД обновляются данные. При поступлении новой порции документов D' словарь пополняется новыми терминами W' и могут образоваться новые

темы T' . Основным принципом выявления научных трендов заключается в том, что новая лексика, появившаяся в новых документах, относится преимущественно к новым темам [11].

Темпоральное исследование кластерообразования, связанного с научной областью квантовых вычислений (quantum computing) в период 1980–2023 гг., приведено на рис. 5.

Анализ динамики тематических данных, интегрированных суммарно в 10 кластеров, показывает активное кластерообразование в новейший период (2020–2023 гг.) в области кластеров С6 («новые технологии»), С7 («квантовая нейронная сеть») и С10 («квантовое шифрование изображения») (рис. 3). Так, область кластера С7 пополняется такими терминами, как «квантовая информация», «квантовые данные», «классификация изображений», «процессор», «глубокая нейронная сеть», «сверточная нейронная сеть». В области кластера С10 можно обнаружить «шифрование цветного изображения», «логистическая карта», «схема квантового шифрования изображения», «квантовая репрезентативная модель». Использование новой, специфичной для данной научной области, лексики релевантно отражает постепенное вхождение новых квантовых информационных технологий, новых архитектур квантовых процессоров и симуляторов, устройств передачи квантовой информации, отражающих сложную динамику новейших квантовых систем и коммуникаций.

Особенностью выбранной методики валидации данных является рассмотрение, помимо классических мер оценки тематического моделирования (перплексии, когерентности, разреженности), критериев, одновременно выступающих индикаторами развития исследовательской модели.

Исходя из выбранного подхода к построению Модели, предложена модифицированная методика количественной оценки релевантности динамически выявленных скрытых тем начальному запросу (ключевому слову или концепту) [22].

Критерием релевантности темы понятию является существование между ними взаимно однозначного соответствия. Для выявле-

Global Number	Title	Authors	Abstract
513	Toward implementing efficient image processing algorithms on quantum computers	[Fei Yan, Salvador E. Venegas-Andrade, Kaoru Hirota]	Quantum information science is an interdisciplinary subject spanning quantum computational complexity, quantum information to black holes and back
527	Hybridization of Moth flame optimization algorithm and quantum computing for gene selection in m	[Shira Chapman, Giuseppe Polcastro]	Ever-increasing data in various fields like Bioinformatics field, which
528	Quantum federated learning through blind quantum computing	[Ali Dabbia, Abdelkamel Tari, Samy Mefrafi]	Private distributed learning studies the problem of how multiple dist
529	Balanced k-means clustering on an adiabatic quantum computer	[Wei-Kang Li, Siro Liu, Dong-Ling Deng]	Adiabatic quantum computers are a promising platform for efficient
530	Unsupervised event classification with graphs on classical and photonic quantum computers	[Davis Arthur, Prassanna Date]	Photonic Quantum Computers provide several benefits over the dis
531	Quantum Computing and Deep Learning Methods for GDP Growth Forecasting	[Andrew Blance, Michael Spornowski]	Precise macroeconomic forecasting is one of the major aims of econ
534	Designing nanotechnology QCA-multiplexer using majority function-based NAND for quantum com	[David Alaminos, M. Belén Salas, Manuel A. Fernández-Gómez]	Quantum-dot cellular automata (QCA) are one of the most signific
535	Collective neutrino oscillations on a quantum computer	[Jun-Cheol Jeon]	We calculate the energy levels of a system of neutrinos undergoing
536	Toward simulating superspring/M-theory on a quantum computer	[Kilira Yeter-Aydeniz, Shikha Bangar, George Sopsis, Raphael C. Poeser]	We present a novel framework for simulating matrix models on a q
537	Constant-depth circuits for dynamic simulations of materials on quantum computers	[Hrant Gharibyan, Masahiro Hanada, Masazumi Honda, Junyu Liu]	A dynamic simulation of materials is a promising application for near-
538	Methods for accelerating geospatial data processing using quantum computers	[Lindsay Bassman O'Reilly, Ruel Van Beurumen, Ed Younis, Ethan Smith, Costin Iancu, Wibe A. T. Dynamic simulation of materials is a promising application for near-	
539	Compiling basic linear algebra subroutines for quantum computers	[Max Henderson, Jared Galina, Michael Brett]	Quantum computing is a transformative technology with the poten
542	Hybrid quantum-classical convolutional neural networks	[Liming Zhao, Zhikuan Zhao, Patrick Rebertrout, Joseph Fitzsimons]	Efficiently processing basic linear algebra subroutines is of great im
544	NISQ computing: where are we and where do we go?	[Junhua Liu, Kwan-Hu Lim, Kristin L. Wood, Wei Huang, Chu Guo, He-Liang Huang]	Deep learning has been shown to be able to recognize data pattern
546	Unconstrained binary models of the travelling salesman problem variants for quantum optimization	[Jonathan Wei Zhong Liu, Kian Hwee Lim, Harshank Shrotriya, Leong Chuan Kwek]	In this short review article, we aim to provide physicists with worki
551	Review of Quantum Image Processing	[Oysem Salehi, Adam Glos, Jaroslaw Adam Mitzczak]	Quantum computing is offering a novel perspective for solving com
552	Emerging Enabling Technologies for Industry 4.0 and Beyond	[Zhaobin Wang, Mizhe Xu, Yaonan Zhang]	As an interdisciplinary between quantum computing and image pro
553	Event Classification with Quantum Machine Learning in High-Energy Physics	[Alexander Sigov, Leonid Ratiuk, Leonid A. Ivanov, U Da Xu]	Rapid advances in technology have spurred tremendous progress in
554	Cloud based QC with Amazon Braket	[Koji Terashi, Michiru Kaneda, Tomoe Kishimoto, Masahiko Saito, Ryu Sawada, Junichi Tanaka]	We present studies of quantum algorithms exploiting machine lear
555	Quantum K-means clustering method for detecting heart disease using quantum circuit approach	[Constantin Gonzalez]	The approaches to solutions in quantum computing technology are
556	Quantum machine learning for particle physics using a variational quantum classifier	[S.S Kavitha, Narasimha Kalugudi]	The development of noisy intermediate-scale quantum computers
557	Quantum machine learning for the constrained portfolio optimization probl	[Andrew Blance, Michael Spornowski]	Quantum machine learning aims to release the prowess of quantum
560	Quantum machine learning: a novel technique for the constrained portfolio optimization probl	[Ameer Tamoor Khan, Xinwei Cao, Shuai Li, Bin Hu, Vasilios N. Katsikis]	In this paper, we have formulated quantum beetle antennae search
561	Quantum machine learning: a novel technique for the constrained portfolio optimization probl	[Zhang Chen, Mordcha Segev]	Let there be light-to change the world we want to be! Over the pas
562	Quantum Error Correction: Noise-Adapted Techniques and Applications	[Bila A. Alkawayk, Hemanth B. Mahajan, Hadi Ilhan, Reza Mohammed Qase]	Secure communication has developed into one of the most promisi
563	Quantum Support Vector Machines for Continuum Suppression in B Meson Decays	[Akshaya Jayashankar, Prabha Mandayam]	The quantum computing devices of today have tens to hundreds of
565	Optimizing quantum heuristics with meta-learning	[Jamie Heredge, Charles Hill, Lloyd Hollenberg, Marcin Seiwior]	Quantum computers have the potential to speed up certain comput
566	Quantum Machine Learning Classification Using Quantum Annealing for Real-World Applications	[Max Wilson, Rachel Stromswold, Filip Wudarski, Stuart Hadfield, Norm M. Tubman, Eleanor C. Variational quantum algorithms, a class of quantum heuristics, are	
567	Quantum Machine Learning Classification Using Quantum Annealing for Real-World Applications	[Rajdeep Kumar Nath, Himanshu Thapliyal, Travis S. Humble]	Optimizing the training of a machine learning pipeline helps in redu
569	Quantum convolutional neural network on NISQ devices	[Shiue Wei, Yanhu Chen, Zengfeng Zhou, Guoli Long]	Quantum machine learning is one of the most promising applicator
570	Quantum Machine Learning Architecture for COVID-19 Classification Based on Synthetic Data Gener	[Javaria Amin, Muhammad Sharif, Nadia Gul, Selfridine Kabry, Chinmay Chakraborty]	COVID-19 is a novel virus that affects the upper respiratory tract, ai
571	Binary quantum-inspired gravitational search algorithm-based multi-criteria scheduling for multi-pro	[Abhijeet Singh Thakur, Tarun Biswas, Pratyay Kulia]	A quantum-inspired hybrid scheduling technique is proposed for mu
572	Probabilistic nonunitary gate in imaginary time evolution	[Tong Liu, Jin-Guo Liu, Heng Fan]	Simulation of quantum materials is a significant application of quan
573	Quantum K-Nearest-Neighbor Image Classification Algorithm Based on K-L Transform	[Han-Sun Zhou, Xu-Xun Liu, Yu-Ling Chen, Xi-Suo Du]	Enlightened by quantum computing theory, a quantum k-nearest-
574	Quantum simulation of gauge theory via orbifold lattice	[Alexander J. Buser, Hrant Gharibyan, Masazumi Honda, Junyu Liu]	We propose a new framework for simulating U(1) Yang-Mills theo
575	Measurements of Quantum Hamiltonians with Locally-Biased Classical Shadows	[Charles Haffner, Sergey Bravyi, Rudy Raymond, Antonio Mezzacapo]	Obtaining precise estimates of quantum observables is a crucial ste
576	Robust implementation of generative modeling with parametrized quantum circuits	[Vicente Lefort-Otteg, Alejandro Pedrono-Oritz, Oscar Pedrono]	Although the performance of hybrid quantum-classical algorithms is
577	Efficient quantum simulation of open quantum dynamics at various Hamiltonians and spectra densi	[Na-Ke Zhang, Yaoliu Chen, Ming-Jie Yao, Jun Fan, Ru Zhang]	Simulation of open quantum dynamics for various Hamiltonians an
578	Quantum secure multi-party summation protocol based on blind matrix and quantum Fourier transfor	[Xin Yi, Dong Cao, Ling Fan, Jun Zhang]	Multi-party quantum summation is the premise for implementing o
580	Color image encryption algorithm based on hyperchaotic system and improved quantum revolution	[Xinwei Sun, Chao Luo, Fuzhong Nian, Lun Tang]	Because quantum commutator can break encryption systems based,

Рис. 4. Фрагмент вывода библиографического описания по запросу «quantum computing»

Примечание: составлено авторами на основании данных, полученных в исследовании.

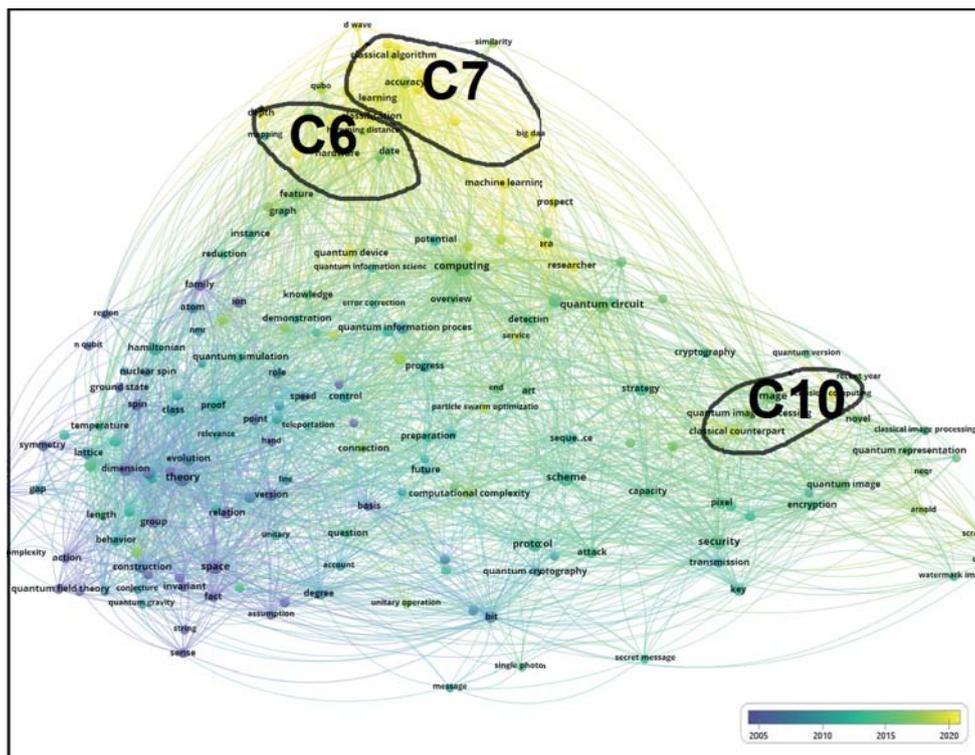


Рис. 5. Карта визуализации VOSviewer, показывающая динамику кластеризации тематических данных в период 1980–2023 гг. в области термина «quantum computing» (квантовые вычисления): цветовая шкала в правом нижнем углу визуализации, отображающая временную оценку данных, варьируется фиолетовым (до 2005 г.), зеленым (период 2010–2015 гг.), желтым (после 2020 г.)
Примечание: составлено авторами на основании данных, полученных в исследовании.

ния степени смещения выявляются 4 типа смещения:

- модель не выявляет темы;
- модель производит смешанные темы;
- ключевые слова в теме отсутствуют;
- повторяются среди скрытых тем.

На первом этапе автоматизированная оценка сходства между темами и концепциями рассчитывается с использованием трех распространенных мер сходства: косинусной, ранговой корреляции Спирмена и KL-дивергенции.

Далее проводится этап сравнительной экспертной оценки совпадений тем и концепций и преобразования оценок сходства в вероятности совпадения.

Составляется матрица размера $n \times m$ всех возможных пар между n ключевыми словами и m выявленными темами. Каждая запись $p(s, t)$ рассматривается как независимая случайная величина Бернулли, представляющая вероятность совпадения того, что эксперт,

изучающий распределения соответствия понятия s с темой t , ответит, что они релевантны. Каждой паре $n \times m$ присваивается рейтинг $\{1, 0,5, 0\}$ для каждого ответа {совпадение, частичное совпадение, отсутствие совпадения}. Пара считается совпадающей, если ее средний рейтинг превышает 0,5.

В результате применения данной методики количественно оценивается и анализируется вероятность соответствия концепта прогнозируемой технологии латентным политематическим потокам, извлеченным и структурированным из научно-технического информационного пространства.

Для дальнейшего анализа качества стохастического моделирования при кластеризации данных и максимизации выявления трендов новых технологий проводится серия экспериментов, где сравнительно рассматриваются вероятностные тематические модели, такие как PLSA, LDA, ARTM [18] с базовыми регуляризаторами матрицы Φ ,

и глубокие нейронные сети, в частности BERTopic [11, 12].

В рамках заданной предметной области будут исследованы возможности интеграции тематического моделирования с глубокими нейросетевыми моделями языка, моделями внимания и архитектуры трансформеров.

ЗАКЛЮЧЕНИЕ

1. На основе анализа методов мягкой кластеризации документов и вероятностных распределений терминов и тем рассмотрены вычислительные методы и инструменты моделирования динамики политематических потоков в многомерном информационном пространстве.

2. Предложена оптимизированная стохастическая модель динамики мягкой кластеризации сетей знаний в информационном пространстве, структурированном на основе семантических связей в текстах определенной предметной области, извлеченных из наукометрических и библиографических баз данных.

3. На основе теоретической модели разработан алгоритм и методика его применения, с помощью которой возможно применение расширенной текстовой аналитики, включая

выявление скрытых тем и прогнозирование трендов.

4. Разработанная методика позволяет с определенным уровнем научной объективности осуществлять прогнозирование новых технологий и актуальных научных направлений в заданной определенной предметной исследовательской области, в том числе для решения теоретических, прикладных и управленческих задач.

5. На основании практических результатов, полученных в работе, разработан глоссарий прогностических терминов «Информационные технологии и коммуникации», который рекомендован к применению в учебном процессе системы общего и профессионального образования в целях наполнения контента при изучении учебных дисциплин.

6. Проведенное темпоральное исследование кластерообразования, связанное с научной областью квантовых вычислений (quantum computing), отражает сложную динамику новейших квантовых систем и коммуникаций и подтверждает постепенное вхождение новых квантовых информационных технологий.

Список источников

1. Shadrova A. Topic models do not model topics: epistemological remarks and steps towards best practices // *Journal of Data Mining & Digital Humanities*. 2021. <https://doi.org/10.46298/jdmdh.7595>.
2. Churchill R., Singh L. The evolution of topic modeling // *ACM Computing Surveys*. 2022. Vol. 54, no. 10s. P. 1–35. <https://doi.org/10.1145/3507900>.
3. Zhao H., Phung D., Huynh V. et al. Topic modelling meets deep neural networks: A survey // *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*. 2021. P. 4713–4720. <https://doi.org/10.48550/arXiv.2103.00498>.
4. Бодрунов С. Д. Ноономика : моногр. М. : Культурная революция, 2018. 432 с.
5. Thilakarathne M., Falkner K., Atapattu T. A systematic review on literature-based discovery: general overview, methodology, & statistical analysis // *ACM Computing Surveys*. 2019. Vol. 52, no. 6. P. 1–34. <https://doi.org/10.1145/3365756>.
6. Zelenkov Yu. The topic dynamics in knowledge management research // *Knowledge Management in Organizations (KMO 2019): Proceedings of the 14th International Conference*. 2019. P. 324–335. https://doi.org/10.1007/978-3-030-21451-7_28.

References

1. Shadrova A. Topic models do not model topics: epistemological remarks and steps towards best practices. *Journal of Data Mining & Digital Humanities*. 2021. <https://doi.org/10.46298/jdmdh.7595>.
2. Churchill R., Singh L. The evolution of topic modeling. *ACM Computing Surveys*. 2022;54(10s):1–35. <https://doi.org/10.1145/3507900>.
3. Zhao H., Phung D., Huynh V. et al. Topic modelling meets deep neural networks: A survey. In: *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*. 2021:4713–4720. <https://doi.org/10.48550/arXiv.2103.00498>.
4. Bodrunov S. D. Noonomika: Monograph. Moscow: Kulturnaya revolyutsiya, 2018. 432 p. (In Russ.).
5. Thilakarathne M., Falkner K., Atapattu T. A systematic review on literature-based discovery: general overview, methodology, & statistical analysis. *ACM Computing Surveys*. 2019;52(6):1–34. <https://doi.org/10.1145/3365756>.
6. Zelenkov Yu. The topic dynamics in knowledge management research. In: *Proceedings of the 14th International Conference “Knowledge Management in Organizations (KMO 2019)”*. 2019:324–335. https://doi.org/10.1007/978-3-030-21451-7_28.

7. Gorshkov S., Ilyushin E., Chernysheva A. et al. Using topic modeling for communities clusterization in the VKontakte social network // *International Journal of Open Information Technologies*. 2021. Vol. 9, no. 5. P. 12–17.
8. Zhang J., Ghahramani Z., Yang Y. A probabilistic model for online document clustering with application to novelty detection // *Advances in neural information processing systems*. 2004. Vol. 17.
9. Zhang Y., Zhang G., Chen H. et al. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research // *Technological forecasting and social change*. 2016. Vol. 105. P. 179–191.
10. Айсина Р. М. Обзор средств визуализации тематических моделей коллекций текстовых документов // *Машинное обучение анализ данных*. 2015. Т. 1, № 11. С. 1584–1618.
11. Герасименко Н. А., Чернявский А. С., Никифорова М. А. и др. Инкрементальное обучение тематических моделей для поиска трендовых тем в научных публикациях // *Доклады Российской академии наук. Математика, информатика, процессы управления*. 2022. Т. 508, № 1. С. 106–108.
12. Герасименко Н. ruSciBench – бенчмарк для оценки эмбедингов научных текстов. URL: <https://habr.com/ru/articles/781032/> (дата обращения: 25.03.2024).
13. Большакова Е. И., Воронцов К. В., Ефремова Н. Э. и др. Автоматическая обработка текстов на естественном языке и анализ данных. М. : НИУ ВШЭ, 2017. 268 с.
14. Воронцов К. В., Потапенко А. А. Регуляризация, робастность и разреженность вероятностных тематических моделей // *Компьютерные исследования и моделирование*. 2012. Т. 4, № 4. С. 693–706.
15. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. 3-е изд., испр. М. : Наука, 1986. 286 с.
16. Hofmann T. Probabilistic latent semantic indexing // *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. 1999. P. 50–57.
17. Blei D. M., Ng A. Y., Jordan M. I. Latent dirichlet allocation // *Journal of Machine Learning Research*. 2003. Vol. 3. P. 993–1022.
18. Воронцов К. В., Потапенко А. А. Аддитивная регуляризация тематических моделей // *Доклады Академии наук*. 2014. Т. 456, № 3. С. 268–271.
19. Попов О. Р. Адаптация мировых практик к проблеме долгосрочного технологического прогнозирования состояния самоорганизующихся интеллектуальных систем // *Интеллектуальные ресурсы – региональному развитию*. 2021. № 2. С. 91–98.
20. Крамаров С. О., Попов О. Р., Джариев И. Э. и др. Динамика формирования связей в сетях, структурированных на основе прогностических терминов // *Russian Technological Journal*. 2023. Т. 11, № 3. С. 17–29. <https://doi.org/10.32362/2500-316X-2023-11-3-17-29>.
7. Gorshkov S., Ilyushin E., Chernysheva A. et al. Using topic modeling for communities clusterization in the VKontakte social network. *International Journal of Open Information Technologies*. 2021;9(5):12–17.
8. Zhang J., Ghahramani Z., Yang Y. A probabilistic model for online document clustering with application to novelty detection. *Advances in neural information processing systems*. 2004;17.
9. Zhang Y., Zhang G., Chen H. et al. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. *Technological forecasting and social change*. 2016;105:179–191.
10. Aysina R. M. Survey of visualization tools for topic models of text corpora. *Machine Learning and Data Analysis*. 2015;1(11):1584–1618. (In Russ.).
11. Gerasimenko N. A., Chernyavskiy A. S., Nikiforova M. A. et al. Inkrementalnoe obuchenie tematicheskikh modeley dlya poiska trendovykh tem v nauchnykh publikatsiyakh. *Doklady Rossijskoj akademii nauk. Matematika, informatika, processy upravleniya*. 2022;508(1):106–108. (In Russ.).
12. Gerasimenko N. ruSciBench – benchmark dlya otsenki embeddingov nauchnykh tekstov. URL: <https://habr.com/ru/articles/781032/> (accessed: 25.03.2024). (In Russ.).
13. Bolshakova E. I., Vorontsov K. V., Efremova N. E. et al. Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i analiz dannykh. Moscow: NIU VSHE, 2017. 268 p. (In Russ.).
14. Vorontsov K. V., Potapenko A. A. Regularization, robustness and sparsity of probabilistic topic models. *Computer Research and Modeling*. 2012;4(4):693–706. (In Russ.).
15. Tikhonov A. N., Arsenin V. Ya. Metody resheniya nekorrektnykh zadach. 3d ed., revised. Moscow: Nauka, 1986. 286 p. (In Russ.).
16. Hofmann T. Probabilistic latent semantic indexing. In: *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. 1999:50–57.
17. Blei D. M., Ng A. Y., Jordan M. I. Latent dirichlet allocation. *Journal of Machine Learning Research*. 2003;3:993–1022.
18. Vorontsov K. V., Potapenko A. A. Additivnaya regulyarizatsiya tematicheskikh modeley. *Doklady Akademii nauk*. 2014;456(3):268–271. (In Russ.).
19. Popov O. R. Adaptatsiya mirovykh praktik k probleme dolgosrochnogo tekhnologicheskogo prognozirovaniya sostoyaniya samoorganizuyushchikhsya intellektualnykh system. *Intellektualnye resursy – regionalnomu razvitiyu*. 2021;(2):91–98. (In Russ.).
20. Kramarov S. O., Popov O. R., Dzhariyev I. E. et al. Dynamics of link formation in networks structured on the basis of predictive terms. *Russian Technological Journal*. 2023;11(3):17–29. <https://doi.org/10.32362/2500-316X-2023-11-3-17-29>. (In Russ.).

21. Попов О. Р., Гросу А., Крамаров С. О. Комплексный сетевой алгоритм формирования глоссария контекстно-близких прогностических терминов // Современные информационные технологии и ИТ-образование. 2023. Т. 19, № 3. URL: <http://sitito.cs.msu.ru/index.php/SITITO/article/view/999> (дата обращения: 25.03.2024).
22. Chuang J., Gupta S., Manning C. et al. Topic model diagnostics: Assessing domain relevance via topical alignment // International conference on machine learning. 2013. P. 612–620.
21. Popov O. P., Grosu A., Kramarov S. O. Complex network algorithm for glossary formation context-related predictive terms. *Modern Information Technologies and IT-Education*. 2023;19(3). URL: <http://sitito.cs.msu.ru/index.php/SITITO/article/view/999> (accessed: 25.03.2024). (In Russ.).
22. Chuang J., Gupta S., Manning C. et al. Topic model diagnostics: Assessing domain relevance via topical alignment. *International conference on machine learning*. 2013:612–620.

Информация об авторах

О. Р. Попов – кандидат технических наук, доцент.

С. О. Крамаров – доктор физико-математических наук, профессор.

About the authors

O. R. Popov – Candidate of Sciences (Engineering), Docent.

S. O. Kramarov – Doctor of Sciences (Physics and Mathematics), Professor.

Научная статья
УДК 621.791.048.5
<https://doi.org/10.35266/1999-7604-2024-3-8>



Анализ влияния шероховатости паяльной маски и формы трафарета на образование пустот в паяных соединениях

**Андрей Владимирович Турецкий¹, Александр Александрович Пирогов²,
Ирина Владимировна Свиридова³, Марина Владимировна Хорошайлова⁴,
Элеонора Викторовна Сёмка⁵**

^{1, 2, 3, 4}Воронежский государственный технический университет, Воронеж, Россия

⁵Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. профессора Н.Е. Жуковского и Ю.А. Гагарина» Минобороны России, Воронеж, Россия

¹tav7@mail.ru, <https://orcid.org/0000-0002-5338-367X>

²Pirogov.alx@gmail.com, <https://orcid.org/0000-0002-5105-0505>

³ri-ss-ka@mail.ru, <https://orcid.org/0000-0001-5279-0807>

⁴pmv2205@mail.ru, <https://orcid.org/0000-0001-9167-9538>

⁵semka_elya@mail.ru, <https://orcid.org/0000-0002-0194-6979>

Аннотация. Проведен анализ причин появления и выявлены пути минимизации пустот, которые образуются внутри паяных соединений. Возникающие в процессе пайки пустоты существенно снижают как механическую, так и электрическую прочность соединения. Кроме того, существенно ухудшается теплопроводность, особенно в корпусах компонентов, где паяное соединение выполняет также роль теплоотвода.

Предложены два новых подхода, ведущих к уменьшению образования пустот в паяных соединениях. Первый подход заключается в изменении шероховатости паяльной маски, влияющей на распределение флюса вокруг контактной площадки. Показано, что большая шероховатость паяльной маски положительно влияет на лучшую растекаемость флюса, выделяющегося при оплавлении припоя. При этом пятно флюса оказывается больше и он меньше попадает в паяное соединение, что приводит к минимизации объема и количества пустот.

Второй подход заключается в нанесении не одного большого пятна паяльной пасты, а нескольких маленьких, разделенных небольшими промежутками. При этом для сохранения объема дозы припоя предложено использовать более толстые трафареты нанесения паяльной пасты. Наличие промежутков между пятнами паяльной пасты способствуют лучшему вытеснению газов, выделяющихся при оплавлении.

Ключевые слова: надежность радиоэлектронных средств, монтаж электронных компонентов, пустоты в паяных соединениях, флюс

Финансирование: работа выполнена в рамках функционирования дизайн-центра «Силовая электроника» ФГБОУ ВО «Воронежский государственный технический университет».

Для цитирования: Турецкий А. В., Пирогов А. А., Свиридова И. В., Хорошайлова М. В., Сёмка Э. В. Анализ влияния шероховатости паяльной маски и формы трафарета на образование пустот в паяных соединениях // Вестник кибернетики. 2024. Т. 23, № 3. С. 70–81. <https://doi.org/10.35266/1999-7604-2024-3-8>.

Original article

Analysis of the influence of solder mask roughness and stencil shape on the formation of voids in solder joints

Andrey V. Turetsky¹, Aleksandr A. Pirogov², Irina V. Sviridova³,
Marina V. Khoroshailova⁴, Eleonora V. Syomka⁵✉

^{1, 2, 3, 4}Voronezh State Technical University, Voronezh, Russia

⁵Military Educational and Scientific Centre of the Air Force N. E. Zhukovsky and Y. A. Gagarin Air Force Academy the Ministry of Defence of the Russian Federation, Voronezh, Russia

¹tav7@mail.ru, <https://orcid.org/0000-0002-5338-367X>

²pirogov.alx@gmail.com, <https://orcid.org/0000-0002-5105-0505>

³ri-ss-ka@mail.ru, <https://orcid.org/0000-0001-5279-0807>

⁴pmv2205@mail.ru, <https://orcid.org/0000-0001-9167-9538>

⁵semka_elya@mail.ru ✉, <https://orcid.org/0000-0002-0194-6979>

Abstract. The analysis of the causes of occurrence and ways of minimizing voids that are formed inside soldered joints is carried out. Voids that form during soldering significantly reduce both the mechanical and electrical strength of the joint. In addition, thermal conductivity significantly decreases, especially in component housings, where the soldered joint also acts as a heatsink.

We propose two alternative approaches leading to a decrease in the formation of voids in soldered joints. The first approach involves changing the roughness of the solder mask, which affects the distribution of flux around the contact pad. Research shows that a higher roughness of the solder mask positively affects the better spreading of the flux released during solder reflow. Here, the flux spot is larger and less of it gets into the soldered joint, which leads to a minimization of the volume and number of voids.

The second approach includes applying not one large amount of solder paste, but several small ones, separated by small gaps. Moreover, to maintain the volume of the solder dose, it is proposed to use thicker stencils to apply solder paste. The presence of gaps between the spots of solder paste contributes to better displacement of gases released during reflow.

Keywords: reliability of radio-electronic equipment, installation of electronic components, voids in solder joints, flux

Funding: the work is carried out alongside the Educational Design Center for Electronics “Silovaya elektronika” at the Voronezh State Technical University.

For citation: Turetsky A. V., Pirogov A. A., Sviridova I. V., Khoroshailova M. V., Syomka E. V. Analysis of the influence of solder mask roughness and stencil shape on the formation of voids in solder joints. *Proceedings in Cybernetics*. 2024;23(3):70–81. <https://doi.org/10.35266/1999-7604-2024-3-8>.

ВВЕДЕНИЕ

Паяные соединения выполняют несколько функций – электрическую коммутацию компонентов, механическую фиксацию их на печатной плате и теплоотвод. Одним из серьезных дефектов паяных соединений является образование пустот (пустоты, образующиеся внутри соединений). Они оказывают негативное влияние на все типы соединений, упомянутые выше, и, таким образом, способствуют снижению надежности. Исследования показывают, что пустоты в паяных

соединениях существенно снижают механическую прочность соединения [1]. Влияние профиля пайки на пустотообразования и механическую прочность получаемых паяных соединений, при использовании четырех паяльных паст (BiSnAg, SAC305, Sn100C, SnSb) и трех покрытий поверхности печатной платы (ENIG, иммерсионное Ag, иммерсионное Sn), рассматривал N. Zhu [2]. Различные виды пустот припоя оказывают совершенно различное влияние на общий тепловой импеданс корпуса. Большие сросшиеся пусто-

ты оказывают более существенное влияние, чем маленькие распределенные. Длинные теплогенерирующие области/объемы типа полосок, связанные с полупроводниковыми лазерными чипами, приводят к сильной чувствительности к ориентации пустот припоя под ними [3]. Качество поверхности паяльной площадки оказывает большое влияние на качество паяных соединений после оплавления [4]. Микроминиатюризация элементов и создание функционально сложных микроэлектронных устройств и микропроцессоров вызывают особые проблемы в области микромонтажа изделий электроники. Поэтому для повышения надежности функционирования изделий электроники необходимо уменьшить количество и размеры пустот в паяных соединениях.

Цель – разработка технологических методов уменьшения пустот в паяных соединениях компонентов на печатной плате с использованием двух типов флюса, трафаретов разного типа и толщины.

МАТЕРИАЛЫ И МЕТОДЫ

1. Анализ пустот и их образование в паяных соединениях компонентов на печатной плате

Пустоты можно разделить на несколько групп. К ним относятся микропустоты, макропустоты (внутренние полости), дырчатые пустоты, усадочные пустоты, пустоты Киркендалла и пустоты в микропереходах. Расположение и форма указанных видов пустот показаны на рис. 1.

Однозначного мнения о допустимых размерах макропустот нет. Пустоты диаметром менее 50 мкм называют микропустотами [5].

Макропустоты возникают в результате технологического процесса пайки. Типичный процесс технологии поверхностного монтажа начинается с трафаретной печати, когда паяльная паста наносится на контактные площадки на печатной плате. Процесс печати имеет множество переменных параметров, таких как условия окружающей среды, форма и размеры апертуры отверстий трафарета, текучесть паяльной пасты [6]. Из-за множества

влияющих факторов процесс образования пустот не до конца изучен. В частности, замечено, что недостаточное качество паяльных паст, их грубое раскатывание и заполнение отверстий трафарета могут стать причиной поглощения пузырьков воздуха. Это может привести к плохому осаждению паяльной пасты на печатной плате и дальнейшему образованию пустот при пайке оплавлением. Д. Хе, Н. Н. Экере, М. А. Карри рассмотрели поведение паяльной пасты под действием вибрирующего ракеля для оптимизации параметров процесса [7].

После процесса нанесения паяльной пасты через трафарет следует установка компонентов с помощью автоматической машины. Компоненты вдавливаются в паяльную пасту с усилием, предварительно заданным монтажной головкой. Неправильно установленная сила прижима может привести к изменению формы ранее нанесенной паяльной пасты и, как следствие, к увеличению количества воздушных пузырьков внутри паяльной пасты [8].

После этого печатная плата с установленными компонентами поступает в конвекционную печь, где происходит оплавление паяльной пасты. Многие исследования показали, что образование пустот сильно зависит от параметров пайки, включая зону предварительного нагрева, пиковую температуру и время выдержки [9]. Направление теплового потока

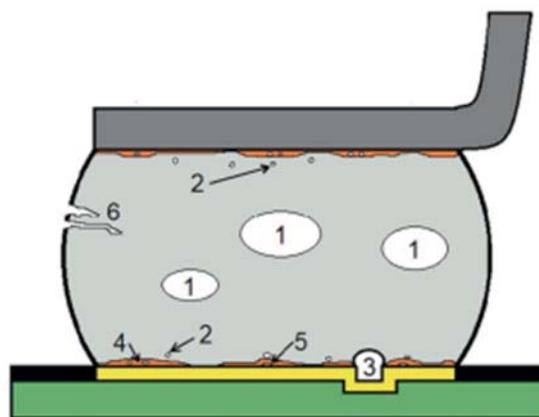


Рис. 1. Расположение различных типов пустот в паяном соединении:

- 1 – макропустоты; 2 – микропустоты; 3 – пустоты в микропереходах; 4 – пустоты Киркендалла;
 - 5 – дырчатые пустоты; 6 – усадочные пустоты
- Примечание: составлено авторами по источнику [5].

и охлаждение также могут оказывать влияние на свойства и структуру получаемых паяных соединений и образование пустот, что было подтверждено в работах [10–12].

Появление пустот вызвано газовойделением флюса и других газов, которые не могут выйти из паяного соединения при оплавлении. Испарение реологических добавок и растворителей из паяльной пасты в процессе нагрева также является источником газовойделения [13].

Еще одним источником образования пустот во время оплавления припоя может быть реакция флюса с металлизацией паяемой детали или с оксидами, присутствующими на поверхностях контактных площадок или в паяльной пасте. Поведение пустот изучалось с использованием печатной платы с микропереходными отверстиями и контактными площадками, компонентов PBGA (пластиковый корпус с матрицей шариковых выводов), а также путем измерения пустот в паяных соединениях [14]. В частности выявлено, что финишное покрытие типа ОЗП (органическое защитное покрытие) способствовало образованию пустот в большей степени, чем ENIG (Ni/Au).

В эксперименте использовались паяльные пасты с флюсами на основе канифоли. Классификация флюсов в соответствии ГОСТ Р МЭК 61190.1-1-2020 приведена в табл. 1 [15].

Стандарт IPC-610 устанавливает критерий приемлемости для пустот: менее 25 % площади на рентгеновском изображении. Паяные соединения являются допустимыми, если размер и количество пустот находятся на низком

уровне, что устраняет потенциальный риск плохого проводящего контакта.

2. Два способа по уменьшению образования пустот в паяных соединениях

Первый способ по уменьшению образования пустот в паяных соединениях основан на идее уменьшения количества флюса, присутствующего в паяном соединении во время фазы жидкого и твердого состояния. Для этого необходим более высокий уровень вывода флюса в окружающую поверхность паяльной площадки, что может быть достигнуто путем изменения равновесия поверхностного натяжения между жидким флюсом, окружающим воздухом и поверхностью паяльной маски. Смачивание и растекание паяльной пасты по контактной площадке в значительной степени зависит от свойств жидкости и поверхности печатной платы, на которую сильно влияет шероховатость паяльной маски. При более шероховатой паяльной маске флюс лучше смачивает окружающую поверхность вокруг контактной площадки, а площадь растекания больше [16]. Это может привести к уменьшению количества флюса в объеме жидкого сплава и, таким образом, снизить вероятность попадания флюса внутрь соединения.

Второй способ заключается в изменении конструкции стального трафарета для нанесения паяльной пасты. Вместо одного отверстия размером с контактную площадку, модифицированный трафарет содержит два или четыре отверстия меньшего размера. С точки зрения сохранения того же объема осаждаемой пасты, что и для обычного трафарета, модифицированный трафарет сделан толще.

Таблица 1

Классификация флюсов, состав, уровни активности

Флюс	Уровни активности флюса		Обозначение флюса в стандарте МЭК*
	Низкий (< 0,01)	L0	
Канифоль	Низкий (< 0,15)	L1	ROL1
	Средний (< 0,01)	MO	ROM0
	Средний (0,15–2,0)	M1	ROM1
	Высокий (< 0,01)	HO	ROH0
	Высокий (> 2,0)	H1	ROH1

Примечание: МЭК* – материалы электронных модулей. Составлено авторами по источнику [15].

Такая конструкция должна вызвать более высокую активность растекания и перемещения паяльной пасты по контактной площадке во время процесса расплавления, что способствует вытеснению большего количества пузырьков газа из припоя. Кроме того, такая конструкция создает воздушные зазоры и каналы, через которые частицы газообразного флюса более эффективно удаляются из паяного соединения.

Для оценки паяемости металлов применяются методы, связанные с определением площади или высоты растекшегося припоя по паяемой поверхности. Коэффициент растекания припоя K_p в первом случае определяется как:

$$K_p = \frac{S_p}{S_0}, \quad (1)$$

где S_p и S_0 – соответственно площади растекшегося припоя и его исходной дозы.

Во втором случае K_p может быть определен с помощью отношения:

$$K_p = \frac{D_K - H_K}{D_K}, \quad (2)$$

где D_K – диаметр капли припоя;

H_K – высота капли растекшегося припоя.

Величина D_K определяется из предположения отсутствия влияния силы тяжести исходной дозы припоя:

$$D_K = \sqrt{\frac{6 \cdot m}{\pi \cdot \rho}}, \quad (3)$$

где m – масса дозы припоя;

ρ – плотность припоя в расплавленном состоянии.

Значения коэффициентов растекания припоя, определенные по приведенным выражениям, оказываются различными, поскольку в первом случае:

$$S_p > S_0 \text{ и } K_p > 1, \quad (4)$$

а во втором случае:

$$H_K > D_K \text{ и } K_p < 1. \quad (5)$$

Для проведения эксперимента была разработана специальная печатная плата размером 46 мм × 33 мм. Всего было изготовлено и про-

анализировано 90 плат, по 5 плат для каждой комбинации типа трафарета, паяльной маски и паяльной пасты.

Печатные платы содержали контактные площадки для 20 резисторов типоразмера 2010. Они были разделены на три группы: одна без паяльной маски и две с паяльными масками, различающимися шероховатостью. Одна из них была матовая, а другая – глянцевая. Для обеспечения одинаковой дозы паяльной пасты разработаны несколько трафаретов, так трафарет 1 имеет толщину 0,1 мм, а два других (трафареты 2 и 3) – 0,2 мм (рис. 2).

Это обеспечило нанесение одинакового объема паяльной пасты.

Использовались две паяльные пасты с одинаковым составом припойного сплава, различающиеся по активности флюса. Флюс ROL1 обладает более высокой химической активностью в отношении восстановления оксидов, чем ROL0. Полные параметры паяльных паст, использованных в эксперименте, приведены в табл. 2, где паяльная паста с флюсом ROL1

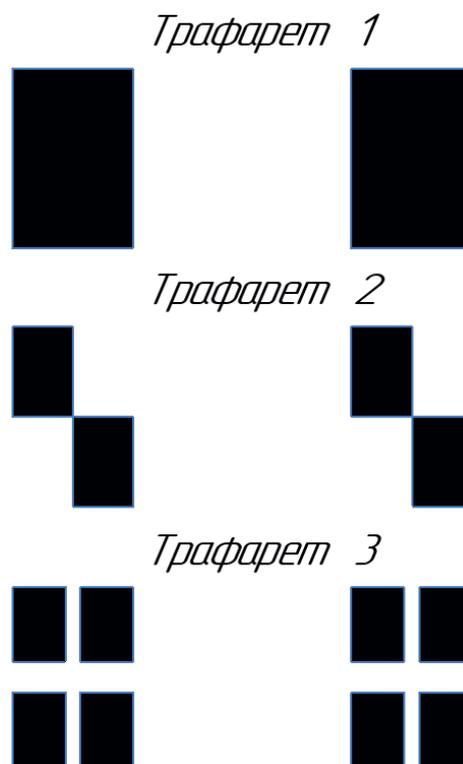


Рис. 2. Тип использованных трафаретов
 Примечание: составлено авторами.

Таблица 2

Параметры используемых паяльных паст

Параметры паст	Паста А	Паста Б
Производитель	АІМ	АІМ
Тип	Бессвинцовая	Бессвинцовая
Обозначение	М8	NC254
Композиция	Sn96.5Ag3Cu0.5 (SAC305)	Sn96.5Ag3Cu0.5 (SAC305)
Содержание флюса (мас.%)	11,50 %	11,50 %
Тип флюса	Натуральная смола – ROL1	Натуральная смола – ROL0
Галогениды	Br: 0,24 %, Cl: 0,0 %	Нет
Температура плавления	230–245 °С	230–245 °С
Вязкость	500–1000 Ксps	300–900 Ксps
Размер частиц	25–45 мкм	25–45 мкм

Примечание: составлено авторами по источнику [15].

обозначена как «паста А», а паяльная паста с флюсом ROL0 – «паста Б».

Печатные платы подвергались пайке в конвекционной печи с непрерывной подачей воздуха. Температурный профиль был одинаковым для обеих паяльных паст и, следовательно, для всех исследуемых образцов. Температурный профиль, измеренный термопарой, прикрепленной к тестируемой печатной плате, показан на рис. 3. Скорость продвижения платы в печи составляет 300 мм/мин. На термопрофиле прослеживаются несколько зон подогрева платы и одна зона оплавления с температурой 230 °С.

Параметры температурного профиля взяты от рекомендаций производителя припоя, чтобы не исказить результаты исследования.

После пайки получены изображения паяных контактов, на которых яркие пятна представляют собой пустоты в соединении (рис. 4). Из-за ограниченного разрешения изображений были обнаружены только пустоты диаметром более 5 мкм. Пустоты обнаружены с ошибкой измерения менее 0,5 %. Минимальный размер макропустоты был установлен на уровне 100 мкм.

Кроме того, с помощью конфокального микроскопа (OPTO EDU, Китай) была измерена шероховатость использованных паяльных масок. По результатам измерений определены три параметра шероховатости: Ra – средняя высота профиля поверхности; RLo – развитая длина профиля шероховатости, выраженная в % рас-

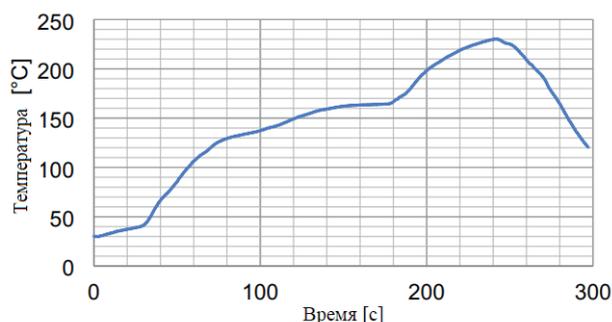


Рис. 3. Температурный режим при оплавлении в печи

Примечание: составлено авторами на основании данных, полученных в исследовании.

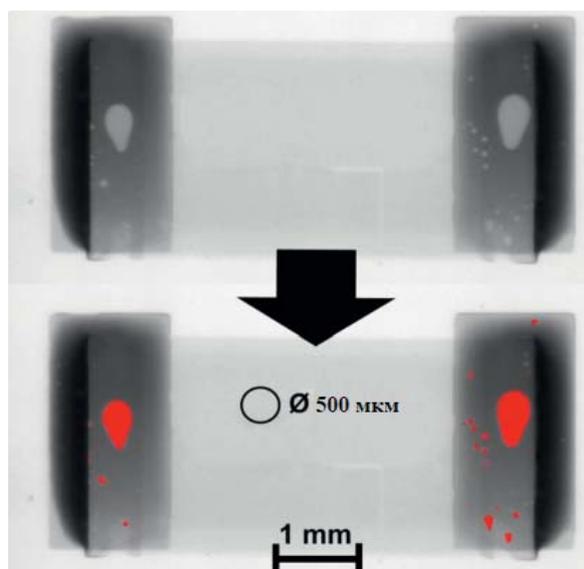


Рис. 4. Исходное рентгеновское изображение и обработанное компьютерное изображение паяных соединений с пустотами, отмеченными красным цветом

Примечание: составлено авторами на основании данных, полученных в исследовании.

ширения над 100 % от гладкого профиля; и R_z – максимальная высота выступов профиля.

В табл. 3 представлены параметры и погрешность измерения шероховатости поверхности испытанных вариантов печатных плат.

Измерение шероховатости используемых паяных масок

Маска для пайки	R_a (мкм)	R_{Lo} (%)	R_z
Без маски	$1,18 \pm 0,03$	141 ± 12	$7,21 \pm 0,05$
Матовая	$1,62 \pm 0,04$	68 ± 7	$1,49 \pm 0,05$
Глянец	$0,18 \pm 0,02$	21 ± 2	$0,73 \pm 0,03$

Примечание: составлено авторами на основании данных, полученных в исследовании.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Для уменьшения количества пустот предложены два новых способа, ведущие к уменьшению образования пустот в паяных соединениях. На рис. 6 представлена оценка площади распространения флюса вокруг паяльных площадок.

В результате эксперимента выявлено, что при первом способе уменьшения пустот (изменение шероховатости поверхности печатной платы) большая площадь распространения припоя приводит к уменьшению уровня пустот внутри соединения (рис. 6).

Рентгеновская съемка ряда макровыступов в паяных соединениях показала соотношения между общей площадью пустот и площадью макровыступов, а также их влияние на общую площадь пустот. Общие результаты средней площади макропористости в паяных соединениях в испытанных образцах в процентах от площади паяного соединения показаны на рис. 7.

Микропустоты составляют лишь минимальную часть пустот, и их количество практически одинаково (в абсолютных цифрах) для всех исследованных образцов. Эти пустоты являются результатом межфазных реакций на границе раздела «подложка/припой». Поэтому они были исключены из итоговой оценки и графиков, так как исследование было сосредоточено в основном на пустотах, связанных с технологическим процессом (макропустотах).

Все данные были статистически обработаны с помощью t -критерия Стьюдента со

Изображение с конфокального микроскопа представлено на рис. 5.

На рис. 5 видна область распространения флюса припоя за пределы паяного соединения и его маркировка.

Таблица 3

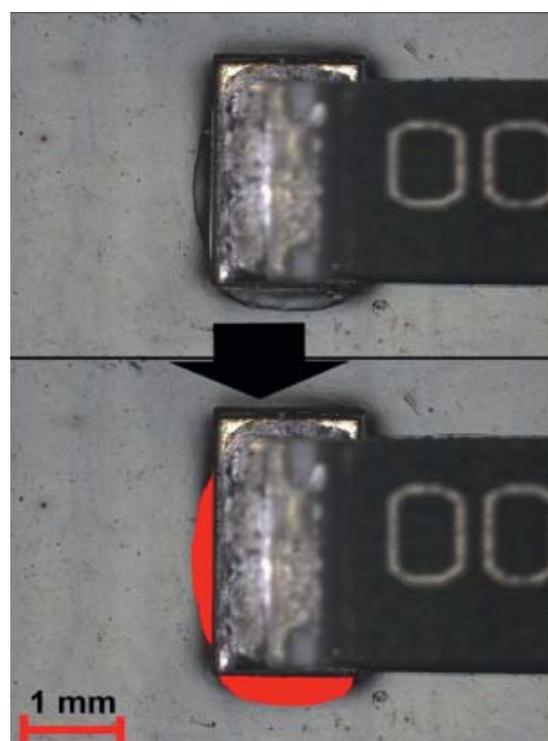


Рис. 5. Изображение с конфокального микроскопа и анализ для определения области растекания припоя

Примечание: составлено авторами на основании данных, полученных в исследовании.

стандартным уровнем значимости (вероятности ошибки) $\alpha = 0,05$. Для выборки использовались результаты оплавления припоя на пяти платах для каждой комбинации типа трафарета, паяльной маски и паяльной пасты.

При рассмотрении влияния типа флюса на образование пустот обнаружено, что образцы, в которых использовалась паяльная паста с ROL1 (паста А), имели меньшую площадь пустот для всех использованных комбинаций

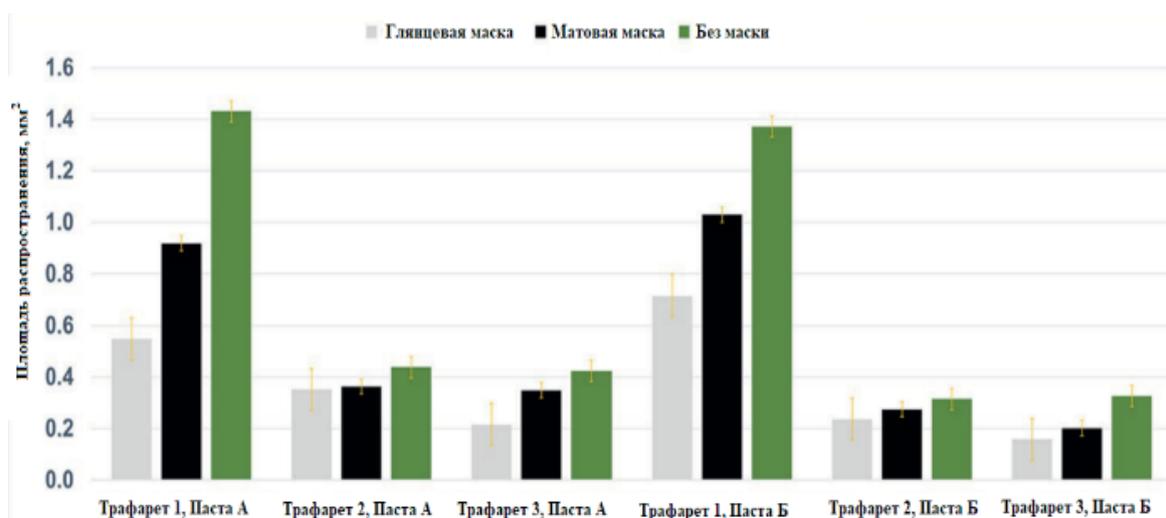


Рис. 6. Область растекания флюса для припоя вокруг контактной площадки
Примечание: составлено авторами на основании данных, полученных в исследовании.

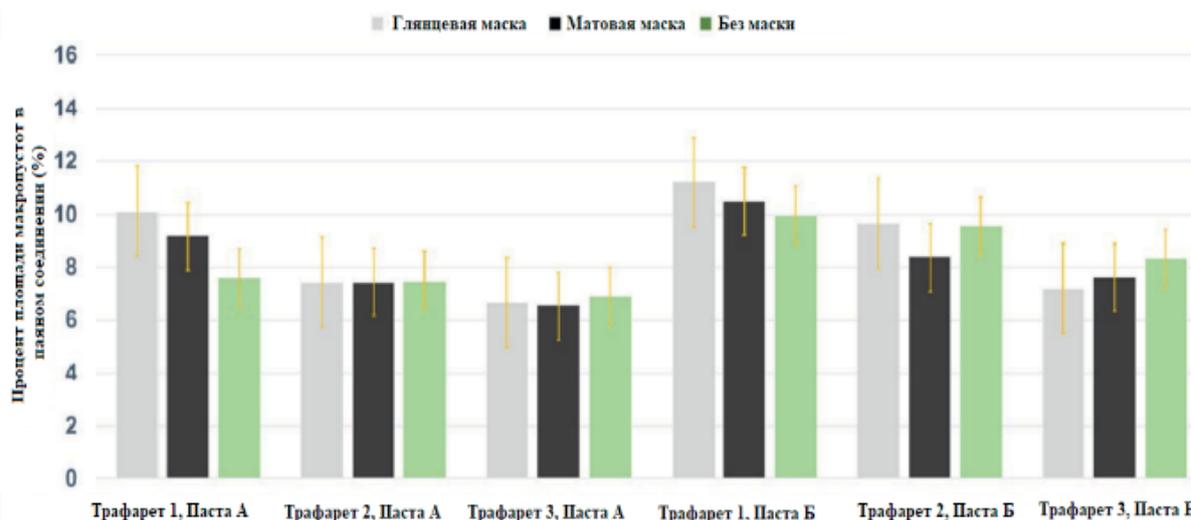


Рис. 7. Общие результаты определения средней площади макропустот в процентах от площади паяного соединения

Примечание: составлено авторами на основании данных, полученных в исследовании.

по сравнению с образцами с паяльной пастой с флюсом ROL0 (паста Б). Содержание пустот в паяных соединениях, изготовленных с использованием пасты А, было на 20 % ниже, чем с пастой Б. Одной из возможных причин меньшей площади пустот в случае ROL1 является его более высокая активность по сравнению с ROL0. Более высокая активность флюса позволяет лучше восстанавливать оксиды с поверхности паяльной площадки. Кроме того, вещества в этом флюсе могут быть более летучими, что может привести к их более быстрому испарению и выходу из объема паяного соединения. Результаты, свидетель-

ствующие о положительном влиянии высокой активности флюса на образование пустот, полностью согласуются с данными, изложенными в источнике [5].

На первый взгляд очевидно, что использование флюса ROL1 (паста А) в целом привело к образованию меньшего количества пустот, но также имело большее растекание вокруг площадки припоя по сравнению с ROL0 (паста Б) (рис. 6, 7). Это означает, что не только активность флюса, но и его способность растекаться на окружающую контактную площадку оказывает влияние на образование пустот.

Согласно t -критерию Стьюдента, разница между пастой А и пастой Б была статистически значимой во всех случаях с вероятностью $p = 0,95$.

Результаты изучения влияния различных конфигураций отверстий трафарета показывают, что образцы, в которых использовались модифицированные конструкции трафаретных отверстий (трафарет 2 и трафарет 3), показывают меньшее образование пустот, чем в случае стандартной конструкции трафаретных отверстий (трафарет 1). В случае трафаретов 2 и 3 паяльная паста (расплавленный сплав, соответственно) имела тенденцию распространяться на первоначально незакрытые участки контактной площадки, что, вероятно, способствовало лучшему газовыделению в процессе пайки. На образцах, подготовленных помощью трафарета 1, растекание расплавленного припоя было практически незначительным, поскольку паяльная паста уже покрыла всю площадь контактной площадки после процесса осаждения паяльной пасты. На рис. 8 представлены различия между образцами, приготовленными с помощью разных трафаретов.

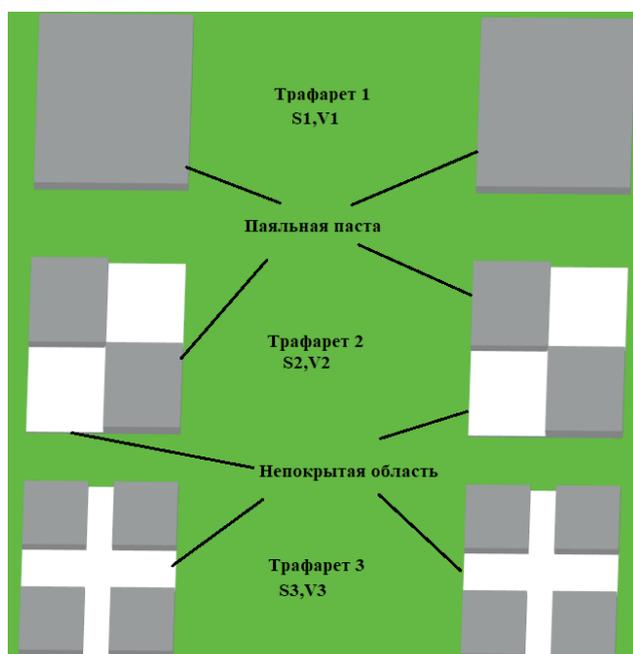


Рис. 8. Одинаковый объем паяльной пасты на контактные площадки с помощью трафарета 1, трафарета 2 и трафарета 3

Примечание: составлено авторами на основании данных, полученных в исследовании.

Очевидно, что объем паяльной пасты сохранился, но площадь покрытия уменьшилась для трафарета 2 и трафарета 3. Влияние использования паяльной пасты Sn96,5Ag3Cu0,5 с различными свойствами от типа точечного флюса и размера порошка на количество и размер пустот внутри паяного соединения исследовали М. Козак и П. Веселый, они подтвердили, что образование несколько более высоких пустот наблюдается в случае паяльной пасты с более мелкими частицами припоя [17]. Однако они не изменяли толщину трафарета, поэтому объем паяльной пасты на площадках уменьшился по сравнению с первоначальной полнотой покрытой областью. Это может привести к механическому ослаблению паяного соединения. Кроме того, исследование проводилось только для соединений, паяемых в вакуумной паяльной печи, в то время как представленный эксперимент моделировал более распространенный на практике процесс пайки.

Другой вопрос, который может возникнуть, касается связи между распространением потока флюса в окружающее пространство и изменением размера отверстий трафарета. Очевидно, что распространение флюса в окружающее пространство контактной площадки больше всего для трафарета 1, хотя объем осажденной паяльной пасты и, следовательно, объем флюса на паяльных площадках был одинаковым для всех случаев.

Чтобы лучше понять это явление, проведен анализ нескольких рельефов паяных соединений, с измерением поперечных размеров профилей паяных соединений на контактных площадках. Измерения проводились на конфокальном микроскопе. На рис. 9 показаны различия в размерном профиле между паяными соединениями, подготовленными с помощью трафарета 1 и трафарета 3.

Для трафарета 1 остатки флюса появляются скорее около контактной площадки, в то время как остатки флюса на образцах, приготовленных по трафарету 3, оставались скорее на верхней части паяного соединения. Такое различие в поведении флюса, вероятно, является

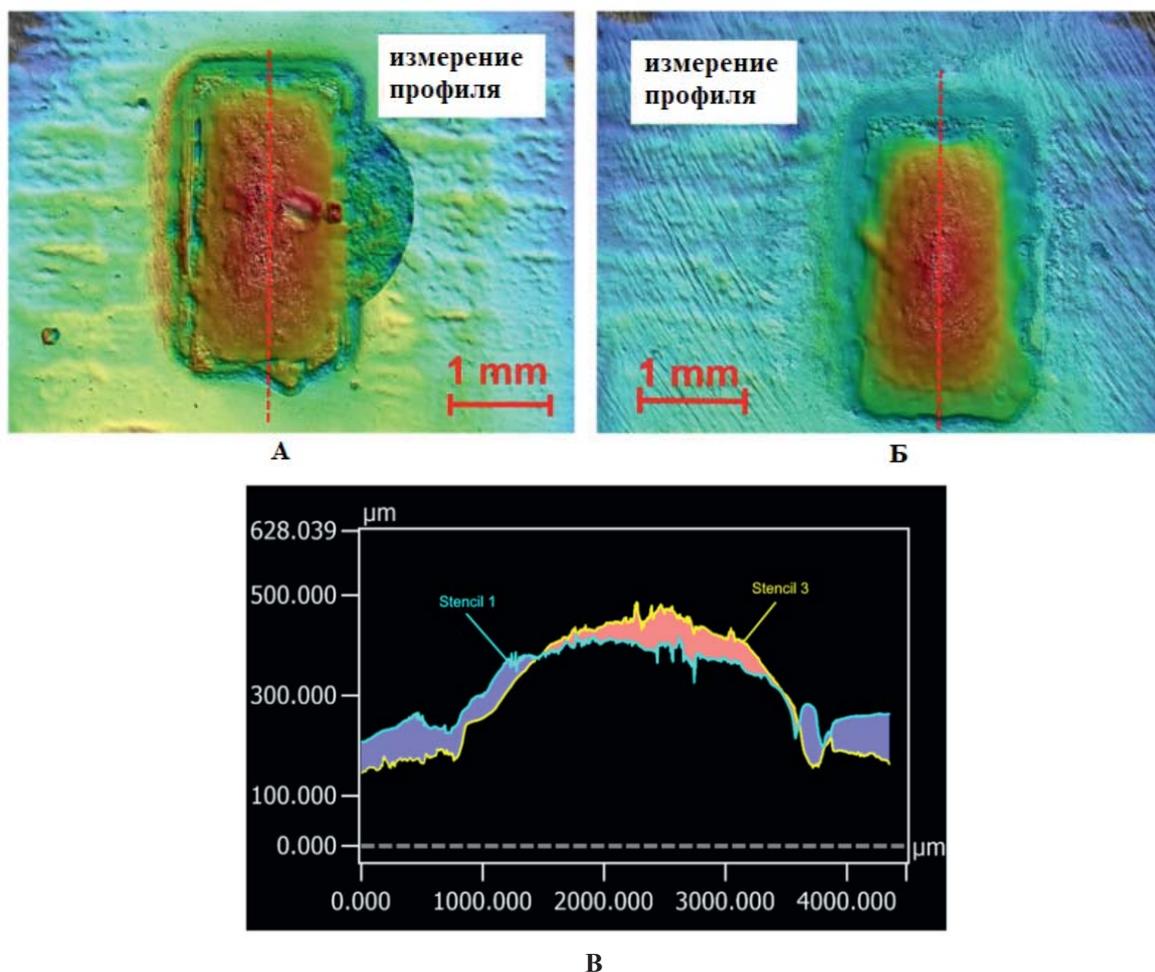


Рис. 9. Рельефы паяного соединения с использованием разных трафаретов:
А – рельеф паяного соединения, подготовленный с использованием трафарета 1 с нанесенной линией измерения профиля размеров; Б – рельеф паяного соединения, подготовленный с использованием трафарета 3 с указанием места измерения профиля размеров; В – полученные размерные профили паяного соединения, подготовленные с помощью трафарета 1 и с помощью трафарета 3
Примечание: составлено авторами на основании данных, полученных в исследовании.

следствием различных процессов смачивания контактных площадок, полностью покрытых паяльной пастой (трафарет 1), и площадок, покрытых лишь частично (трафарет 3). Таким образом, можно утверждать, что основной принцип уменьшения пустот основан в первую очередь на подходе, заключающемся в создании воздушных зазоров и каналов, через которые газовые вещества флюса могут выходить из объема паяного соединения. Кроме того, жидкие вещества вытесняются на верхнюю часть паяного соединения во время процесса пайки.

t-критерий Стьюдента показал, что только разница между трафаретом 1 и трафаретом 2 не была статистически значимой ($p = 0,099$).

ЗАКЛЮЧЕНИЕ

В ходе проведенного научного исследования был произведен анализ влияния различных типов паяльной маски на количество пустот внутри паяных соединений. В итоге было отмечено, что в случае осаждения паяльной пасты через немодифицированный трафарет шероховатость поверхности печатной платы существенно влияет на образование пустот в паяном соединении (рис. 7). Наименьшее количество пустот в паяных соединениях наблюдается в тех случаях, когда печатные платы были без паяльной маски. В этом случае отношение пустот к общей площади паяного соединения составило в среднем 10,2 %, что примерно на 20 % меньше по сравнению с пе-

чатной платой с глянцевой паяльной маской. Подложка без паяльной маски продемонстрировала самую высокую шероховатость поверхности из всех рассмотренных вариантов печатных плат. Более высокая шероховатость поверхности обычно увеличивает смачиваемость жидкости [18], таким образом, площадь распространения флюса была больше (рис. 6), и флюс стекал из объема паяного соединения в большей степени. Различия в выделении флюса также можно обнаружить

между матовыми и глянцевыми масками для припоя, хотя влияние на образование пустот не столь существенно – согласно результатам статистического анализа с использованием t -критерия Стьюдента, значение p было равно 0,08.

В случае применения трафарета 2 и трафарета 3 влияние паяльной маски не столь очевидно, как в случае с трафаретом 1. Причина заключается в гораздо меньшем выделении флюса в окрестности паяльной площадки.

Список источников

1. Steiner F., Wirth V., Hirman M. Relationship of soldering profile, voids formation and strength of soldered joints // 42nd International Spring Seminar on Electronics Technology (ISSE), Wrocław, 2019. P. 1–6. <https://doi.org/10.1109/ISSE.2019.8810303>.
2. Zhu N. Thermal impact of solder voids in the electronic packaging of power devices // Fifteenth Annual IEEE Semiconductor Thermal Measurement and Management Symposium, San Diego, CA, USA, 1999. P. 22–29. <https://doi.org/10.1109/STHERM.1999.762424>.
3. Chen L., Paulasto-Krockel M., Frohler U. et al. Thermal impact of randomly distributed solder voids on Rth-JC of MOSFETs // 2nd Electronics System-Integration Technology Conference, 2008. P. 237–244. <https://doi.org/10.1109/ESTC.2008.4684356>.
4. Zhou Y., Ding D., Han B. et al. Influence of reflow atmosphere on SAC305 solder joints // International Symposium on Advanced Packaging Materials (APM), 2011. P. 122–128.
5. Bušek D., Dušek K., Růžička D. et al. Flux effect on void quantity and size in soldered joints // Microelectronics Reliability, 2016. Vol. 60. P. 135–140. <https://doi.org/10.1016/j.microrel.2016.03.009>.
6. Ekere N. N., Lo E. K. New challenges in solder-paste printing // Journal of Electronics Manufacturing, 1991. Vol. 01. P. 29–40. <https://doi.org/10.1142/S0960313191000059>.
7. He D., Ekere N. N., Currie M. A. The behavior of solder pastes in stencil printing with vibrating squeegee // IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part C, 1998. Vol. 21, no. 4. P. 317–324. <https://doi.org/10.1109/TCPMC.1998.7102530>.
8. Hong L.-Y., Li Y.-T., Li H.-F. Minimize Bottom Termination Component Voids by Board Assembly Process & Design Optimization // 17th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), Taipei, Taiwan, 2022. P. 1–6. <https://doi.org/10.1109/IMPACT56280.2022.9966717>.
9. Sweatman K., Nishimura T., Sugimoto K. et al. Controlling Voiding Mechanisms in the Reflow Soldering Process // Proceedings IPC APEX Expo, 2016. P. 1–11.

References

1. Steiner F., Wirth V., Hirman M. Relationship of soldering profile, voids formation and strength of soldered joints. In: *42nd International Spring Seminar on Electronics Technology (ISSE)*, Wrocław. 2019:1–6. <https://doi.org/10.1109/ISSE.2019.8810303>.
2. Zhu N. Thermal impact of solder voids in the electronic packaging of power devices. In: *Fifteenth Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, San Diego, CA, USA. 1999:22–29. <https://doi.org/10.1109/STHERM.1999.762424>.
3. Chen L., Paulasto-Krockel M., Frohler U. et al. Thermal impact of randomly distributed solder voids on Rth-JC of MOSFETs. In: *2nd Electronics System-Integration Technology Conference*. 2008:237–244. <https://doi.org/10.1109/ESTC.2008.4684356>.
4. Zhou Y., Ding D., Han B. et al. Influence of reflow atmosphere on SAC305 solder joints. In: *International Symposium on Advanced Packaging Materials (APM)*. 2011:122–128.
5. Bušek D., Dušek K., Růžička D. et al. Flux effect on void quantity and size in soldered joints. *Microelectronics Reliability*. 2016;60:135–140. <https://doi.org/10.1016/j.microrel.2016.03.009>.
6. Ekere N. N., Lo E. K. New challenges in solder-paste printing. *Journal of Electronics Manufacturing*. 1991;01:29–40. <https://doi.org/10.1142/S0960313191000059>.
7. He D., Ekere N. N., Currie M. A. The behavior of solder pastes in stencil printing with vibrating squeegee. *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part C*. 1998;21(4):317–324. <https://doi.org/10.1109/TCPMC.1998.7102530>.
8. Hong L.-Y., Li Y.-T., Li H.-F. Minimize Bottom Termination Component Voids by Board Assembly Process & Design Optimization. In: *17th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT)*, Taipei, Taiwan. 2022:1–6. <https://doi.org/10.1109/IMPACT56280.2022.9966717>.
9. Sweatman K., Nishimura T., Sugimoto K. et al. Controlling Voiding Mechanisms in the Reflow Soldering Process. In: *Proceedings IPC APEX Expo*. 2016:1–11.

10. Otáhal A., Somer J., Szendiuch I. Influence of heating direction on BGA solder balls structure // *Microelectronics and Packaging Conference (EMPC) & Exhibition*, Warsaw. 2017. P. 1–4. <https://doi.org/10.23919/EMPC.2017.8346878>.
11. Otáhal A., Somer J., Szendiuch I. Influence of heat flow direction on solder ball interfacial layer // *Journal of Electrical Engineering*. 2018. Vol. 69, no. 4. P. 305–310. <https://doi.org/10.2478/jee-2018-0043>.
12. Wang D., Panton R. L. Effect of Reversing Heat Flux Direction During Reflow on Void Formation in High-Lead Solder Bumps // *Journal of Electronic Packaging*. 2005. Vol. 127, no. 4. P. 440–445. <https://doi.org/10.1115/1.2070047>.
13. Yunus M., Primavera A., Srihari K. et al. Effect of voids on the reliability of BGA/CSP solder joints // *Twenty Sixth IEEE/CPMT International Electronics Manufacturing Technology Symposium*, Santa Clara, CA, USA, 2000. P. 207–213. <https://doi.org/10.1109/IEMT.2000.910730>.
14. Nurmi S. T., Sundelin J. J., Ristolainen E. O. et al. The effect of PCB surface finish on lead-free solder joints // *Soldering & Surface Mount Technology*. 2005. Vol. 17, no. 1. P. 13–23. <https://doi.org/10.1108/09540910510579203>.
15. ГОСТ Р МЭК 61190-1-1-2020. Материалы для электронных модулей. Требования к паяльным флюсам для высококачественных межсоединений в электронных сборках. Ч. 1–1. М., 2020. 16 с. URL: <https://www.elec.ru/viewer?url=/files/2021/08/09/GOST-R-MEK-61190-1-1-2020.pdf> (дата обращения: 18.09.2024).
16. Veselý P., Bušek D., Krammer O. et al. Analysis of noclean flux spatter during the soldering process // *Journal of Materials Processing Technology*, 2020. Vol. 275. <https://doi.org/10.1016/j.jmatprotec.2019.116289>.
17. Kozák M., Veselý P. Study of voids inside solder joints based on SAC305 solder paste with different properties // *43rd International Spring Seminar on Electronics Technology (ISSE)*, 2020. P. 1–5. <https://doi.org/10.1109/ISSE49702.2020.9120951>.
18. Hirman M., Steiner F. Optimization of solder paste quantity considering the properties of solder joints // *Soldering & Surface Mount Technology*. 2017. Vol. 29, no. 1. P. 15–22. <https://doi.org/10.1108/SSMT-10-2016-0025>.
10. Otáhal A., Somer J., Szendiuch I. Influence of heating direction on BGA solder balls structure. In: *Microelectronics and Packaging Conference (EMPC) & Exhibition*, Warsaw. 2017:1–4. <https://doi.org/10.23919/EMPC.2017.8346878>.
11. Otáhal A., Somer J., Szendiuch I. Influence of heat flow direction on solder ball interfacial layer. *Journal of Electrical Engineering*. 2018;69(4):305–310. <https://doi.org/10.2478/jee-2018-0043>.
12. Wang D., Panton R. L. Effect of Reversing Heat Flux Direction During Reflow on Void Formation in High-Lead Solder Bumps. *Journal of Electronic Packaging*. 2005;127(4):440–445. <https://doi.org/10.1115/1.2070047>.
13. Yunus M., Primavera A., Srihari K. et al. Effect of voids on the reliability of BGA/CSP solder joints. In: *Twenty Sixth IEEE/CPMT International Electronics Manufacturing Technology Symposium*, Santa Clara, CA, USA. 2000:207–213. <https://doi.org/10.1109/IEMT.2000.910730>.
14. Nurmi S. T., Sundelin J. J., Ristolainen E. O. et al. The effect of PCB surface finish on lead-free solder joints. *Soldering & Surface Mount Technology*. 2005;17(1):13–23. <https://doi.org/10.1108/09540910510579203>.
15. GOST R MEK 61190-1-1-2020. Materialy dlya elektronnykh moduley. Trebovaniya k payalnym flyusam dlya vysokokachestvennykh mezhsoedineniy v elektronnykh sborkakh. Pt. 1–1. Moscow. 2020. 16 p. URL: <https://www.elec.ru/viewer?url=/files/2021/08/09/GOST-R-MEK-61190-1-1-2020.pdf> (accessed: 18.09.2024). (In Russ.).
16. Veselý P., Bušek D., Krammer O. et al. Analysis of noclean flux spatter during the soldering process. *Journal of Materials Processing Technology*. 2020;275. <https://doi.org/10.1016/j.jmatprotec.2019.116289>.
17. Kozák M., Veselý P. Study of voids inside solder joints based on SAC305 solder paste with different properties. In: *43rd International Spring Seminar on Electronics Technology (ISSE)*. 2020:1–5. <https://doi.org/10.1109/ISSE49702.2020.9120951>.
18. Hirman M., Steiner F. Optimization of solder paste quantity considering the properties of solder joints. *Soldering & Surface Mount Technology*. 2017;29(1):15–22. <https://doi.org/10.1108/SSMT-10-2016-0025>.

Информация об авторах

А. В. Турецкий – кандидат технических наук, доцент.

А. А. Пирогов – кандидат технических наук, доцент.

И. В. Свиридова – старший преподаватель.

М. В. Хорошайлова – кандидат технических наук, доцент.

Э. В. Сёмка – кандидат физико-математических наук, доцент.

About the authors

A. V. Turetsky – Candidate of Sciences (Engineering), Docent.

A. A. Pirogov – Candidate of Sciences (Engineering), Docent.

I. V. Sviridova – Senior Lecturer.

M. V. Khoroshailova – Candidate of Sciences (Engineering), Docent.

E. V. Syomka – Candidate of Sciences (Physics and Mathematics), Docent.

Научная статья
УДК 517.443:338.5
<https://doi.org/10.35266/1999-7604-2024-3-9>



Применение метода вейвлет-анализа данных для построения прогнозной модели

Олеся Викторовна Чубарова¹, Анатолий Викторович Чубаров²,
Дарья Игоревна Ликсонова³✉

¹Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева, Красноярск, Россия

^{2,3}Сибирский федеральный университет, Красноярск, Россия

¹kuznetcova_o@mail.ru, <https://orcid.org/0000-0002-9331-6682>

²chubarovav@hotmail.ru, <https://orcid.org/0009-0005-1317-9649>

³LiksonovaDI@yandex.ru ✉, <https://orcid.org/0000-0001-9663-6481>

Аннотация. Работа посвящена разработке модели прогнозирования цены на любой продукт на основе исследования входных данных с помощью метода вейвлет-анализа. Данный метод является удобным инструментом для обработки и преобразования различных типов исходных данных. Каждый элемент рыночного механизма связан с ценой, которая является основным критерием, воздействующим на спрос и предложение. Процессы планирования, стимулирования и управления экономической деятельностью на всех стадиях воспроизводства связаны с ценообразованием, и предвидение последствий принимаемых решений является достаточно важной задачей. В ходе проведенных экспериментов на данных из открытых источников были получены удовлетворительные результаты; установлено, что строить прогноз возможно не более чем на три последующих дня.

Ключевые слова: временные ряды, вейвлет-анализ, математическое моделирование, прогнозирование, коэффициенты разложения

Для цитирования: Чубарова О. В., Чубаров А. В., Ликсонова Д. И. Применение метода вейвлет-анализа данных для построения прогнозной модели // Вестник кибернетики. 2024. Т. 23, № 3. С. 82–89. <https://doi.org/10.35266/1999-7604-2024-3-9>.

Original article

Application of wavelet data analysis for forecast model development

Olesya V. Chubarova¹, Anatoly V. Chubarov², Darya I. Liksonova³✉

¹Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia

^{2,3}Siberian Federal University, Krasnoyarsk, Russia

¹kuznetcova_o@mail.ru, <https://orcid.org/0000-0002-9331-6682>

²chubarovav@hotmail.ru, <https://orcid.org/0009-0005-1317-9649>

³liksonovadi@yandex.ru ✉, <https://orcid.org/0000-0001-9663-6481>

Abstract. The work is devoted to the development of a model for forecasting the price of any product based on the study of input data using the wavelet analysis. This method is a convenient tool for processing and transforming various types of initial data. Each element of the market mechanism is associated with a price, which is the key criterion affecting supply and demand. The processes of planning, stimulating and managing economic activity at all stages of reproduction are associated with pricing. Forecasting the consequences of decisions taken is a very important task. In the course of experiments on data from open sources, satisfactory results were obtained. It was established that it is possible to build a forecast for no more than three subsequent days.

Keywords: time series, wavelet analysis, math modeling, forecasting, expansion coefficients

For citation: Chubarova O. V., Chubarov A. V., Liksonova D. I. Application of wavelet data analysis for forecast model development. *Proceedings in Cybernetics*. 2024;23(3):82–89. <https://doi.org/10.35266/1999-7604-2024-3-9>.

ВВЕДЕНИЕ

В настоящее время используется большое количество традиционных методов анализа для прогнозирования цен. К ним можно отнести методы анализа данных и машинного обучения (например, алгоритм градиентного бустинга, факторная регрессионная модель, метод корреляционно-регрессионного анализа и др.). Безусловно, такие методы имеют преимущества и недостатки. При рассмотрении процесса изменения цен ценовой ряд можно отнести к временному ряду. Одним из основных методов обработки такого ряда является спектральный анализ (Фурье-анализ), который позволяет охарактеризовать частотный состав исследуемого сигнала [1, 2]. Недостатком Фурье-анализа является то, что частотные компоненты несут глобальный характер в силу существования в пределах от $-\infty$ до $+\infty$, это условие естественным образом ограничивает его использование на практике. Существует форма кратковременного преобразования Фурье. Однако и у этой формы существует недостаток, происходящий от небезызвестного принципа Гейзенберга. Поэтому решением задачи исследования временных рядов могут стать преобразования с переменным разрешением, такие как вейвлет-преобразования [3]. Таким образом, основной задачей данной работы является моделирование временных рядов с применением математического аппарата вейвлет-анализа.

МАТЕРИАЛЫ И МЕТОДЫ

Сравнительно молодой термин «вейвлет» введен в середине 1980-х гг. в статье А. Гроссмана и Ж. Морле, посвященной анализу свойств сейсмических и акустических сигналов [3]. Сейчас это известно как «вейвлет-преобразование», которое переводит сигнал из временного представления в частотно-временное. При решении практических задач анализ различных зависимостей на основе вейвлетов получил широкое применение [4].

Анализ одномерного сигнала с помощью вейвлет-преобразования состоит в его разложении по базису. Базисные функции должны обладать определенными свойствами солитоноподобной функции (вейвлета) за счет масштабных изменений и переносов. Каждая из функций этого базиса характеризует как определенную пространственную (временную) частоту, так и ее локализацию в физическом пространстве (времени). Преобразование выполняется отдельно для различных сегментов сигнала во временной области. Базисные функции определены компактно, это позволяет говорить о финитности получаемых последовательностей.

В случае когда сигнал представляет собой нестационарный случайный процесс, для задач анализа и синтеза, оптимизации параметров моделей сигналов и выделения сигналов на фоне шумов различного происхождения целесообразно применить вейвлет-преобразование. Благодаря высокому разрешению во временной области возможно выделять и анализировать импульсные составляющие, проектируя на основе полученных данных различные графические модели, удобные для восприятия [5].

Для временного ряда с помощью методов вейвлет-анализа необходимо получить количественные и качественные оценки, характеризующие нестационарность (хаотичность) исследуемого процесса, затем построить авторегрессионную математическую модель типа ARMA [6] для коэффициентов разложения, сделать их прогноз, после чего выполнить вейвлет-реконструкцию для предсказания исходного сигнала.

В первую очередь необходимо выполнить вейвлет-разложение исходного временного ряда. В результате этого на каждом уровне получаемый на предыдущем шаге низкочастотный сигнал разделяется на низкочастотную и высокочастотную части.

Исходный сигнал (обозначим его как x) после разложения на N слоев можно представить следующим образом:

$$x = D_1 + D_2 + \dots + D_N + A_N, \quad (1)$$

где D_1, D_2, \dots, D_N – это первый, второй и другие N слоев соответственно (высокочастотные сигналы);

A_N – это низкочастотный сигнал, полученный путем разложения N -го слоя.

Таким образом, необходимо получить прогнозные значения $D1, D2, \dots, D_N$ и A_N , после чего выполнить вейвлет-реконструкцию для предсказания исходного сигнала [7].

Представим блок-схему работы алгоритма прогнозирования (рис. 1).

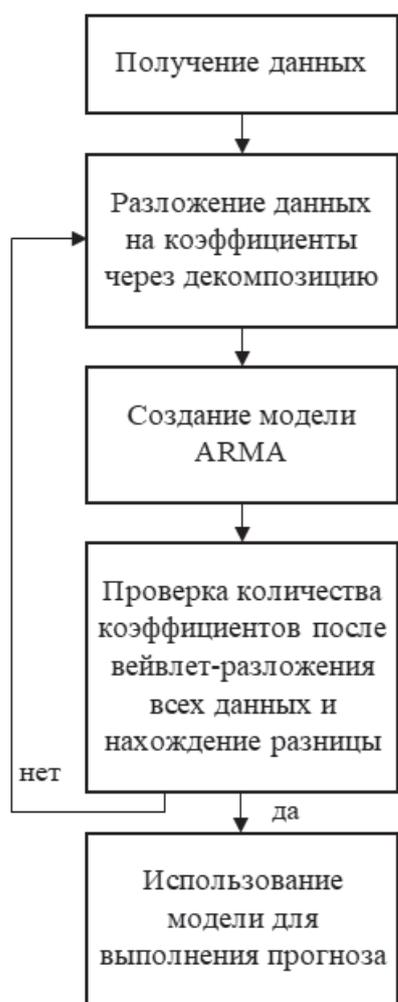


Рис. 1. Блок-схема работы алгоритма прогнозирования

Примечание: составлено авторами на основании разработанного алгоритма.

Рассмотрим последовательность шагов для разложения по методу вейвлет-анализа.

1. Необходимо разложить выбранные данные с использованием вейвлет-анализа. Выбираем в качестве вейвлет-функции вейвлеты Добеши с $n = 4$, которые являются достаточно известными типами и используются во многих практических приложениях. Вейвлеты порядка N (dbN) отличны от нуля лишь на интервале длиной $(2N - 1)$ и имеют $2N$ отличных от нуля коэффициентов фильтров [8]. Количество уровней разложения – 2.

2. Создаем и восстанавливаем модель ARMA для каждого коэффициента слоя.

3. По построенным моделям определяем вейвлет-коэффициенты, для чего определяем количество шагов, которое следует предсказать каждой модели ARMA. Для этого проверяем количество коэффициентов после вейвлет-разложения всех исходных данных и находим разницу.

4. Для предсказания исходного сигнала производим вейвлет-реконструкцию.

В результате выполнения данной последовательности шагов получаем модель, которая позволит построить необходимый прогноз временного ряда.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В качестве исходной информации для исследования были взяты данные из открытых источников, в частности цена закрытия акций на торгах некоторых компаний. Были сформированы обучающая и тестовая выборки.

Из исходной выборки выделим значения «цена закрытия» и разделим их на тикеры. Для каждого тикера выделяем обучающую выборку и 5 точек для построения прогноза. На обучающей выборке применяем вейвлет-преобразование с уровнем глубины 2, в результате чего получаем коэффициенты разложения $D1, D2, A_2$, которые затем используем для настройки авторегрессионной модели ARMA.

Покажем пример полученной модели по обучающей выборке для тикера «CLL» (рис. 2). Значение по оси абсцисс – номер точ-

ки (соответствует временному промежутку равному суткам), по оси ординат – значение цены закрытия.

Построим прогноз цены закрытия на 5 дней вперед по полученной модели. Результаты прогнозирования и реальные значения цены представлены в табл. 1: показаны реальные значения из теста (real), предсказанные (predict), сигнал рассогласования (err) и относительная ошибка в процентах (err_rate).

Приведем на рис. 3 предсказание на тестовой выборке (укрупненно для 5 точек) по модели ARMA, предсказанные точки обозначены кружочками, реальные – звездочками.

Было проведено моделирование на разных тиккерах и для разных периодов прогнозирования. Полученные результаты позволяют говорить об эффекте прогнозирования модели на небольшие промежутки времени, то есть

чем длительней период прогнозирования, тем больше растет ошибка. В результате было принято решение не строить прогноз на период более чем 3 дня.

Покажем на примере тиккера «A-share», как выглядят модели ARMA для глубины равной 2 (рис. 4–6). На рисунках пунктирной линией представлены истинные данные, сплошной – значения модели.

Анализируя модели вейвлет-коэффициентов, можно понять необходимость фильтрации исходной последовательности. Цель состоит в том, чтобы подавить шумовую часть сигнала и восстановить основной тренд.

При вейвлет-анализе сигнал разделяется на сглаженный (аппроксимирующие коэффициенты) и шумовой (детализирующие коэффициенты). К шумовой части целесообразно применить вейвлет-фильтрацию. Поскольку шумовая компонента больше отражается в детализирующих коэффициентах D_j , обра-

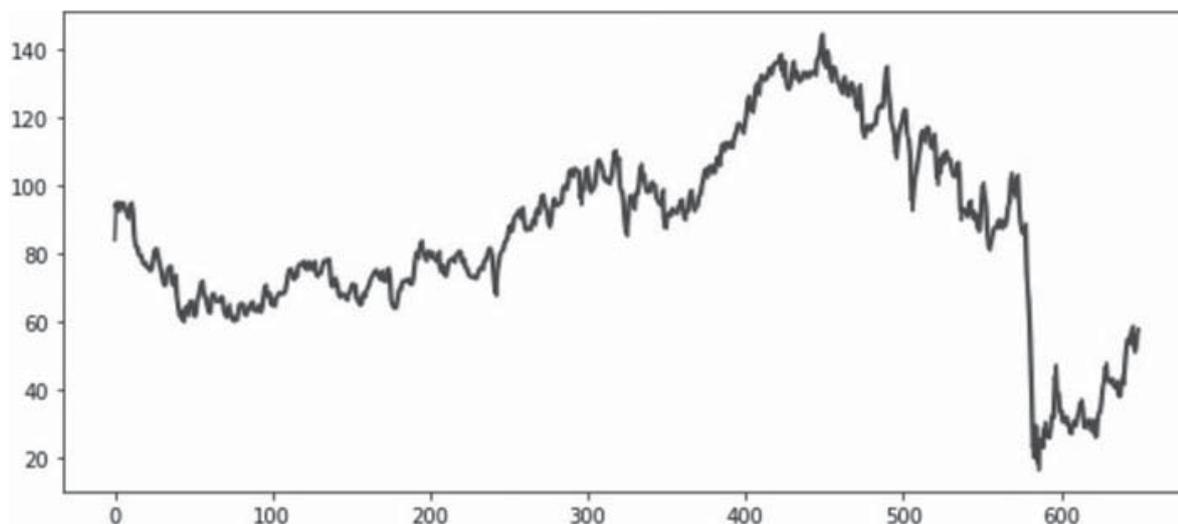


Рис. 2. Модель прогнозирования тиккера «CLL» с использованием вейвлет-анализа

Примечание: составлено авторами на основании данных, полученных в исследовании.

Таблица 1

Результаты прогнозирования

№	real	predict	err	err_rate
0	28,56	28,80	0,24	0,86
1	29,30	28,75	-0,54	1,86
2	27,75	29,06	1,32	4,75
3	27,82	29,17	1,35	4,87
4	27,94	29,20	1,26	4,50

Примечание: составлено по модели ARMA.

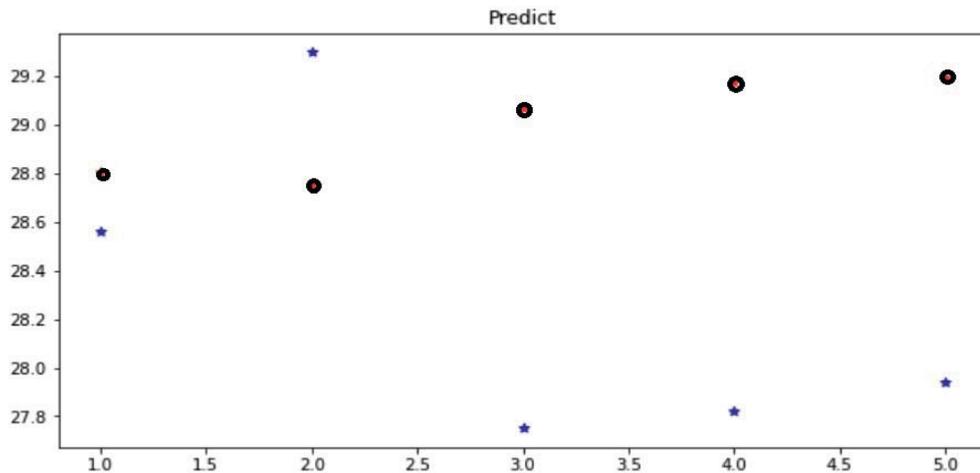


Рис. 3. Предсказание на тестовой выборке по модели ARMA

Примечание: составлено авторами на основании данных, полученных в исследовании.

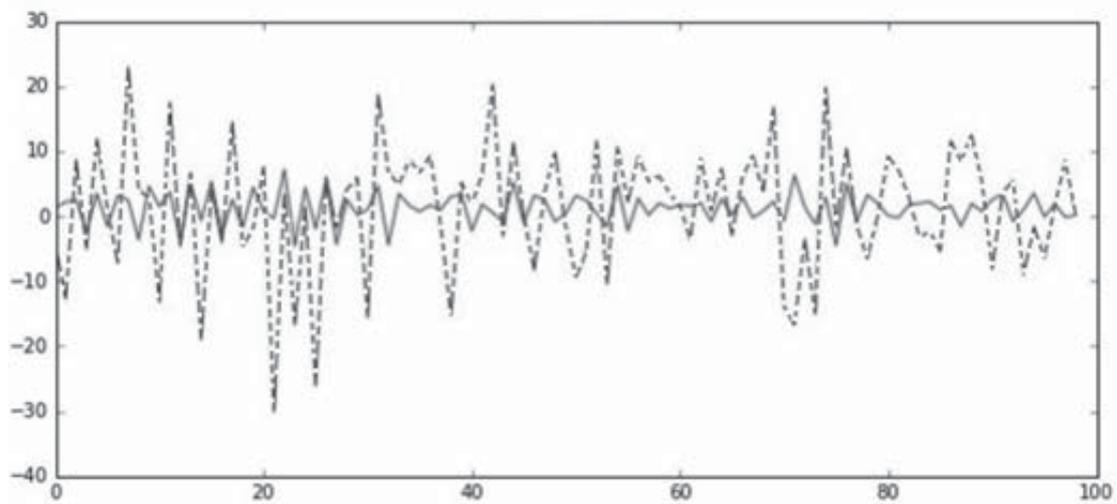


Рис. 4. Модель ARMA для слоя коэффициента D_1

Примечание: составлено авторами на основании данных, полученных в исследовании.

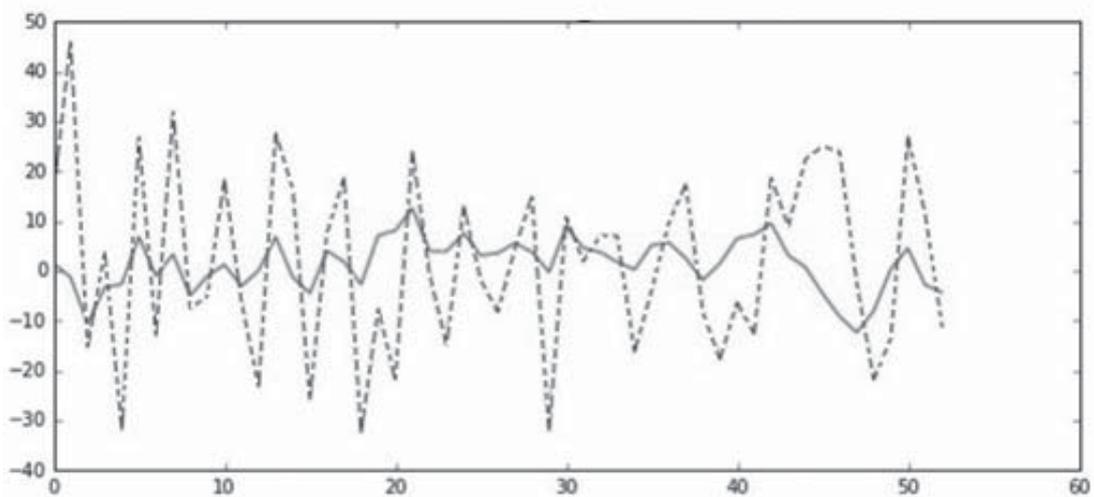


Рис. 5. Модель ARMA для слоя коэффициента D_2

Примечание: составлено авторами на основании данных, полученных в исследовании.

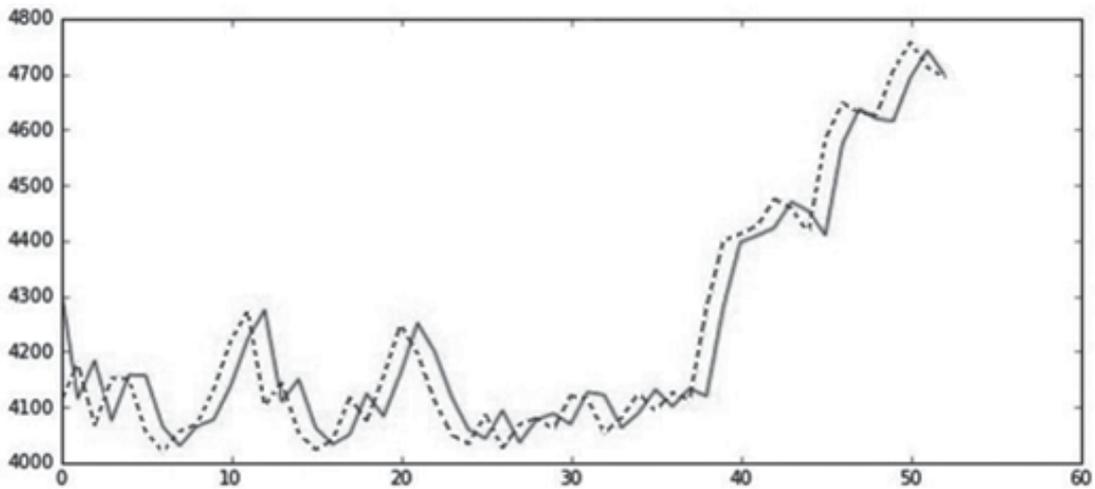


Рис. 6. Модель ARMA для слоя коэффициента A ,
Примечание: составлено авторами на основании данных, полученных в исследовании.

батывают обычно их. Самый простой способ удаления шума – придать нулевые значения коэффициентам, меньшим некоторой пороговой величины. Эта процедура называется пороговой обработкой или трешхолдингом коэффициентов [8].

Полученные детализирующие вейвлет-коэффициенты $D_1, D_2, \dots, D_{N-2}, D_{N-1}$ обнуляют все положительные (отрицательные) значения, имеющие амплитуду, меньшую (большую), чем некоторые положительные (отрицательные) пороги, устанавливаемые симметрично относительно нулевого уровня [8]. Для задачи фильтрации лучшим вариантом будет ситуация, когда остальные элементы векторов остаются без изменения, то есть реализуется так называемый жесткий порог. При решении практических задач также можно столкнуться с использованием мягкого порога, когда элементы последовательностей, находящиеся вне ограниченной порогом полосы, сдвигаются по направлению к нулевому уровню на величину высоты порога.

Продемонстрируем результат применения фильтрации исходной последовательности и покажем результат прогнозирования на 3 дня. На рис. 7 представлена полученная модель для тиккера «ХОМ».

Результаты прогнозирования и реальные значения цены для выбранных данных представлены в табл. 2.

Приведем на рис. 8 предсказание на тестовой выборке (укрупненно для 3 точек).

Видно, что применение трешхолдинга позволило уменьшить ошибку прогнозирования. Высокой точности прогнозирования временного ряда, безусловно, не ожидается, но важным является отслеживание тенденции роста или падения цены закрытия, что вполне отслеживается в прогнозе.

ЗАКЛЮЧЕНИЕ

Проведя исследования данного подхода на разных тиккерах за разный период, времени можно сделать следующие выводы:

- использование вейвлет-анализа для временных рядов является эффективным инструментом;
- модель прогнозирования имеет более высокую точность прогнозов на небольших временных промежутках;
- при вейвлет-реконструкции для предсказания исходного сигнала целесообразно использовать фильтрацию.

В перспективе для моделирования вейвлет-коэффициентов каждого слоя после вейвлет-разложения можно попробовать иные методы, например, нейронные сети.

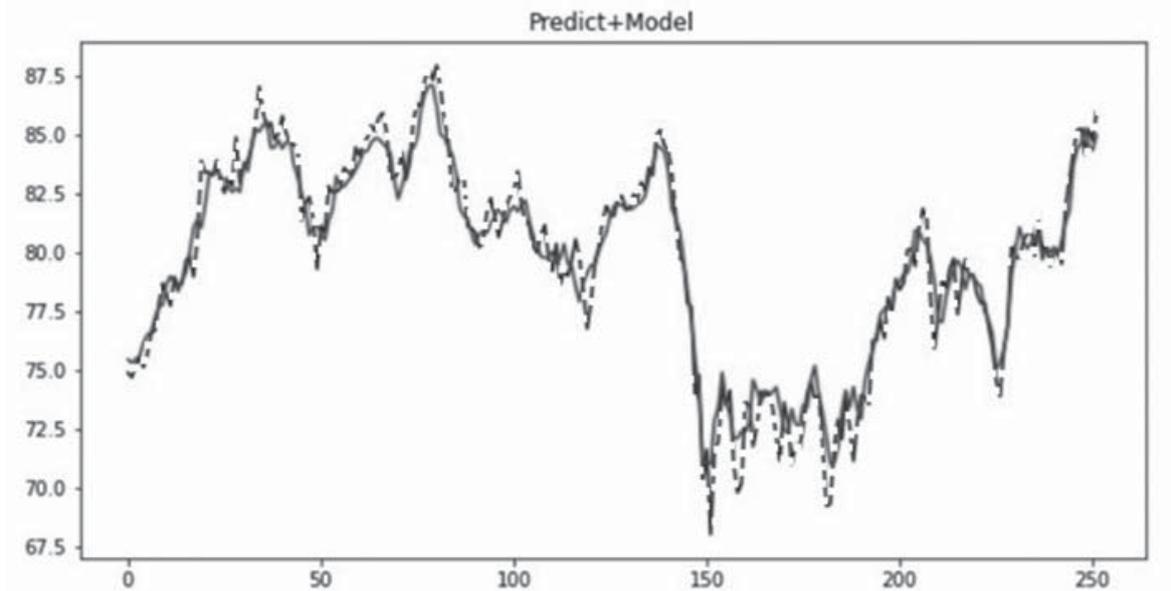


Рис. 7. Модель с фильтрацией для тиккера «ХОМ»

Примечание: составлено авторами на основании данных, полученных в исследовании.

Таблица 2

Результаты прогнозирования для тиккера «ХОМ»

№	real	predict	err	err_rate
0	87,05	87,53	0,47	0,55
1	86,10	87,75	1,66	1,92
2	85,04	87,23	2,18	2,57

Примечание: составлено по модели ARMA.

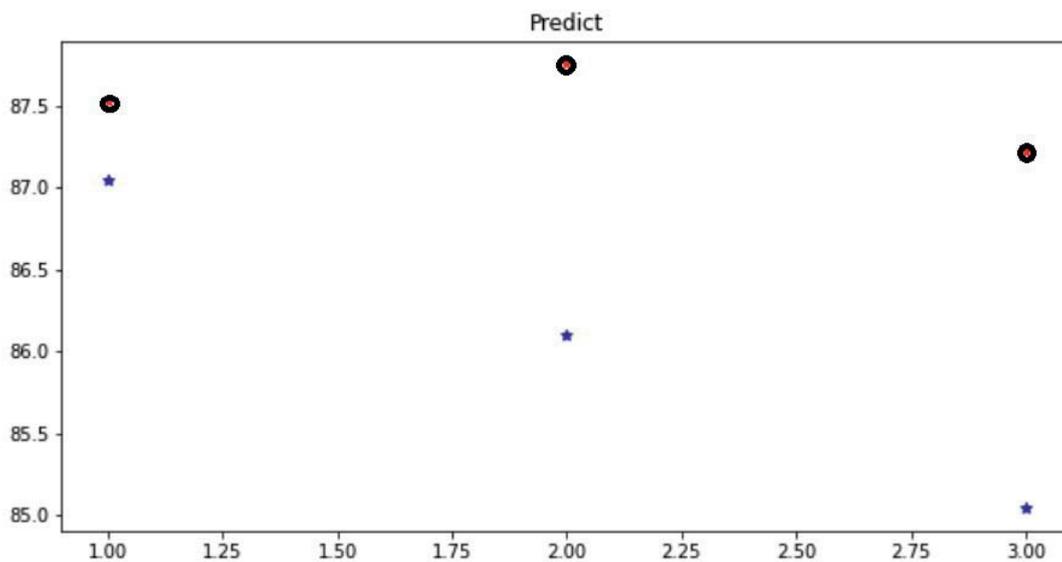


Рис. 8. Предсказание на тестовой выборке для тиккера «ХОМ»

Примечание: составлено авторами на основании данных, полученных в исследовании.

Список источников

1. Латыпова Н. В., Тучинский Л. И. Ряды Фурье. Ижевск : Удмуртский университет, 2011. 80 с.
2. Жук В. В., Натансон Г. И. Тригонометрические ряды Фурье и элементы теории аппроксимации. Л. : Изд-во ЛГУ, 1983. 188 с.
3. Grossman A., Morlet J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape // *SIAM Journal on Mathematical Analysis*. 1984. Vol. 15, no. 4. P. 723–736.
4. Нагорнов О. В., Никитаев В. Г., Простоквашин В. М. и др. Вейвлет-анализ в примерах. М. : НИЯУ МИФИ, 2010. 120 с.
5. Новиков И. Я., Протасов В. Ю., Скопина М. А. Теория всплесков : моногр. М. : Физматлит, 2005. 615 с.
6. Афанасьев В. Н., Юзбашев М. М. Анализ временных рядов и прогнозирование. М. : Финансы и статистика, 2001. 228 с.
7. Анализ и модели временных рядов. URL: <https://www.statmethods.ru/statistics-metody/modeli-vremennykh-ryadov/> (дата обращения: 15.05.2024).
8. Петров Г. А., Шуранов Е. В. Практика использования вейвлет-анализа в дефектоскопии. СПб. : Балт. гос. техн. ун-т, 2012. 64 с.

Информация об авторах

О. В. Чубарова – кандидат технических наук, доцент.

А. В. Чубаров – кандидат технических наук.

Д. И. Ликсонова – кандидат технических наук.

References

1. Latypova N. V., Tuchinskiy L. I. Ryady Fure. Izhevsk: Udmurtskiy universitet, 2011. 80 p. (In Russ.).
2. Zhuk V. V., Natanson G. I. Trigonometricheskie ryady Fure i elementy teorii approksimatsii. Leningrad: Izd-vo LGU, 1983. 188 p. (In Russ.).
3. Grossman A., Morlet J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM Journal on Mathematical Analysis*. 1984;15(4):723–736.
4. Nagornov O. V., Nikitaev V. G., Prostokvashin V. M. et al. Veyvlet-analiz v primerakh. Moscow: NIYAU MIFI, 2010. 120 p. (In Russ.).
5. Novikov I. Ya., Protasov V. Yu., Skopina M. A. Teoriya vspleskov: Monograph. Moscow: Fizmatlit, 2005. 615 p. (In Russ.).
6. Afanasev V. N., Yuzbashev M. M. Analiz vremennykh ryadov i prognozirovanie. Moscow: Finansy i statistika, 2001. 228 p. (In Russ.).
7. Analiz i modeli vremennykh ryadov. URL: <https://www.statmethods.ru/statistics-metody/modeli-vremennykh-ryadov/> (accessed: 15.05.2024). (In Russ.).
8. Petrov G. A., Shuranov E. V. Praktika ispolzovaniya veyvlet-analiza v defektoskopii. Saint Petersburg: Balt. gos. tekhn. un-t, 2012. 64 p. (In Russ.).

About the authors

O. V. Chubarova – Candidate of Sciences (Engineering), Associate Professor.

A. V. Chubarov – Candidate of Sciences (Engineering).

D. I. Liksonova – Candidate of Sciences (Engineering).

ФИЗИКО-МАТЕМАТИЧЕСКИЕ НАУКИ / PHYSICS AND MATHEMATICS



Научная статья

УДК 519.142.6:621.391

<https://doi.org/10.35266/1999-7604-2024-3-10>

Кодирование информации линейными перестановками дискретного преобразования Уолша

Михаил Сергеевич Беспалов¹, Кирилл Андреевич Фролов²✉

^{1,2}Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых, Владимир, Россия

¹bespalov@vlsu.ru, <https://orcid.org/0000-0003-0661-337X>

²golegoga33rus@gmail.com✉, <https://orcid.org/0000-0001-8691-8151>

Аннотация. В работе предложен метод использования кодовой матрицы в виде закрытого ключа для кода Уолша – Адамара. Посчитано число возможных кодовых матриц. Например, для матрицы Уолша порядка 32, то есть уровня 5, возможных ключей почти 10 миллионов. Показан способ декодирования информации и процедура выделения кодовой матрицы из зашифрованной матрицы Уолша.

Ключевые слова: линейный код, код Уолша – Адамара, дискретные преобразования Уолша, кодовая матрица, линейные перестановки

Для цитирования: Беспалов М. С., Фролов К. А. Кодирование информации линейными перестановками дискретного преобразования Уолша // Вестник кибернетики. 2024. Т. 23, № 3. С. 90–95. <https://doi.org/10.35266/1999-7604-2024-3-10>.

Original article

Information encoding by linear permutations of discrete Walsh transform

Mikhail S. Bespalov¹, Kirill A. Frolov²✉

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

¹bespalov@vlsu.ru, <https://orcid.org/0000-0003-0661-337X>

²golegoga33rus@gmail.com✉, <https://orcid.org/0000-0001-8691-8151>

Abstract. The paper proposes a method for using a code matrix as a private key for the Walsh-Hadamard code. The number of possible code matrices has been calculated. For example, there are almost 10 million possible keys for a Walsh matrix of the order 32, they are also level 5. The paper describes the method of decoding information and the procedure for isolating the code matrix from the encrypted Walsh matrix.

Keywords: linear code, Walsh-Hadamard code, discrete Walsh transform, code matrix, linear permutations

For citation: Bespalov M. S., Frolov K. A. Information encoding by linear permutations of discrete Walsh transform. *Proceedings in Cybernetics*. 2024;23(3):90–95. <https://doi.org/10.35266/1999-7604-2024-3-10>.

ВВЕДЕНИЕ

При передаче сигналов по неустойчивому каналу связи используют метод кодирования с исправлением ошибок, который применяется для устранения ошибок замещения символа [1]. Рассмотрим линейные блочные коды,

известные как (n, k) -коды, где n – длина кодового слова, k – длина кодируемого сообщения. Подробное обозначение кода (n, k, d) , где добавлен третий параметр d – кодовое расстояние, равное минимальному расстоянию Хэмминга между кодовыми словами [1].

Существует много работ, в которых приведено сравнение эффективности (n, k, d) -кодов [2, 3], подробно рассмотрена их классификация [4–6] и методы их декодирования [7].

При зашумленном канале (например, космическая связь) берут метод кодирования с максимально возможным d . К кодам такого типа относится код Уолша – Адамара, который служит линейным $(2^k, k, 2^{k-1})$ -кодом. Для него разработаны и известны процедуры обнаружения и исправления ошибок.

Космическая связь служит открытым каналом связи. Поскольку результатами разработчиков могут воспользоваться конкурентные структуры, перехватившие передаваемый сигнал, то код Уолша – Адамара желательно совместить с элементами шифрования сигнала.

В статье предлагается дополнить этот код матричным методом кодирования в виде линейной перестановки дискретного преобразования Уолша, разработанной одним из авторов [8, 9].

МАТЕРИАЛЫ И МЕТОДЫ

Применение кода Уолша – Адамара при передаче информации

Множество кодовых слов кода Уолша – Адамара совпадает с множеством дискретных функций Уолша в аддитивной записи. Поясним понятие аддитивной записи.

Группу из двух элементов алгебраисты изучают: в аддитивной форме $\mathbb{Z}_2 = \{0, 1; \oplus\}$ с операцией \oplus сложения по модулю 2; или в мультипликативной форме $\mathbb{Z}_2 = \{1, -1; \cdot\}$ с операцией \cdot умножения \cdot . К группе \mathbb{Z}_2 добавляется вторая операция, превращая ее в поле F_2 . Множество n -мерных векторов с элементами из группы \mathbb{Z}_2^n составляют n -мерное векторное пространство над полем F_2 с операциями \oplus покомпонентного сложения по модулю 2 (в аддитивном случае) или с операцией \bullet умножения по Адамару (в мультипликативном случае). В аддитивном случае вторая внешняя операция есть операция умножения на элементы поля 0 или 1, что реализуется при составлении линейной комбинации векторов как включение вектора в сумму (при коэффициенте 1) или невключение (при коэффициенте 0). Аналогично составляются линейные комбинации и в мультипликативном

случае, что позволяет специально не описывать вторую операцию. В этой терминологии можно предложить следующее алгебраическое описание дискретных функций Уолша.

Теорема 1. Для $N = 2^n$ множество дискретных функций Уолша есть подмножество элементов мультипликативного представления векторного пространства \mathbb{Z}_2^N , изоморфное векторному пространству \mathbb{Z}_2^n .

Все дискретные функции Уолша уровня n (порядка $N = 2^n$) записаны в строках матрицы Сильвестра – Адамара $H_n = H_n^{\otimes}$, полученной в виде n -й кронекеровой [10] степени матрицы:

$$H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (1)$$

По этому определению получаем рекуррентную формулу вычисления матриц Сильвестра – Адамара [9]:

$$H_{n+1} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}. \quad (2)$$

Линейное преобразование с матрицей H_n называется дискретным преобразованием Уолша (в нумерации Адамара) и встречается под аббревиатурой ДПУ (ДПУ – Адамара).

Проведя в H_n обратную перекодировку $1 \rightarrow 0, -1 \rightarrow 1$ из мультипликативной формы записи в аддитивную, в строках матрицы A_n получим все кодовые слова кода Уолша – Адамара соответствующего уровня. Например, при $k = 2$ кодовые слова в строках матрицы:

$$A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}. \quad (3)$$

Известен следующий способ непосредственного вычисления матриц A_n типа (3) без обращения к матрицам Сильвестра – Адамара (2).

Рекуррентно определяются матрицы C_m размера $m \times 2^m$:

$$C_1 = (0 \ 1), C_m = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ C_{m-1} & C_{m-1} \end{pmatrix}, \quad (4)$$

где $\mathbf{0} = (00 \dots 0)$, $\mathbf{1} = (11 \dots 1)$ – постоянные векторы соответствующей длины.

По правилам действия в поле F_2 вычисляем матрицу Уолша – Адамара (в аддитивной записи) как произведение матриц:

$$A_n = C_n^T \cdot C_n. \quad (5)$$

В этой формуле первый множитель произведения (5) (транспонированная матрица C_n^T) трактуем как набор всех возможных коэффициентов линейных комбинаций, а второй множитель трактуем как набор образующих, составляющих базис векторного подпространства $S \subset \mathbb{Z}_2^N$ кодовых слов. Строки матрицы A_n служат кодовыми словами длины $N = 2^n$, а соответствующие строки матрицы C_n^T являются соответствующими кодируемыми сообщениями длины n , что и приводит к $(2^n, n)$ -коду. Приведенные рассуждения можно рассматривать и как доказательство теоремы 1.

Основное преимущество этого кода состоит в максимально возможном кодовом расстоянии кода $d(K) = 2^{n-1}$, что вытекает из ортогональности [11] дискретных функций Уолша.

Кодовые сообщения легко получаются из кодовых слов выделением информационных координат с номерами 1, 2, 4, ..., 2^{n-1} . Точнее процедура декодирования формулируется следующим образом.

Утверждение 1. Если кодовые сообщения есть строки матрицы C_n^T вида (4), то информационными координатами являются координаты с номерами в порядке $2^{n-1}, 2^{n-2}, \dots, 4, 2, 1$ строк матрицы A_n кодовых слов.

Доказательство. Найдем матрицу R , выделяющую из матрицы A_n информационные столбцы. Это условие в матричном виде запишется $C_n^T = A_n \cdot R$, что можно переписать в виде $C_n^T = C_n^T \cdot C_n \cdot R$. Значит, матрицу R ищем из условия $C_n \cdot R = E$ равенства единичной матрице порядка n . Столбцы стандартного базиса в матрице C_n упорядочены под номерами $2^{n-1}, 2^{n-2}, \dots, 4, 2, 1$, поскольку порядок координат берется слева направо. Поэтому в столбцах матрицы R размера $2^n \times n$ единицы на тех же указанных позициях, а остальные элементы – нули. Согласно формуле $C_n^T = A_n \cdot R$ номера позиций в столбцах матрицы R (при умножении справа) и есть указание на порядок информационных координат.

Именно тем, что порядок координат в общепринятом описании метода кодирования матрицами Уолша – Адамара получается обратный, вызван другой подход в определении матриц C_n , в статье [8] и пособии [9].

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Кодирование информации линейными перестановками дискретного преобразования Уолша

Если вернуться к теореме 1 и формуле (5), то можно заменить упорядоченный базис в виде строк матрицы C_n на любой другой упорядоченный базис того же n -мерного пространства кодовых слов S .

Теорема 2. Произвольная невырожденная над полем F_2 матрица K порядка n формулой $K \cdot C_n$ задает упорядоченный базис n -мерного пространства кодовых слов S , а строки матрицы

$$A(K) = C_n^T \cdot K \cdot C_n \quad (6)$$

повторяют (в другом порядке) строки матрицы A_n , то есть множество кодовых слов S .

Доказательство. Невырожденность матрицы K гарантирует линейную независимость строк матрицы $K \cdot C_n$. Базисность строк доказана. Упорядоченность базиса берется в обратном порядке, снизу вверх по нумерации строк.

Строки матрицы C_n^T составляют полный набор возможных линейных комбинаций в виде всех 2^n возможных кодовых сообщений длины n . Произведение $C_n^T \cdot K$ дает тот же набор линейных комбинаций (за счет невырожденности K), но в другом порядке.

Замечание. Если в матрице $A(K)$ произвести перекодировку $0 \rightarrow 1, 1 \rightarrow -1$ из аддитивного представления в мультипликативное, то получим [9] матрицу линейной перестановки ДПУ.

Таким образом, предлагается использовать аддитивное представление матрицы линейной перестановки ДПУ для шифрования информации. Отображение $C_n^T \rightarrow C_n^T \cdot K$ задает перестановку строк $A_n \rightarrow A(K)$. Обратное отображение $C_n^T \leftarrow C_n^T \cdot K$ задает обратную перестановку строк $A_n \leftarrow A(K)$, которая применяется при декодировании. При этом процедура поиска и исправления ошибок остается прежней, что позволяет использовать все существующие аппаратные средства. Секретным ключом служат кодовая матрица K и тесно связанная с ней матрица $A(K)$.

Если в одном направлении этим методом передается ключ, то для передачи информации в обратном направлении процедура шифрования может пройти по составленной программе без участия оператора. Знание ключа позволяет легко организовать декодирование.

Теорема 3. Если кодовые сообщения есть строки матрицы C_n^T вида (4), то упорядоченными информационными координатами описанного метода кодирования кодовыми словами $A(K)$ вида (5) будут упорядоченные двоичные числа в столбцах матрицы K^{-1} .

Доказательство по схеме рассуждений в утверждении 1. Найдем матрицу T , выделяющую из матрицы $A(K)$ информационные столбцы. Это условие в матричном виде запишется $C_n^T = A(K) \cdot T$, что перепишем в виде $C_n^T = C_n^T \cdot K \cdot C_n \cdot T$. Значит, матрицу T ищем из условия $K \cdot C_n \cdot T = E$. Домножив слева на обратную K^{-1} к кодовой матрице K , получим, что $C_n \cdot T = K^{-1}$. Значит матрица T при умножении справа выделяет из столбцов матрицы C_n столбцы матрицы K^{-1} . Поскольку столбцы матрицы C_n упорядочены по возрастанию чисел в двоичной системе счисления, представленных в ее столбцах, то единицы в матрице T на тех позициях, которые в двоичной системе счисления записаны в столбцах матрицы K^{-1} . Согласно формуле $C_n^T = A(K) \cdot T$, номера позиций в столбцах и есть указание на порядок информационных координат.

Пример. Действие теоремы 3 и принцип кодирования продемонстрируем примером для нумерации Уолша уровня 3 с кодовой матрицей $K = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. Матрица C_3 имеет вид:

$$C_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

По формуле (6) получается следующая матрица:

$$A(K) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Поскольку обратная к кодовой есть матрица $K^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, то информационные столбцы идут в порядке 1, 3, 7 (нумерация с нуля), что также легко заметить из вида матрицы $A(K)$.

Предположим, что при кодировании этим методом по каналу удаленной связи получили сообщение 00111100010110111111100 из 24 знаков. Разобьем сообщение на блоки (восьмерки бит) и получим набор кодовых сообщений: 00111100, 01011011, 11111100.

Для каждого сообщения вычисляем синдром, равный расстоянию Хэмминга этого сообщения до всех кодовых слов. Для первого сообщения 00111100 синдром будет равен: 4,4,0,4,4,4,4,4. Данный набор означает, что сообщение не содержит ошибок и совпадает с кодовым словом 00111100 под номером 2. Синдром второго сообщения 01011011 равен 5,3,5,3,5,3,1,3. Минимальное расстояние равно 1 до кодового сообщения с номером 6. Значит, в сообщении одна ошибка, исправив которую, получим 01011010. Синдром третьего сообщения 11111100 равен 6,6,2,6,4,4,4,4. Минимальное расстояние равно 2 до кодового сообщения с номером 2. Значит, в сообщении две ошибки, исправив которые, получим 00111100. После исправления ошибок и разделения на кодовые слова получим переданное сообщение 00111100, 01011010, 00111100.

Выделением координат с номерами 1, 3 и 7 получим кодируемое сообщение: 010, 110, 010.

Поскольку сообщение передавалось по открытому каналу связи, то перехватить его и исправить до 00111100, 01011010, 00111100 стандартным методом может любой пользователь. А вот выделить кодируемое сообщение – 010, 110, 010 – способен лишь обладатель ключа в виде матрицы K или K^{-1} .

На практике данный метод кодирования, естественно применять для уровня большего чем 3. Тогда сложность декодирования определяется числом возможных кодовых матриц.

В источнике [9] приведена следующая формула для числа различных невырожденных бинарных матриц порядка m :

$$sb(m) = (2^m - 1)(2^m - 2)(2^m - 4) \dots (2^m - 2^{m-1}).$$

В частности: $sb(2) = 6$, $sb(3) = 168$, $sb(4) = 20\ 160$, $sb(5) = 9\ 999\ 360$. В источнике [12] тот же результат (до уровня 4) вычислен компьютерным перебором. Экспоненциальный рост числа возможных кодовых матриц, а также отсутствие приемов декодирования перехваченного сообщения превращают этот метод в сложную математическую задачу.

Закрытым ключом данного метода кодирования служит кодовая матрица K . Также к секретной информации относятся тесно взаимосвязанные с ней матрицы K^{-1} и $A(K)$. Покажем метод восстановления кодовой матрицы K по матрице $A(K)$.

Сначала введем новое понятие инверсии квадратной матрицы, где принята нумерация с нуля. При операции инверсии все элементы матрицы порядка N симметрично отражаются относительно центра матрицы:

$$\alpha_{kj} \rightarrow \alpha_{N-k-1, N-j-1}.$$

Другой вариант определения операции инверсии: это сначала инверсия всех строк, а потом инверсия всех столбцов.

Например, инверсия матрицы $K \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ будет $Inv(K) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

Теорема 4. Главная подматрица матрицы $A(K)$, выделенная выборкой $(1, 2, \dots, 2^{n-1})$, есть инверсия ее кодовой матрицы.

Доказательство. В строках матрицы C_n^T с номерами $1, 2, \dots, 2^{n-1}$ записаны векторы стандартного базиса в обратном порядке. Например, для $n = 3$ это векторы $(0, 0, 1)$, $(0, 1, 0)$ и $(1, 0, 0)$. При умножении матрицей C_n^T на матрицу $K \cdot C_n$ слева, j -я строка произведения есть линейная комбинация строк матрицы $K \cdot C_n$ с коэффициентами, указанными в j -й строке матрицы C_n^T (метод описан в источнике [13]). Поэтому на месте строки с номером 1 матрицы $A(K)$ будет последняя строка матрицы $K \cdot C_n$, на месте строки с номером 2 – предпоследняя строка матрицы $K \cdot C_n$, на месте строки с номером 4 – третья снизу строка матрицы $K \cdot C_n$ и т. д.

Теперь аналогично рассмотрим произведение матриц $K \cdot C_n$ как умножение справа C_n на матрицу K . При правом умножении m -й столбец матрицы $K \cdot C_n$ есть линейная комбинация столбцов матрицы K с коэффициентами, указанными в m -м столбце матрицы C_n . Столбцы, выделенные той же выборкой, также будут элементами стандартного базиса. Но ввиду порядка следования этих столбцов в матрице C_n , при изменении значения m по элементам выборки $(1, 2, \dots, 2^{n-1})$ будут получаться столбцы матрицы K в порядке $n - 1, n - 2, \dots, 1, 0$ (обратный порядок их степеней).

Доказанная теорема объясняет, почему общепринятый в теории кодирования метод определения матриц C_n , использованный в данной статье, хуже, чем новый метод определения аналогичных (но других) матриц, обозначенных тем же символом C_n [8, 9].

Если бы в качестве матриц C_n брали матрицы, определенные в источниках [8, 9], то теорема 4 формулировалась бы проще: указанная выборка выделяет кодовую матрицу (а не ее инверсию).

При применении рассмотренного метода кодирования, то есть при распространении метода с двух известных нумераций на примерно 10 миллионов (при уровне 5) нумераций ДПУ, естественно применять более удобную конструкцию определения C_n [8, 9].

ЗАКЛЮЧЕНИЕ

Итак, зная кодовое слово в качестве закрытого ключа, отправитель может построить аддитивную матрицу дискретного преобразования Уолша $A(K)$, с помощью которой закодировать сообщение и передать получателю. При декодировании любой получатель посредством существующей стандартной аппаратуры обнаруживает и исправляет ошибки передачи информации. Для уровня 5 кодовой матрицы гарантируется исправление семи ошибок в передаваемом сообщении из 32 символов. Получатель, зная закрытый ключ K , выделяет из полученного и исправленного сообщения исходное сообщение. В случае перехвата без знания закрытого ключа существует около 10 миллионов различных вариантов

дискретных преобразований Уолша для минимальной для практики размерности кодовой матрицы порядка 5. В реальных каналах связи

используется значительно больший порядок t кодовых матриц, что делает нереализуемым процесс дешифровки сигнала.

Список источников

1. Питерсон У. У., Уэлдон Э. Коды, исправляющие ошибки / пер. с англ. ; под ред. Р. Л. Добрушина, С. И. Самойленко. М. : Мир, 1976. 594 с.
2. Дворников С. В., Устинов А. А., Дворников С. С. и др. Анализ эффективности блочных кодов // Вопросы радиоэлектроники. Серия: Техника телевидения. 2011. № 1. С. 63–73.
3. Леонтьев В. К., Мовсисян Г. Л., Маргарян Ж. Г. Верхняя и нижняя границы мощности кода, исправляющего ошибки алгебраического канала // ՀՀ ԳԱԱ Ձեռնարկներ : доклады НАН РА. 2020. Т. 120, № 1. С. 7–14.
4. Tinnirello C. Cyclic Codes: Low-Weight Codewords and Locators. PhD thesis, University of Trento, 2016. 133 p.
5. Костюков А. С., Башкиров А. В., Никитин Л. Н. и др. Помехоустойчивое кодирование в современных форматах связи // Вестник Воронежского государственного технического университета. 2019. Т. 15, № 2. С. 132–138.
6. Ceria M., Mora T., Sala M. Groebner bases and error correcting codes: from Cooper Philosophy to Degrobnerization // International Conference on Polynomial Computer Algebra. Saint Petersburg, October, 2020. P. 45–48.
7. Рацеев С. М., Иванцов А. М., Булдаковский П. А. Об алгоритмах декодирования циклических кодов // Ученые записки УлГУ. Серия Математика и информационные технологии. 2021. № 1. С. 87–101.
8. Беспалов М. С. Собственные подпространства дискретного преобразования Уолша // Проблемы передачи информации. 2010. Т. 46, № 3. С. 60–79.
9. Беспалов М. С., Склярченко В. А. Дискретные функции Уолша и их приложения. Владимир : ВлГУ, 2014. 67 с.
10. Беллман Р. Введение в теорию матриц / пер. с англ. В. Я. Катковника, Р. А. Полуэкова, М. С. Эпельмана ; под ред. В. Б. Лидвского. М. : Наука, 1969. 368 с.
11. Ахмед Н., Рао К. Р. Ортогональные преобразования при обработке цифровых сигналов. М. : Связь, 1980. 248 с.
12. Зяблицева Л. В., Корабельщикова С. Ю., Чесноков А. И. Линейные коды, исправляющие ошибки, и алгоритмы их подсчета // Эвристические алгоритмы и распределенные вычисления. 2014. Т. 1, № 3. С. 47–59.
13. Малоземов В. Н. Линейная алгебра без определителей. Квадратичная функция. СПб. : СПбГУ, 1997. 80 с.

Информация об авторах

М. С. Беспалов – доктор физико-математических наук, профессор.

К. А. Фролов – аспирант.

References

1. Piterson U. U., Ueldon E. Kody, ispravlyayushchie oshibki. Trans., eds. R. L. Dobrushin, S. I. Samoylenko. Moscow: Mir, 1976. 594 p. (In Russ.).
2. Dvornikov S. V., Ustinov A. A., Dvornikov S. S. et al. Analiz effektivnosti blokovykh kodov. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2011;(1):63–73. (In Russ.).
3. Leontiev V. K., Movsisian G. L., Margaryan Zh. G. Upper and lower bounds of the power of the error correction code of an algebraic channel. *ՀՀ ԳԱԱ Ձեռնարկներ : reports of the National academy of sciences of Armenia*. 2020;120(1):7–14. (In Russ.).
4. Tinnirello C. Cyclic Codes: Low-Weight Codewords and Locators. PhD thesis, University of Trento, 2016. 133 p.
5. Kostyukov A. S., Bashkirov A. V., Nikitin L. N. et al. Anti-interference coding in modern communication formats. *Bulletin of Voronezh State Technical University*. 2019;15(2):132–138. (In Russ.).
6. Ceria M., Mora T., Sala M. Groebner bases and error correcting codes: from Cooper Philosophy to Degrobnerization. In: *International Conference on Polynomial Computer Algebra*. Saint Petersburg, October. 2020. P. 45–48.
7. Ratseev S. M., Ivantsov A. M., Buldakovskiy P. A. Ob algoritmakh dekodirovaniya tsiklicheskih kodov. *Uchenye zapiski UIGU. Seriya Matematika i informatsionnye tekhnologii*. 2021;(1):87–101. (In Russ.).
8. Bespalov M. S. Sobstvennye podprostranstva diskretnogo preobrazovaniya Uolsha. *Problems of Information Transmission*. 2010;46(3):60–79. (In Russ.).
9. Bespalov M. S., Sklyarenko V. A. Diskretnye funktsii Uolsha i ikh prilozheniya. Vladimir: VIGU, 2014. 67 p. (In Russ.).
10. Bellman R. Vvedenie v teoriyu matrits. Trans. V. Ya. Katkovnik, R. A. Poluekov, M. S. Epelman; Ed. V. B. Lidvskiy. Moscow: Nauka, 1969. 368 p. (In Russ.).
11. Akhmed N., Rao K. R. Ortogonalnye preobrazovaniya pri obrabotke tsifrovyykh signalov. Moscow: Svyaz, 1980. 248 p. (In Russ.).
12. Zyablitseva L. V., Korabelshchikova S. Y., Chesnokov A. I. Linear error-correcting codes and algorithms for their calculations. *Evristsicheskie algoritmy i raspredelennye vychisleniya*. 2014;1(3):47–59. (In Russ.).
13. Malozemov V. N. Lineynaya algebra bez opredeliteley. Kvadrachnaya funktsiya. Saint Petersburg: SPbGU, 1997. 80 p. (In Russ.).

About the authors

M. S. Bespalov – Doctor of Sciences (Physics and Mathematics), Professor.

K. A. Frolov – Postgraduate.

Научная статья
УДК 621.396.962.3
<https://doi.org/10.35266/1999-7604-2024-3-11>



Оценка минимального достижимого уровня боковых лепестков функции неопределенности радиоимпульсных последовательностей

Анатолий Ильич Дёмко^{1✉}, Елена Геннадьевна Алмазова²

^{1,2}Сургутский государственный университет, Сургут, Россия

¹dai321@mail.ru[✉], <https://orcid.org/0000-0003-0747-0294>

²egalmazova@mail.ru, <https://orcid.org/0000-0001-5523-7222>

Аннотация. В радиолокации и связи широко используются сложные сигналы, что определяет актуальность синтеза, исследования и оптимизации подобных сигналов. Целью статьи является получение расчетных формул оценки минимальной границы достижимого уровня боковых лепестков нулевого сечения функции неопределенности дискретно кодированных радиоимпульсных последовательностей. Введено понятие коэффициента оптимальности, показывающего превышение уровня боковых лепестков над потенциально достижимым минимальным значением. Показано, что радиоимпульсные последовательности на основе кодов Баркера являются оптимальными по критерию минимального уровня боковых лепестков нулевого сечения функции неопределенности. Приведены результаты проведенного анализа оптимальности многофазных радиоимпульсных последовательностей на основе кода Фрэнка, многофазных последовательностей на основе кода Чу и фазоманипулированных М-последовательностей. Полученные результаты могут быть использованы в радиолокации и связи.

Ключевые слова: нулевое сечение, функция неопределенности, боковые лепестки, радиоимпульсные последовательности, коэффициент оптимальности

Для цитирования: Дёмко А. И., Алмазова Е. Г. Оценка минимального достижимого уровня боковых лепестков функции неопределенности радиоимпульсных последовательностей // Вестник кибернетики. 2024. Т. 23, № 3. С. 96–102. <https://doi.org/10.35266/1999-7604-2024-3-11>.

Original article

Assessment of minimum achievable sidelobes level of ambiguity function for radio pulse sequences

Anatoly I. Dyomko^{1✉}, Elena G. Almazova²

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

¹dai321@mail.ru[✉], <https://orcid.org/0000-0003-0747-0294>

²egalmazova@mail.ru, <https://orcid.org/0000-0001-5523-7222>

Abstract. The wide usage of complex signals in radiolocation and communication determines the relevance of synthesizing, researching, and optimizing such signals. This article aims to derive calculation formulas for estimating the minimum limit of the achievable level of sidelobes of the zero section for the ambiguity function of discretely encoded radio pulse sequences. The introduction of the optimality coefficient provides a measure of the excess of the sidelobe level above the potentially achievable minimum. Radio pulse sequences based on Barker codes are shown to be optimal according to the criterion of the minimum level of the sidelobes of the zero section for the ambiguity function. The results of the optimality analysis of multiphase radio pulse sequences based on the Frank code, multiphase sequences based on the Chu code and phase-manipulated M-sequences are presented. The results can be used in radiolocation and communication.

Keywords: zero cross section, ambiguity function, sidelobes, radio pulse sequences, optimality coefficient

For citation: Dyomko A. I., Almazova E. G. Assessment of minimum achievable sidelobes level of ambiguity function for radio pulse sequences. *Proceedings in Cybernetics*. 2024;23(3):96–102. <https://doi.org/10.35266/1999-7604-2024-3-11>.

ВВЕДЕНИЕ

В радиолокации, радионавигации и электрической связи широко применяются сложные сигналы, имеющие произведение длительности на ширину спектра больше единицы. Способность совместного разрешения по дальности и скорости определяется функцией неопределенности сигнала [1], которая характеризует степень отличия сигнала от своих копий, сдвинутых по времени и частоте во всем диапазоне возможных значений. Функция неопределенности характеризуется областью сильной корреляции (главный максимум) и областью слабой корреляции (боковые лепестки). Наличие боковых лепестков является нежелательным фактором, ухудшающим качество решения радиолокационных, радионавигационных и связных задач. В связи с этим актуальной является оценка оптимальности сигнала с точки зрения степени близости уровня боковых лепестков функции неопределенности к минимально достижимой величине.

МАТЕРИАЛЫ И МЕТОДЫ

Для радиотехнических систем, обработка сигнала в которых позволяет использовать несколько периодов повторения, можно использовать периодическую функцию неопределенности (ФН, ПФН), существуют сигналы с нулевым уровнем боковых лепестков нулевого сечения ПФН (сигналы с идеальными периодическими корреляционными свойствами) [2]. При обработке одного периода сигнала следует рассматривать импульсную (одиночную) ФН (1), уровень боковых лепестков нулевого сечения которой ненулевой и зависит от конкретного сигнала.

$$\chi(\tau, F) = \int_{\tau}^{T_c} \dot{S}(t) \cdot \dot{S}^*(t - \tau) e^{j2\pi Ft} dt, \quad 0 \leq \tau \leq T_c, \quad (1)$$

где $\dot{S}(t)$ – комплексная форма представления сигнала;

$\dot{S}^*(t)$ – комплексно-сопряженный сигнал;

τ – временной сдвиг;

F – частотный сдвиг;

T_c – длительность сигнала.

Для примера на рис. 1 приведено тело неопределенности (трехмерное графическое изображение импульсной ФН) многофазной радиоимпульсной последовательности на основе кода Фрэнка из 64 элементов [3, 4].

Нормирование графика выполнено относительно начала координат $x(0,0)$. Центр изображения ($\tau = 0, F = 0$) соответствует главному максимуму ФН, где оба сигнала в подынтегральном выражении (1) максимально похожи друг на друга. Остальная область тела неопределенности меньшего уровня содержит боковые лепестки (в идеальном случае должны отсутствовать), характеризующие степень непохожести сигналов подынтегрального выражения (1) при отдалении от начала координат ($\tau \neq 0, F \neq 0$). Наблюдается сложная структура боковых лепестков во всем интервале возможных временных задержек и частотных сдвигов (1).

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Определим нижнюю оценку уровня боковых лепестков нулевого сечения ФН, ниже которой уровень боковых лепестков получить невозможно для дискретно кодированных сигналов в виде радиоимпульсных последовательностей.

Радиоимпульсная последовательность из N радиоимпульсов:

$$s(t) = \begin{cases} \sum_{n=0}^{N-1} S_m(t) e^{j\psi_n(t)}, & nT + \tau_n \leq t \leq nT + \tau_n + \tau_u; \\ 0, & nT + \tau_n + \tau_u < t < (n+1)T + \tau_{n+1}, \end{cases} \quad (2)$$

где $S_m(t)$ – амплитуда радиоимпульса (возможна внутриимпульсная амплитудная, частотная или фазовая модуляция);

n – порядковый номер радиоимпульса;

T – период следования радиоимпульсов;

$\psi_n(t)$ – начальная фаза n -го радиоимпульса;

τ_n – сдвиг радиоимпульса относительно тактового (nT) момента времени;

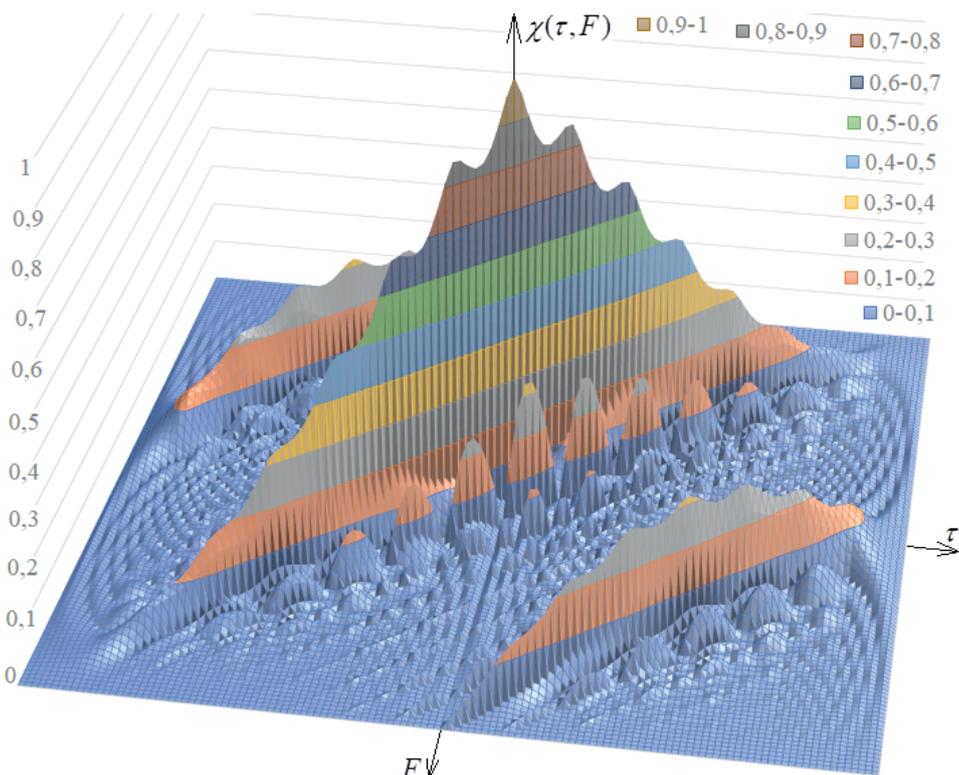


Рис. 1. Огибающая нормированного тела неопределенности сигнала на основе кода Фрэнка для последовательности из 64 элементов

Примечание: составлено авторами на основании данных, полученных в исследовании.

τ_u – длительность радиоимпульса;

N – число радиоимпульсов в последовательности.

$$T_c = NT. \quad (3)$$

Можно не ограничивать вид последовательности, допуская как однородные последовательности (следующие с одинаковым сдвигом радиоимпульсов относительно предыдущих, с модуляцией одного параметра, например начальной фазы радиоимпульсов), так и различные варианты неоднородных и разнопараметрических радиоимпульсных последовательностей.

$$\chi(\tau, F) = \int_{\tau}^{T_c} \left\{ \sum_{n=0}^{N-1} S_m(t) \exp j\psi_n(t) \right\} \times \left\{ \sum_{k=0}^{N-1} S_m(t) \exp j[-\psi_k(t-\tau)] \right\} e^{j2\pi Ft} dt. \quad (4)$$

Найдем значения ФН при нулевом частотном сдвиге F (нулевое сечение функции неопределенности) при двух значениях сдвига по оси задержек:

а) нулевом ($\tau = 0$), когда при вычислении интеграла (4) пересекаются все радиоимпульсы с одинаковыми номерами;

б) максимальном ($\tau = NT - T$), когда пересекаются только последний радиоимпульс исходной последовательности с первым радиоимпульсом задержанной последовательности.

Значение функции неопределенности при нулевом ($\tau = 0$) временном сдвиге:

$$\chi(0,0) = \int_0^{T_c} \dot{S}(t) \cdot \dot{S}^*(t) dt, \quad 0 \leq \tau \leq T_c. \quad (5)$$

$$\chi(0,0) = \int_0^{T_c} \left\{ \sum_{n=0}^{N-1} S_m(t) \exp j\psi_n(t) \right\} \left\{ \sum_{k=0}^{N-1} S_m(t) \exp j[-\psi_k(t)] \right\} dt \quad (6)$$

$$\chi(0,0) = \int_0^{T_c} \sum_{n=0}^{N-1} S_m^2(t) dt = \sum_{n=0}^{N-1} \int_0^{\tau_u} S_m^2(t) dt. \quad (7)$$

Для получения замкнутых форм выражения (7) ограничимся рассмотрением вариантов радиоимпульсных последова-

тельностью с прямоугольной огибающей радиоимпульсов:

$$S_m(t) = S_m = \text{const}, \quad (8)$$

$$\chi(0,0) = \sum_{n=0}^{N-1} S_m^2 \int_0^{\tau_u} dt = \sum_{n=0}^{N-1} S_m^2 \tau_u = NS_m^2 \tau_u. \quad (9)$$

Абсолютное значение нулевого сечения функции неопределенности при нулевом временном сдвиге (9) пропорционально энергии радиоимпульса и числу радиоимпульсов в последовательности.

Вычислим значение нулевого сечения функции неопределенности при максимальном временном сдвиге, когда пересекаются только два радиоимпульса:

$$\tau_{\max} = (N-1)T. \quad (10)$$

При этом в подынтегральном выражении (6) исчезают суммы, поскольку ненулевым остается лишь интеграл от произведения последнего радиоимпульса первой последовательности с первым радиоимпульсом второй последовательности.

$$\chi((N-1)T,0) = \int_{(N-1)T}^{(N-1)T+\tau_u} \left\{ \sum_{n=N-1}^{N-1} S_m \exp j\psi_n(t) \right\} \left\{ \sum_{k=0}^0 S_m \exp j[-\psi_k(t)] \right\} dt \quad (11)$$

$$\chi((N-1)T,0) = \int_{(N-1)T}^{(N-1)T+\tau_u} \{S_m \exp j\psi_{N-1}(t)\} \{S_m \exp j[-\psi_0(t)]\} dt \quad (12)$$

$$\chi((N-1)T,0) = S_m^2 \int_{(N-1)T}^{(N-1)T+\tau_u} \exp j\{\psi_{N-1}(t) - \psi_0(t)\} dt. \quad (13)$$

Дальнейшие преобразования интеграла (13) требуют конкретизации закона изменения фазы $\psi(t)$, однако модуль интеграла от гармонической функции не зависит от аргумента гармонической функции, а определяется только пределами интегрирования, поэтому верхнюю оценку интеграла можно записать в виде:

$$|\chi((N-1)T,0)| = S_m^2 \tau_u. \quad (14)$$

На основании выражений (9) и (14) можно произвести оценку нормированной величины бокового лепестка нулевого сечения функции неопределенности при сдвиге по оси задержки на величину $(N-1)T$. Назовем эту величину коэффициентом оценки уровня боковых лепестков дискретно кодированных сигналов

$$k_{N-1} = \frac{|\chi((N-1)T,0)|}{\chi(0,0)} = \frac{S_m^2 \tau_u}{NS_m^2 \tau_u} = 1/N = N^{-1}. \quad (15)$$

Расчетная формула для вычисления коэффициента оценки уровня боковых лепестков k_{N-1} в логарифмических единицах.

$$k_{N-1} = 201g(N^{-1}) = -201g(N), \quad \text{дБ}. \quad (16)$$

Очевидно, что при любых меньших временных сдвигах kT ($k \neq N-1$) уровень боковых лепестков нулевого сечения функции неопределенности может как больше, так и меньше, но максимальное значение M_{\max} уровня боковых лепестков не может быть меньше, чем величина k_{N-1} . Для примера на рис. 2 изображено нормированное нулевое сечение ФН многофазной радиоимпульсной последовательности на основе кода Фрэнка из 256 элементов и коэффициент оценки уровня боковых лепестков k_{255} .

Уровень боковых лепестков при таком числе элементов в последовательности составляет $M_{\max} = -34$ дБ. Боковые лепестки располагаются симметрично относительно главного максимума (вне зависимости от знака временной задержки), имеют сложную структуру и большой динамический диапазон (на рисунке 29 дБ – от минус 34 дБ до минус 63 дБ, хотя минимальные значения достигают минус 100 дБ).

Практическая полезность коэффициента оценки уровня боковых лепестков k_{N-1} определяется двумя факторами:

а) k_{N-1} определяет минимально достижимый уровень боковых лепестков нулевого сечения функции неопределенности при данном числе элементов в последовательности для данной радиоимпульсной последовательности;

б) k_{N-1} соответствует значению уровня бокового лепестка при максимально возможной величине временного сдвига $\pm (N-1)T$.

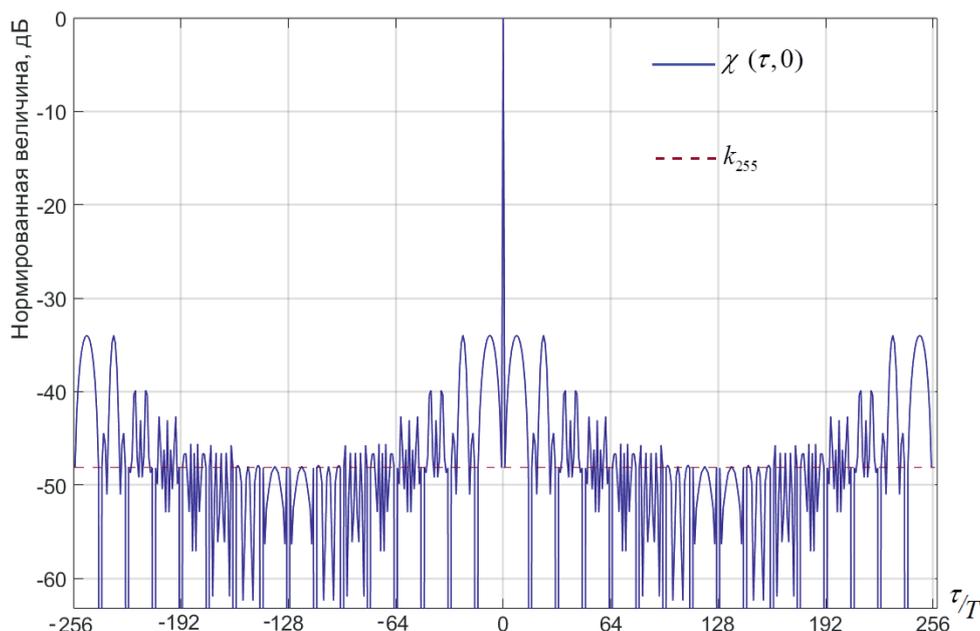


Рис. 2. Функция неопределенности многофазной радиоимпульсной последовательности на основе кода Фрэнка из 256 элементов и коэффициент оценки уровня боковых лепестков k_{255}
 Примечание: составлено авторами на основании данных, полученных в исследовании.

Введем параметр, характеризующий величину отличия уровня боковых лепестков M_{max} нулевого сечения функции неопределенности от коэффициента оценки уровня боковых лепестков k_{N-1} и назовем его коэффициентом оптимальности k_{opt} :

$$k_{opt} = k_{N-1} - M_{max} = -20 \lg(N) - M_{max}, \text{ дБ.} \quad (17)$$

Коэффициент оптимальности большинства сигналов является отрицательным: чем больше по модулю его значение, тем в большей степени корреляционные свойства отличаются от оптимальных.

Известны бинарные коды Баркера [5] для 3, 4, 5, 7, 11 и 13 элементов в последовательности. Уровень боковых лепестков нулевого сечения функции неопределенности радиоимпульсных последовательностей на основе кодов Баркера совпадает с (16), поэтому для них коэффициент оптимальности k_{opt} равен нулю, следовательно такие сигналы имеют минимально возможный уровень боковых лепестков.

На рис. 3 приведены зависимости коэффициента оптимальности для многофазных последовательностей на основе кода Фрэнка, многофазных последовательностей на основе

кода Чу [6, 7] и псевдослучайных последовательностей максимальной длины (М-последовательностей) [8].

Как видно, для многофазных и фазоманипулированных радиоимпульсных последовательностей наиболее оптимальными (корреляционные параметры которых максимально приближаются к потенциально достижимым) являются последовательности небольшой длины. При увеличении длины последовательности становятся менее оптимальными, следовательно увеличиваются возможности для дополнительной оптимизации с целью улучшения корреляционных свойств сигналов [9].

Аналогично можно оценить степень отличия от оптимальных (по критерию максимального уровня боковых лепестков нулевого сечения функции неопределенности) для других радиоимпульсных последовательностей.

ЗАКЛЮЧЕНИЕ

Для радиоимпульсных сигналов уровень бокового лепестка на границах максимальной задержки не зависит от вида модуляции, определяясь числом элементов в последовательности N .

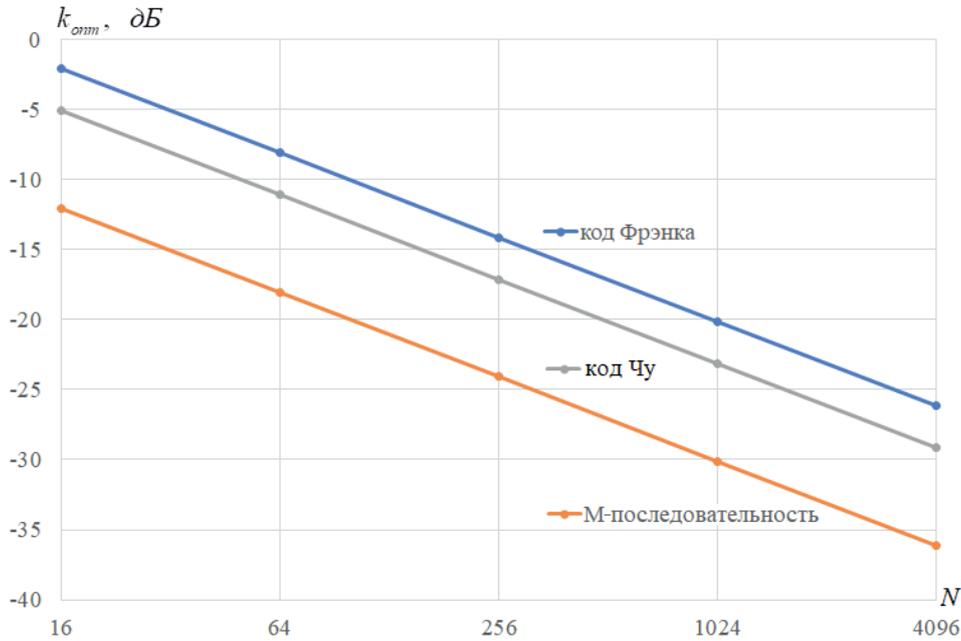


Рис. 3. Зависимость коэффициента оптимальности от числа элементов в последовательности

Примечание: составлено авторами на основании данных, полученных в исследовании.

Коэффициент оценки уровня боковых лепестков радиоимпульсной последовательности k_{N-1} (18):

$$k_{N-1} = -201g(N), \quad \text{дБ} \quad (18)$$

определяет минимально-достижимый уровень боковых лепестков нулевого сечения функции неопределенности при данном числе элементов N в последовательности для данной радиоимпульсной последовательности и соответствует значению уровня бокового лепестка при максимально возможной величине временного сдвига $\pm (N-1)T$.

Коэффициент оптимальности k_{opt} (19) характеризует отличие максимального уровня боковых лепестков радиоимпульсных последовательностей

M_{max} нулевого сечения функции неопределенности от потенциально достижимого:

$$k_{opt} = k_{N-1} - M_{max} = -201g(N) - M_{max}, \quad \text{дБ.} \quad (19)$$

Коэффициент оптимальности k_{opt} радиоимпульсных сигналов на основе бинарных кодов Баркера равен нулю, т. е. такие сигналы имеют минимально достижимый уровень боковых лепестков нулевого сечения функции неопределенности.

Для многофазных последовательностей на основе кода Фрэнка, многофазных последовательностей на основе кода Чу и M-последовательностей коэффициент оптимальности k_{opt} уменьшается при увеличении числа радиоимпульсов N в последовательности со скоростью 6 дБ/окт.

Список источников

1. Вакман Д. Е. Сложные сигналы и принцип неопределенности в радиолокации. М. : Сов. Радио, 1965. 304 с.
2. Кук Ч., Бернфельд М. Радиолокационные сигналы. Теория и применение / пер. с англ. ; под ред. В. С. Кельзона. М. : Сов. Радио, 1971. 567 с.
3. Дёмко А. И. О связи корреляционных свойств сигналов на основе кодов Фрэнка с погрешностями их

References

1. Vakman D. E. Slozhnye signaly i printsip neopredelenosti v radiolokatsii. Moscow: Sov. Radio, 1965. 304 p. (In Russ.).
2. Kuk Ch., Bernfeld M. Radiolokatsionnye signaly. Teoriya i primeneniye. Trans. and ed. V. S. Kelzon. Moscow: Sov. Radio, 1971. 567 p. (In Russ.).
3. Dyomko A. I. On relation between correlation characteristics of signals based on the frank codes and

- формирования // Вестник кибернетики. 2023. Т. 22, № 3. С. 23–31.
4. Frank R., Zadoff S., Heimiller R. Phase shift pulse codes with good periodic correlation properties (corresp.) // *IRE Transactions on Information Theory*. 1962. Vol. 8, no. 6. P. 381–382.
 5. Дворников С. В., Дворников С. С., Марков Е. В. Модифицированные импульсные последовательности на основе кодов Баркера // Труды учебных заведений связи. 2022. Т. 8, № 1. С. 8–14. <https://doi.org/10.31854/1813-324X-2022-8-1-8-14>.
 6. Chu D. Polyphase codes with good periodic correlation properties (corresp.) // *IEEE Transactions on information theory*. 1972. Vol. 18, no. 4. P. 531–532.
 7. Zhou Z., Hellesteth T., Paramalli U. A family of polyphase sequences with asymptotically optimal correlation // *IEEE Transactions on Information Theory*. 2018. Vol. 64. no. 4. P. 2896–2900.
 8. Сенин А. И. Корреляционные свойства последовательностей, построенных на основе М-последовательностей и последовательностей Уолша // Вестник Московского государственного технического университета им. Н. Э. Баумана. Серия «Приборостроение». 2014. № 5. С. 88–97.
 9. Вакман Д. Е., Седлецкий Р. М. Вопросы синтеза радиолокационных сигналов. М. : Сов. радио, 1973. 312 с.
4. Frank R., Zadoff S., Heimiller R. Phase shift pulse codes with good periodic correlation properties (corresp.). *IRE Transactions on Information Theory*. 1962;8(6):381–382.
 5. Dvornikov S., Dvornikov S., Markov E. Modified Pulse Sequences Based on Barker Codes. *Proceedings of Telecommunication Universities*. 2022;8(1):8–14. <https://doi.org/10.31854/1813-324X-2022-8-1-8-14>. (In Russ.).
 6. Chu D. Polyphase codes with good periodic correlation properties (corresp.). *IEEE Transactions on information theory*. 1972;18(4):531–532.
 7. Zhou Z., Hellesteth T., Paramalli U. A family of polyphase sequences with asymptotically optimal correlation. *IEEE Transactions on Information Theory*. 2018;64(4):2896–2900.
 8. Senin A. I. Correlation properties of sequences formed on the basis of m-sequences and Walsh sequences. *Herald of the Bauman Moscow State Technical University. Series Instrument Engineering*. 2014;(5):88–97. (In Russ.).
 9. Vakman D. E., Sedletskiy R. M. Voprosy sinteza radiolokatsionnykh signalov. Moscow: Sov. radio, 1973. 312 p.

Информация об авторах

А. И. Дёмко – кандидат технических наук, доцент, почетный работник высшего образования РФ.

Е. Г. Алмазова – старший преподаватель.

About the authors

A. I. Dyomko – Candidate of Sciences (Engineering), Docent, Honored Worker of Higher Education of the Russian Federation.

E. G. Almazova – Senior Lecturer.

Сетевое издание. Полные тексты статей размещаются на официальном сайте издания <https://www.vestcyber.ru> и в базе данных Научной электронной библиотеки на сайте elibrary.ru, сведения о публикуемых материалах включаются в Российский индекс научного цитирования (РИНЦ).

Адрес учредителя и издателя:
бюджетное учреждение высшего образования Ханты-Мансийского автономного округа – Югры
«Сургутский государственный университет»,
628412, Ханты-Мансийский автономный округ – Югра, Сургут, пр. Ленина, 1. Тел. +7 (3462) 76-29-29

Дата опубликования 31.10.2024.