

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

ИНФОРМАТИКА

INFORMATICS

TOM 19
VOL.

4 | 2022

ОТ РЕДАКЦИИ

В журнале «Информатика» публикуются оригинальные и обзорные статьи, описывающие результаты фундаментальных и прикладных исследований специалистов академического и вузовского профиля в области информатики и информационных технологий.

Основной целью журнала является публикация наиболее значимых новых результатов в указанной области. Приветствуются статьи, описывающие заключительные результаты научных проектов и диссертационных исследований, открывающие новые направления исследований, которые находятся на стыке информатики и других наук.

Журнал рассчитан на широкий круг специалистов в области информатики и информационных технологий.

Основные разделы журнала:

- биоинформатика;
- математическое моделирование;
- защита информации и надежность систем;
- информационные технологии;
- логическое проектирование;
- обработка сигналов, изображений, речи, текста и распознавание образов;
- автоматизация проектирования;
- интеллектуальные системы.

Префикс DOI: 10.37661

Условия распространения материалов:

контент доступен под лицензией Creative Commons Attribution 4.0 License

Индексирование:

Высшей аттестационной комиссией Республики Беларусь журнал «Информатика» был включен в список научных изданий для опубликования результатов диссертационных исследований.

В декабре 2017 г. включен в базу данных Российского индекса научного цитирования (РИНЦ). С помощью инструментов и сервисов, доступных на платформе eLIBRARY (раздел «Личный кабинет»), можно самостоятельно корректировать список своих публикаций и цитирований в РИНЦ.

В июле 2017 г. включен в базу журналов открытого доступа Directory of Open Access Journals (DOAJ).

С помощью поисковых систем Google Scholar, WorldCat, Соционет можно получить свободный доступ к полному тексту научных публикаций журнала.

Адрес редакции:

ул. Сурганова, 6, к. 305, г. Минск, 220012, Беларусь
Тел. +375 (017) 351 26 22

Editorial address:

Surganova str., 6, of. 305, Minsk, 220012, Belarus
Phone +375 (017) 351 26 22

E-mail: rio@newman.bas-net.by

<https://inf.grid.by/jour>

THE EDITOR'S NOTE

The journal «Informatics» is a scientific publication in computer sciences and information technologies which reviews the results in basic and applied research of scientists from the universities and scientific centers.

The journal focuses on the most significant and modern papers of research projects results and PhD/DSc thesis in computer sciences.

The journal is edited for the specialists in IT and computer sciences research and application.

The main sections of the journal:

- bioinformatics;
- mathematical modeling;
- information protection and system reliability;
- information technology;
- logical design;
- signal, image, speech, text processing and pattern recognition;
- computer-aided design;
- artificial intelligence methods.

DOI Prefix: 10.37661

Distribution:

content is distributed under Creative Commons Attribution 4.0 License

Indexation:

the journal «Informatics» is in the list of scientific publications recommended by the Higher Attestation Commission of the Republic of Belarus for scientists to publish the results of PhD/DSc research.

In December 2017 the journal was included in the database of the Russian Science Citation Index (RISC) and provides free access to reviewed electronic scientific paper, improving scientific information traffic and also raising quotation of works of the authors (please use <https://elibrary.ru> or section https://elibrary.ru_author_tools).

In July 2017 included in the database of open access journals Directory of Open Access Journals (DOAJ).

Using the Google Scholar, WorldCat, Соционет search engine, you can get free access to full text of scientific publications of magazine.

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

ИНФОРМАТИКА

Informatika

Том 19, № 4, октябрь-декабрь 2022

Ежеквартальный научный журнал

Издается с января 2004 г.

Учредитель и издатель – Объединенный институт проблем информатики
Национальной академии наук Беларуси (ОИПИ НАН Беларуси)

Г л а в н ы й р е д а к т о р

Тузиков Александр Васильевич, д-р физ.-мат. наук, проф., чл.-корр. НАН Беларуси,
ОИПИ НАН Беларуси (Минск, Беларусь)

З а м е с т и т е л ь г л а в н о г о р е д а к т о р а

Ковалев Михаил Яковлевич, д-р физ.-мат. наук, проф., чл.-корр. НАН Беларуси,
ОИПИ НАН Беларуси (Минск, Беларусь)

Р е д а к ц и о н н а я к о л л е г и я

Абламейко Сергей Владимирович, д-р техн. наук, проф., академик НАН Беларуси, БГУ (Минск, Беларусь)

Анищенко Владимир Викторович, канд. техн. наук, доцент, ООО «СофтКлуб» (Минск, Беларусь)

Бибило Петр Николаевич, д-р техн. наук, проф., ОИПИ НАН Беларуси (Минск, Беларусь)

Бобов Михаил Никитич, д-р техн. наук, проф., БГУИР (Минск, Беларусь)

Долгий Александр Борисович, д-р техн. наук, проф., Высшая инженерная школа Бретани (Нант, Франция)

Дудин Александр Николаевич, д-р физ.-мат. наук, проф., БГУ (Минск, Беларусь)

Карпов Алексей Анатольевич, д-р техн. наук, доцент, СПИИРАН (Санкт-Петербург, Россия)

Килин Сергей Яковлевич, д-р физ.-мат. наук, проф., академик НАН Беларуси, Президиум НАН
Беларуси (Минск, Беларусь)

Краснопрошин Виктор Владимирович, д-р техн. наук, проф., БГУ (Минск, Беларусь)

Крот Александр Михайлович, д-р техн. наук, проф., ОИПИ НАН Беларуси (Минск, Беларусь)

Кругликов Сергей Владимирович, д-р воен. наук, канд. техн. наук, доцент, ОИПИ НАН Беларуси
(Минск, Беларусь)

Лиходед Николай Александрович, д-р физ.-мат. наук, проф., БГУ (Минск, Беларусь)

Матус Петр Павлович, д-р физ.-мат. наук, проф., Институт математики НАН Беларуси (Минск, Беларусь)

Скляр Валерий Анатольевич, д-р техн. наук, проф., Университет Авейру (Авейру, Португалия)

Сотсков Юрий Назарович, д-р физ.-мат. наук, проф., ОИПИ НАН Беларуси (Минск, Беларусь)

Стемпковский Александр Леонидович, д-р техн. наук, проф., академик РАН, ИПИМ РАН (Москва, Россия)

Харин Юрий Семенович, д-р физ.-мат. наук, проф., академик НАН Беларуси, НИИ ППМИ БГУ
(Минск, Беларусь)

Чернявский Александр Федорович, д-р техн. наук, проф., академик НАН Беларуси, НИИ ПФП
им. А. Н. Севченко БГУ (Минск, Беларусь)

Ярмолик Вячеслав Николаевич, д-р техн. наук, проф., БГУИР (Минск, Беларусь)

Редакционный совет

Ефанов Дмитрий Викторович, Российский университет транспорта (Московский институт инженеров транспорта) (Москва, Россия)

Кумари Мадху, Университетский центр исследований и разработок, Университет Чандигарха (Мохали, Пенджаб, Индия)

Лазарев Александр Алексеевич, Институт проблем управления им. В. А. Трапезникова РАН (Москва, Россия)

Лай Цунг-Чьян, Азиатский университет в Тайчжуне (Китайская Народная Республика, Тайвань)

Марина Нинослав, Университет информационных наук и технологий им. Св. апостола Павла (Охрид, Македония)

Меликян Вазген Шаваршович, Национальный политехнический университет Армении (Ереван, Армения)

Пеш Эрвин, Зигенский университет (Зиген, Германия)

Сингх Таджиндер, Институт инженерии и технологий Сант Лонговал (Лонговал, Пенджаб, Индия)

Ходаченко Максим Леонидович, Институт космических исследований Австрийской академии наук (Грац, Австрия)

Чиулла Карло, Университет Эпока (Тирана, Албания)

Штейнберг Борис Яковлевич, Институт математики, механики и компьютерных наук Южного федерального университета (Ростов-на-Дону, Россия)

ИНФОРМАТИКА

Том 19, № 4, октябрь-декабрь 2022

Ответственный за выпуск *Мойсейчик Светлана Сергеевна*

Редактор *Гончаренко Галина Борисовна*

Корректор *Михайлова Анна Антоновна*

Компьютерная верстка *Бутевич Ольга Борисовна*

Сдано в набор 15.11.2022. Подписано в печать 19.12.2022. Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс. Ризография. Усл. печ. л. 12,8. Уч.-изд. л. 12,5. Тираж 40 экз. Заказ 10.

Государственное научное учреждение «Объединенный институт проблем информатики
Национальной академии наук Беларуси».

Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 1/274 от 04.04.2014. ЛП № 02330/444 от 18.12.13. Ул. Сурганова, 6, 220012, Минск, Беларусь.

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

THE UNITED INSTITUTE OF INFORMATICS PROBLEMS
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

INFORMATICS

Vol. 19, no. 4, October-December 2022

Published quarterly

Issued since January 2004

Founder and publisher – the United Institute of Informatics Problems
of the National Academy of Sciences of Belarus (UIIP NASB)

Editor-in-Chief

Alexander V. Tuzikov, D. Sc. (Phys.-Math.), Prof., Corr. Member of NASB,
UIIP NASB (Minsk, Belarus)

Deputy Editor-in-Chief

Mikhail Y. Kovalyov, D. Sc. (Phys.-Math.), Prof., Corr. Member of NASB,
UIIP NASB (Minsk, Belarus)

Editorial Board

Sergey V. Ablameyko, D. Sc. (Eng.), Prof., Academician of NASB, BSU (Minsk, Belarus)

Uladimir V. Anishchanka, Ph. D. (Eng.), Assoc. Prof., SoftClub Ltd. (Minsk, Belarus)

Petr N. Bibilo, D. Sc. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Mikhail N. Bobov, D. Sc. (Eng.), Prof., BSUIR (Minsk, Belarus)

Alexandre B. Dolgui, D. Sc. (Eng.), Prof., IMT Atlantique (Nantes, France)

Alexander N. Dudin, D. Sc. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

Alexey A. Karpov, D. Sc. (Eng.), Assoc. Prof., SPII RAS (Saint Petersburg, Russia)

Sergey Ya. Kilin, D. Sc. (Phys.-Math.), Prof., Academician of NASB, Presidium of NASB (Minsk, Belarus)

Viktor V. Krasnoproshin, D. Sc. (Eng.), Prof., BSU (Minsk, Belarus)

Alexander M. Krot, D. Sc. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Sergey V. Kruglikov, D. Sc. (Milit.), Ph. D. (Eng.), Assoc. Prof., UIIP NASB (Minsk, Belarus)

Nikolai A. Likhoded, D. Sc. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

Petr P. Matus, D. Sc. (Phys.-Math.), Prof., Institute of Mathematics of NASB (Minsk, Belarus)

Valery A. Sklyarov, D. Sc. (Eng.), Prof., University of Aveiro (Aveiro, Portugal)

Yuri N. Sotskov, D. Sc. (Phys.-Math.), Prof., UIIP NASB (Minsk, Belarus)

Alexander L. Stempkovsky, D. Sc. (Eng.), Prof., Academician of RAS, IPPM RAS (Moscow, Russia)

Yuriy S. Kharin, D. Sc. (Phys.-Math.), Prof., Academician of NASB, RI APMI BSU (Minsk, Belarus)

Alexander F. Cherniavsky, D. Sc. (Eng.), Prof., Academician of NASB, A. N. Sevchenko IAPP BSU (Minsk, Belarus)

Vyacheslav N. Yarmolik, D. Sc. (Eng.), Prof., BSUIR (Minsk, Belarus)

Editorial Council

Dmitry V. Efanov, Russian University of Transport (Moscow Institute of Transport Engineers) (Moscow, Russia)

Madhu Kumari, University Center for Research & Development, Chandigarh University (Mohali, Punjab, India)

Alexander A. Lazarev, V. A. Trapeznikov Institute of Control Sciences of the RAS (Moscow, Russia)

Tsung-Chyan Lai, Asia University at Taichung (The People's Republic of China, Taiwan)

Ninoslav Marina, St. Paul the Apostle University of Information Sciences and Technology (Ohrid, Macedonia)

Vazgen Sh. Melikyan, National Polytechnic University of Armenia (Yerevan, Armenia)

Erwin Pesch, University of Siegen (Siegen, Germany)

Tajinder Singh, Sant Longowal Institute of Engineering & Technology (Longowal, Punjab, India)

Maxim L. Khodachenko, Space Research Institute, Austrian Academy of Sciences (Graz, Austria)

Carlo Ciulla, Epoka University (Tirana, Albania)

Boris Steinberg, Institute of Mathematics, Mechanics and Computer Science Southern Federal University (Rostov-on-Don, Russia)

INFORMATICS

Vol. 19, no. 4, October-December 2022

Issue Head *Sviatlana S. Maiseichyk*

Editor *Halina B. Hancharenka*

Corrector *Hanna A. Mikhailava*

Computer Imposition *Volha B. Butsevich*

Sent for press 15.11.2022. Output 19.12.2022. Format 60×84 1/8. Offset paper. Headset Times. Riesography. Printed sheets 12,8. Publisher's signatures 12,5. Circulation 40 copies. Order 10.

State Scientific Institution "The United Institute of Informatics Problems of the National Academy of Sciences of Belarus".

Certificate on the state registration of the publisher, manufacturer, distributor of printing editions no. 1/274 dated 04.04.2014. License for the press no. 02330/444 dated 18.12.13.

6, Surganov Str., 220012, Minsk, Belarus.

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

СОДЕРЖАНИЕ

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

Ярмолик В. Н., Петровская В. В., Мрозек И. Мера различия для тестовых наборов при генерировании управляемых вероятностных тестов 7

ЗАЩИТА ИНФОРМАЦИИ И НАДЕЖНОСТЬ СИСТЕМ

Шамына А. Ю., Иванюк А. А. Построение и балансировка путей физически неклонированной функции типа арбитр на FPGA 27

Бойправ В. А., Утин Л. Л. Методика и программное средство для проведения аудита систем менеджмента информационной безопасности 42

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

Семенюк В. В., Складчиков М. В. Разработка алгоритма распознавания эмоций человека с использованием сверточной нейронной сети на основе аудиоданных 53

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ

Петухов А. В. Информационное обеспечение САПР технологических процессов опытных образцов структурно-сложных изделий 69

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Марков А. Н. Модель балансировки нагрузки кластерной системы с учетом аппаратных характеристик серверного оборудования 84

Лукашевич М. Н., Ковалев М. Я. Модели и методы машинного обучения для решения задач оптимизации и прогнозирования работы морских портов 94

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

CONTENTS

LOGICAL DESIGN

Yarmolik V. N., Petrovskaya V. V., Mrozek I. A measure of the difference between test sets for generating controlled random tests 7

INFORMATION PROTECTION AND SYSTEM RELIABILITY

Shamyna A. Yu., Ivaniuk A. A. Creating and balancing the paths of arbiter-based physically unclonable functions on FPGA 27

Boiprav V. A., Utin L. L. Methodology and software development for auditing information security management systems 42

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION

Semenuk V. V., Skladchikov M. V. Algorithm development for recognizing human emotions using a convolutional neural network based on audio data 53

COMPUTER-AIDED DESIGN

Petukhov A. V. Information support of CAD of technological processes of structurally complex product prototypes 69

INFORMATION TECHNOLOGIES

Markov A. N. Cluster system load balancing model with consideration of hardware characteristics of server hardware 84

Lukashevich M. N., Kovalyov M. Y. Machine learning models and methods for solving optimization and forecasting problems of the work of seaports 94

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

LOGICAL DESIGN



УДК 004.33.054
<https://doi.org/10.37661/1816-0301-2022-19-4-7-26>

Оригинальная статья
Original Paper

Мера различия для тестовых наборов при генерировании управляемых вероятностных тестов

В. Н. Ярмолик^{1✉}, В. В. Петровская¹, И. Мрозек²

¹Белорусский государственный университет информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь
✉E-mail: yarmolik10ru@yahoo.com

²Белостоцкий технический университет,
ул. Вейска, 45А, 15-351, Белосток, Польша

Аннотация

Цели. Решается задача построения характеристик различия тестовых наборов, представляющих собой наборы символов, включая двоичные наборы. Обосновывается ее актуальность для генерирования управляемых вероятностных тестов и сложность нахождения мер различия для символьных тестов. Показывается ограниченность применения расстояния Хэмминга и Дамерау – Левенштейна для получения меры различия тестовых наборов.

Методы. На основе характеристики интервала, применяемого в теории строя цепи последовательных событий, определяется новая мера различия двух символьных тестовых наборов. В качестве меры различия рассчитывается расстояние $AD(T_i, T_k)$ между тестовыми наборами T_i и T_k , использующее характеристику интервала и основанное на определении независимых пар одинаковых (тождественных) символов, принадлежащих двум наборам, и вычислении интервалов между ними.

Результаты. Показывается комбинаторный характер вычисления предложенной меры различия для символьных тестовых наборов произвольного алфавита и размерности. Приводится пример вычисления данной меры для различных видов тестовых наборов, в том числе таких, как адресные тестовые наборы. Показываются возможные ее модификации и определяются некоторые свойства и ограничения. Рассматривается применение данной меры различия для случая многократного тестирования запоминающих устройств на основе адресных последовательностей pA с четным p повторением адресов. Для случая $p = 2$ приводятся математические соотношения вычисления интервалов и расстояния $AD(T_i, T_k)$ для последовательностей адресов $2A$, используемых для управляемого вероятностного тестирования запоминающих устройств. Основное внимание уделяется двоичным тестовым наборам, для которых задача вычисления данной метрики различия сводится к классической задаче о назначениях с использованием венгерского алгоритма. Вычислительная сложность венгерского алгоритма оценивается соотношением $O(n^4)$. Как альтернатива венгерскому алгоритму предлагается алгоритм вычисления рассматриваемой меры, сложность которого существенно меньше и имеет оценку $O(n^2)$. Проведенные экспериментальные исследования подтверждают эффективность рассмотренного алгоритма.

Заключение. Предложенная мера различия расширяет возможности генерирования тестовых последовательностей при генерировании управляемых вероятностных тестов. Показано, что тестовые наборы, неразличимые при использовании в качестве меры различия расстояния Хэмминга, имеют различные значения $AD(T_i, T_k)$, позволяющие более точно классифицировать формируемые случайным образом наборы, которые являются кандидатами в тестовые наборы.

Ключевые слова: управляемые вероятностные тесты, мера различия символьных наборов, адресные последовательности, задача о назначениях, венгерский алгоритм

Благодарности. Авторы выражают искреннюю благодарность старшему преподавателю кафедры программного обеспечения информационных технологий БГУИР Н. С. Петюкевич за участие в обсуждении результатов статьи, советы и рекомендации.

Для цитирования. Ярмолик, В. Н. Мера различия для тестовых наборов при генерировании управляемых вероятностных тестов / В. Н. Ярмолик, В. В. Петровская, И. Мрозек // Информатика. – 2022. – Т. 19, № 4. – С. 7–26. <https://doi.org/10.37661/1816-0301-2022-19-4-7-26>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 28.07.2022

Подписана в печать | Accepted 02.09.2022

Опубликована | Published 29.12.2022

A measure of the difference between test sets for generating controlled random tests

Vyacheslav N. Yarmolik¹✉, Vita V. Petrovskaya¹, Ireneusz Mrozek²

¹*Belarusian State University of Informatics and Radioelectronics,
st. P. Brovki, 6, Minsk, 220013, Belarus*

✉*E-mail: yarmolik10ru@yahoo.com*

²*Bialystok University of Technology,
Wiejska, 45A, 15-351, Białystok, Poland*

Abstract

Objectives. The problem of constructing the characteristics of the difference between test sequences is solved. Its relevance for generating controlled random tests and the complexity of finding measures of difference for symbolic tests are substantiated. The limitations of using the Hamming and Damerau – Levenshtein distances to obtain a measure of the difference between test sets are shown.

Methods. Based on the characteristic of the interval used in the theory of the chain of successive events, a new measure of the difference between two symbolic test sets is determined. As a difference measure, the distance $AD(T_i, T_k)$ between the test sets T_i and T_k is calculated using the interval characteristic, which is based on determining independent pairs of same (identical) symbols belonging to two sets and calculating the intervals between them.

Results. The combinatorial nature of the calculation of the proposed difference measure for symbolic test sets of an arbitrary alphabet and dimension is shown. An example of calculating this measure for various types of test sets, including such as address test sets, is given. Possible modifications are shown and some properties and limitations are determined. The application of the measure of difference is considered for the case of repeated testing of storage devices based on address sequences pA with even p repetition of addresses. For the case $p = 2$, mathematical relations are given for calculating the intervals and distances $AD(T_i, T_k)$ for address sequences $2A$ used for controlled random testing of storage devices. The main attention is paid to binary test sets, when the task of calculating given difference metric is reduced to the classical assignment problem using the Hungarian algorithm. The computational complexity of the Hungarian algorithm is estimated by the relation $O(n^4)$. As an alternative to the Hungarian algorithm, an algorithm for calculating the considered difference measure is proposed, the complexity of which is much less and has an estimate equal to $O(n^2)$. The experimental studies confirm the effectiveness of the proposed algorithm.

Conclusion. The proposed difference measure extends the possibilities of generating test sequences when generating controlled random tests. It is shown that test sets, which are indistinguishable when Hamming distance is used as a measure of difference, have different values of $AD(T_i, T_k)$ that allows to make more accurate classification of randomly generated sets as candidates for test sets.

Keywords: controlled random tests, character set difference measure, address sequences, assignment problem, Hungarian algorithm

Acknowledgements. The authors express their sincere gratitude to N. S. Petyukevich, Senior Lecturer of the Department of Information Technology Software at BSUIR, for participating in the discussion of the results of the article, advice and recommendations.

For citation. Yarmolik V. N., Petrovskaya V. V., Mrozek I. *A measure of the difference between test sets for generating controlled random tests*. Informatika [Informatics], 2022, vol. 19, no. 4, pp. 7–26 (In Russ.). <https://doi.org/10.37661/1816-0301-2022-19-4-7-26>

Conflict of interest. The authors declare of no conflict of interest.

Введение. Фундаментальный подход к тестированию современных вычислительных систем включает в себя построение тестовых наборов случайным образом из всех возможных входных данных объекта тестирования. Такой подход называется *вероятностным тестированием* (Random Testing) [1, 2]. Вероятностное тестирование зачастую бывает единственно возможным подходом к тестированию не только на стадии эксплуатационного тестирования, где оценивается надежность вычислительной системы, но и на стадии отладочного и других видов тестирования. Вероятностный тест задается количеством q тестовых наборов $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,n-1}$, $i \in \{0, 1, \dots, q-1\}$, данных $t_{i,j}$, $j \in \{0, 1, \dots, n-1\}$, их множеством, определяемым заданным алфавитом данных и числом n в наборе, а также законом распределения наборов.

Для повышения эффективности вероятностных тестов были предложены различные их модификации, которые получили общее название *управляемые (адаптивные) вероятностные тесты* (Adaptive Random Tests) [1–3]. Основная предпосылка появления подобных тестов заключается в том, что в целях достижения более высокого покрытия неисправностей, обнаруживаемых тестом, и минимизации его временной сложности необходимо целенаправленно выбирать очередной тестовый набор в зависимости от ранее сгенерированных наборов. Существует большое разнообразие как принципов генерирования подобных тестов, так и учета специфики объекта тестирования при формировании тестовых наборов для управляемых вероятностных тестов [3–7]. Под управляемыми вероятностными тестами понимают вероятностные тесты, в которых очередной тестовый набор формируется с учетом ранее сгенерированных наборов и которые формально определяются следующим образом [3, 4, 7]: управляемым вероятностным является тест, состоящий из q наборов $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,n-1}$ тестовых данных $t_{i,j}$, $j \in \{0, 1, \dots, n-1\}$, сгенерированных случайным образом так, что очередной тестовый набор T_i удовлетворяет заданным критериям, полученным на основании ранее сгенерированных наборов T_0, T_1, \dots, T_{i-1} .

Согласно общепринятому представлению об управляемых вероятностных тестах ключевой особенностью генерирования тестовых наборов T_i является информация, которая извлекается в виде некоторых характеристик (метрик) из ранее сгенерированных тестовых наборов и используется для формирования очередного набора [3, 7]. Основная идея управляемых вероятностных тестов заключается в том, что очередной тестовый набор T_i формируется максимально удаленным (высокой степени различия) от ранее сгенерированных наборов T_0, T_1, \dots, T_{i-1} в терминах заранее определенных мер различия. Таким образом, принимается гипотеза, что для двух тестовых наборов T_i и T_k , имеющих минимальное различие, количество обнаруживаемых неисправностей будет минимальным и, наоборот, для максимально различных тестовых наборов обнаруживающая способность будет максимальной [3, 7]. В качестве меры различия тестового набора T_i от предыдущих наборов $T_0, T_1, T_2, \dots, T_{i-1}$ чаще всего используются *расстояние Хэмминга* и *расстояние Евклида* [3–9]. Эти характеристики наиболее эффективны для двоичного случая тестовых данных и малопродуктивны для произвольных данных [10, 11].

Проблема сравнения тестовых наборов, которые в общем случае представляют собой символьные последовательности, актуальна для различных областей науки. В работе [10] показано, что в пространстве символьных последовательностей сложно ввести метрику различия. Формально подобная метрика в таком пространстве существует – это *расстояние Хэмминга* [3, 10, 11]. Отметим, что расстояние Хэмминга $HD(T_i, T_k)$ между двумя наборами T_i и T_k равняется числу их несовпадающих компонент (символов) $t_{i,j}$ и $t_{k,j}$. В ряде работ отмечалось, что эта метрика малоэффективна, так как позволяет различать лишь полностью совпадающие последовательности

при $HD(T_i, T_k) = n$ и все остальные несовпадающие [10–13]. В настоящее время чаще всего обсуждается метод сравнения символьных последовательностей, основанный на их выравнивании (метод *редакционного расстояния*), и его модификации [10]. Наиболее известными из них являются методы, использующие *расстояние Дамерау – Левенштейна* (Damerau – Levenshtein distance) [14]. Они основаны на сравнении одной последовательности символов T_i с другой T_k с помощью операций вставки, замены (либо удаления) и транспозиции так, чтобы эти две последовательности совпали. Недостатком методов выравнивания является необходимость назначения и обоснования системы штрафных (весовых) функций и выбора опорной последовательности, относительно которой проводится выравнивание. Кроме того, их эффективность снижается экспоненциально при увеличении длины строк символов [10, 14]. Оценку схожести (подобия) символьных последовательностей можно получить, используя *сходство Джаро – Винклера* (Jaro – Winkler distance), представляющее собой меру схожести строк для измерения расстояния между двумя последовательностями символов [15]. Чем меньше расстояние Джаро – Винклера для двух строк, тем больше сходства между этими строками. Данная мера расстояния чаще всего применяется для оценки схожести наборов символов, а не их различия и также характеризуется достаточно большой вычислительной сложностью.

Основной проблемой управляемых вероятностных тестов является их большая вычислительная сложность, связанная с необходимостью перечисления возможных тестовых наборов и вычисления меры различия для каждого потенциального кандидата в тесты [10–13]. Чаще всего при формировании очередного тестового набора первоначально генерируется достаточное количество кандидатов в тесты (ориентировочно в диапазоне от 10 до 100), представляющих собой равномерно распределенные случайные наборы. Для каждого из них вычисляются метрики различия, с учетом которых и выбирается наилучший из кандидатов в тесты в качестве очередного тестового набора T_i .

Таким образом, главная задача управляемого вероятностного тестирования состоит в нахождении меры различия для тестовых наборов T_i и T_k , которая максимально адекватно показывает их различия и характеризуется невысокой вычислительной сложностью. Вычисление мер различия тестовых наборов, в общем случае представляющих собой символьные последовательности, в свою очередь, сводится к задаче их сравнения [10].

1. Мера различия для построения управляемых вероятностных тестов. Рассматриваемая мера различия основана на *теории строя*, которая предназначена для формального описания и анализа последовательностей данных (символов) любой природы [16]. В общем случае строй может представлять собой последовательность символов (*цепь событий*) любого алфавита, произвольной структуры и длины, а его анализ может осуществляться по различным правилам. Однако основной числовой характеристикой строя является *интервал*, определяемый как расстояние от выделенного в строе символа до другого ближайшего, отмеченного в направлении просмотра такого же символа. Содержание характеристики интервала весьма близко к *числу транспозиций*, используемому для вычисления расстояния Джаро – Винклера [15].

Характеристика интервала, которая служит основой для оценки различных свойств последовательностей символов, была использована для определения меры различия (степени несовпадения) двух тестовых наборов, показывая их удаленность (либо близость) друг от друга [11]. В общем случае для тестовых наборов T_i и T_k , каждый из которых состоит из n_i и n_k данных $t_{i,j}$, $j \in \{0, 1, \dots, n_i - 1\}$, и $t_{k,r}$, $r \in \{0, 1, \dots, n_k - 1\}$, интервалом для пары $t_{i,j} = t_{k,r}$ является значение $D(t_{i,j}, t_{k,r})$, которое в дальнейшем будем отождествлять с расстоянием между данными $t_{i,j}$ и $t_{k,r}$. Предполагая, что определяется расстояние (интервал) между $t_{i,j}$ и $t_{k,r}$, первоначально вычисляются значения $|j - r|$ и $\max(n_i, n_k) - |j - r|$. Минимальное $\min[|j - r|, \max(n_i, n_k) - |j - r|]$ из приведенных значений принимается в качестве расстояния $D(t_{i,j}, t_{k,r})$ между $t_{i,j}$ и $t_{k,r}$.

Как отмечалось ранее, подобная оценка расстояния необходима для синтеза управляемых вероятностных тестов, когда очередной тестовый набор формируется максимально удаленным от ранее сгенерированных наборов. Формально эта характеристика описана в работе [11] и соответствует следующему определению.

Определение 1. Мера различия $AD(T_i, T_k)$ тестовых наборов T_i и T_k , каждый из которых состоит из n_i и n_k данных $t_{i,j}, j \in \{0, 1, \dots, n_i - 1\}$, и $t_{k,r}, r \in \{0, 1, \dots, n_k - 1\}$, использующая характеристику интервала, основана на определении независимых пар одинаковых данных, принадлежащих двум наборам. Независимость пар означает участие каждого значения данных $t_{i,j}$ и $t_{k,r}$ тестовых наборов T_i и T_k только в одной паре. Процедура формирования подобных пар носит комбинаторный характер и заключается в нахождении такого их сочетания, для которого сумма расстояний $D(t_{i,j}, t_{k,r})$ также минимальна. При отсутствии пары для очередного значения данных $t_{i,j}$ в наборе T_k разность величин индексов, т. е. расстояние $D(t_{i,j}, -)$, принимается равным $\min(n_i, n_k)$. Такое же значение расстояний задается и для данных набора T_k , для которых отсутствует пара в T_i .

Минимальное и максимальное значения меры различия $AD(T_i, T_k)$ зависят от алфавита данных $t_{i,j}$ и $t_{k,r}$ и соотношения их значений в тестовых наборах T_i и T_k , а также размерности n_i и n_k сравниваемых наборов. Минимальное значение метрики $AD(T_i, T_k)$ определяется выражением $\min AD(T_i, T_k) = |n_i - n_k| \cdot \min(n_i, n_k)$, что свидетельствует о максимальном сходстве двух наборов T_i и T_k . При выполнении равенства $n_i = n_k$ значение $\min AD(T_i, T_k)$ равняется нулю, что свидетельствует о тождественности сравниваемых наборов. Максимальное значение $AD(T_i, T_k)$ определяется как $\max AD(T_i, T_k) = (n_i + n_k) \cdot \min(n_i, n_k)$, что свидетельствует о полном (максимальном) отличии наборов T_i и T_k . Чем ближе численное значение $AD(T_i, T_k)$ к его максимальной величине $\max AD(T_i, T_k)$, тем больше степень различия между тестовыми наборами T_i и T_k , и, наоборот, близость этой характеристики к $\min AD(T_i, T_k)$ говорит о их максимальном сходстве. В качестве примера рассмотрим тестовые данные, представляющие собой адресные тестовые наборы, удовлетворяющие следующему определению [7, 17].

Определение 2. Под тестовым набором адресов $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,n-1}$ (последовательностью адресов A) понимают последовательность из $n = 2^m$ m -битовых векторов $t_{i,j} \in \{0, 1, \dots, 2^m - 1\}$, $j \in \{0, 1, \dots, 2^m - 1\}$, каждый из которых принимает одно из 2^m возможных значений.

Приведенное выше определение меры различия позволяет сравнить два тестовых набора адресов T_i и T_k . В качестве примера набора T_i рассмотрим счетчиковую последовательность для $m = 3$, а в качестве T_k используем последовательность отраженного кода Грея (табл. 1) [18].

Таблица 1
Примеры адресных тестовых последовательностей

Table 1
Examples of address test sequences

T_i	j	0	1	2	3	4	5	6	7
	$t_{i,j}$	000	001	010	011	100	101	110	111
T_k	r	0	1	2	3	4	5	6	7
	$t_{k,r}$	000	001	011	010	110	111	101	100

Анализ тестовых наборов, приведенных в табл. 1, показывает их одинаковую размерность $n_i = n_k = n = 2^m = 8$ и структуру, характеризующуюся наличием в каждом наборе по одному трехбитовому вектору из восьми возможных. Таким образом, для каждого данного в тестовых последовательностях адресов T_i и T_k существует пара, и она единственная. Для $t_{i,0} = 000$ набора T_i тождественное значение 000 в наборе T_k имеет тот же индекс $r = j = 0$. Соответственно, расстояние $D(t_{i,0}, t_{k,0})$, равное разности их индексов, принимает значение 0, т. е. $D(t_{i,0}, t_{k,0}) = 0$. Аналогично и для 001 имеем $D(t_{i,1}, t_{k,1}) = 0$. Для $t_{i,2} = t_{k,3} = 010$ существуют уже два значения положительных разностей индексов $|2 - 3| = 1$ и $8 - |2 - 3| = 7$, минимальное из которых и определяет расстояние $D(t_{i,2}, t_{k,3}) = 1$. Продолжая пример, приведенный в табл. 1, получим $D(t_{i,3}, t_{k,2}) = \min(|3 - 2|, 8 - |3 - 2|) = 1$. Далее имеем $D(t_{i,4}, t_{k,7}) = \min(|4 - 7|, 8 - |4 - 7|) = 3$; $D(t_{i,5}, t_{k,6}) = \min(|5 - 6|, 8 - |5 - 6|) = 1$; $D(t_{i,6}, t_{k,4}) = \min(|6 - 4|, 8 - |6 - 4|) = 2$ и $D(t_{i,7}, t_{k,5}) = \min(|7 - 5|, 8 - |7 - 5|) = 2$. Окончательно для рассматриваемого примера мера различия $AD(T_i, T_k)$ вычисляется как $AD(T_i, T_k) = D(t_{i,0}, t_{k,0}) + D(t_{i,1}, t_{k,1}) + D(t_{i,2}, t_{k,3}) + D(t_{i,3}, t_{k,2}) + D(t_{i,4}, t_{k,7}) + D(t_{i,5}, t_{k,6}) + D(t_{i,6}, t_{k,4}) + D(t_{i,7}, t_{k,5}) = 0 + 0 + 1 + 1 + 3 + 1 + 2 + 2 = 10$.

Мера различия $AD(T_i, T_k)$ для двух произвольных T_i и T_k тестовых адресных последовательностей, удовлетворяющих определению 2, вычисляется согласно соотношению [11]

$$AD(T_i, T_k) = \sum_{j=0}^{2^m-1} \sum_{r=0}^{2^m-1} I(t_{i,j} = t_{k,r}) \cdot \min[|j-r|, 2^m - |j-r|]. \quad (1)$$

Выражение $I(t_{i,j} = t_{k,r})$ представляет собой индикаторную функцию, равную единице при $t_{i,j} = t_{k,r}$ и нулю в противном случае. Минимальное значение $\min AD(T_i, T_k) = 0$ при совпадении последовательностей T_i и T_k , а максимальное значение $\max AD(T_i, T_k) = 2^{2^m-1}$ для случая максимальных величин $\min(|j-r|, 2^m - |j-r|) = 2^{m-1}$ для всех данных в сравниваемых последовательностях.

В ряде случаев формирование тестового набора адресов T_k , функционально зависящего от набора T_i , позволяет аналитически вычислять значение меры различия $AD(T_i, T_k)$. Например, в случае генерирования набора $T_k = T_i(s)$ – циклически сдвинутой вправо на $s \in \{0, 1, \dots, 2^m-1\}$ позиций копии исходного набора T_i – существенно упрощается соотношение (1). Соответственно, получим равенство

$$AD(T_i, T_i(s)) = \begin{cases} s \cdot 2^m, & s = 0, 1, \dots, 2^m-1; \\ (2^{m-1} - s \bmod 2^{m-1}) \cdot 2^m, & s = 2^{m-1} + 1, 2^{m-1} + 2, \dots, 2^m - 1. \end{cases} \quad (2)$$

Очевидно, что $\min AD(T_i, T_i(s)) = 0$ при $s = 0$ и $\max AD(T_i, T_i(s)) = 2^{2^m-1}$ при $s = 2^{m-1}$.

Рассматриваемая мера различия основана на определении независимых пар тождественных данных, принадлежащих наборам T_i и T_k . Структура адресных тестовых наборов, соответствующих определению 2, упрощает задачу нахождения независимых пар, так как для каждого $t_{i,j}$ набора T_i всегда существует единственное значение $t_{k,r} = t_{i,j}$ в наборе T_k , где $j, r \in \{0, 1, \dots, 2^m-1\}$, что и определяет независимость пар тождественных данных.

Новым развитием адресных последовательностей являются последовательности pA для произвольного четного p , удовлетворяющие следующему определению [19, 20].

Определение 3. Под тестовым набором адресов $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,p-1}$ (последовательностью pA) понимают последовательность из $n = p \cdot 2^m$ m -битовых векторов $t_{i,j} \in \{0, 1, \dots, 2^m-1\}$, $j \in \{0, 1, \dots, p \cdot 2^m-1\}$, каждый из которых формируется ровно p раз.

Для случая адресных наборов T_i и T_k , соответствующих определению 3, вычисление метрики $AD(T_i, T_k)$ сопряжено с нахождением минимальной суммы минимальных расстояний для всех сочетаний p независимых пар одинаковых значений $t_{i,j} = t_{k,r}$ адресов $t_{i,j}, t_{k,r} \in \{0, 1, \dots, 2^m-1\}$; $j, r \in \{0, 1, \dots, p \cdot 2^m-1\}$, входящих в T_i и T_k . Основная сложность определения данной метрики заключается в получении $p!$ сочетаний пар для каждого адреса $t_{i,j} = t_{k,r}$ в T_i и T_k . Для $p = 2$ количество подобных сочетаний равняется двум. Рассмотрим задачу вычисления $AD(T_i, T_k)$ для случая последовательностей $2A$, представленных в табл. 2.

Таблица 2
Примеры T_i и T_k тестовых последовательностей адресов $2A$

Table 2
Examples of $2A$ test sequences for addresses T_i and T_k

T_i	j	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	$t_{i,j}$	000	001	010	011	000	001	010	011	100	101	110	111	100	101	110	111
T_k	r	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	$t_{k,r}$	110	111	101	100	101	000	001	011	010	000	010	011	100	110	001	111

В приведенных тестовых наборах адресов T_i и T_k каждый из восьми адресов повторяется дважды и каждый из них входит в два возможных сочетания пар адресов. Например, адрес 010

входит в последовательность T_i как два ее адреса $t_{i,2}$ и $t_{i,6}$, а в последовательность T_k , соответственно, как $t_{k,8}$ и $t_{k,10}$. Таким образом, первое и второе сочетания пар адресов имеют следующий вид: $\{(t_{i,2}, t_{k,8}), (t_{i,6}, t_{k,10})\}$ и $\{(t_{i,6}, t_{k,8}), (t_{i,2}, t_{k,10})\}$.

В общем случае для каждого значения адреса $A \in \{0, 1, 2, \dots, 2^m - 1\}$ в адресных последовательностях T_i и T_k при $p = 2$ используются четыре одинаковых значения $t_{ij} = t_{i,l} = t_{k,r} = t_{k,g}$, образующие два сочетания пар адресов. Следовательно, можно предположить, что пара $\{(t_{ij}, t_{k,r}), (t_{i,l}, t_{k,g})\}$ представляет собой первое сочетание, а пара $\{(t_{i,l}, t_{k,r}), (t_{ij}, t_{k,g})\}$ – второе. Каждое сочетание пар адресов характеризуется суммой минимальных интервалов для каждой пары сочетания: в первом случае $D(t_{ij}, t_{k,r}) + D(t_{i,l}, t_{k,g})$ и во втором $D(t_{i,l}, t_{k,r}) + D(t_{ij}, t_{k,g})$. Далее для каждого адреса $A = t_{ij} = t_{i,l} = t_{k,r} = t_{k,g}$ определяется минимальное значение суммы минимальных интервалов согласно выражению [11]

$$D(A) = \min \left\{ \begin{array}{l} \min[|j-r|, 2^{m+1}-|j-r|] + \min[|l-g|, 2^{m+1}-|l-g|], \\ \min[|l-r|, 2^{m+1}-|l-r|] + \min[|j-g|, 2^{m+1}-|j-g|] \end{array} \right\}. \quad (3)$$

Значение 2^{m+1} в выражении (3) представляет собой размерность n последовательностей T_i и T_k , которая для примеров, приведенных в табл. 2, принимает значение $2^{3+1} = 16$. Соотношение $\min[|j-r|, 2^{m+1}-|j-r|]$ в выражении (3) является минимальным расстоянием $D(t_{ij}, t_{k,r})$ между адресами t_{ij} и $t_{k,r}$.

Ранее показывалось (см. табл. 2), что адрес 010 образует в последовательностях T_i и T_k два сочетания одинаковых адресов $\{(t_{i,2}, t_{k,8}), (t_{i,6}, t_{k,10})\}$ и $\{(t_{i,6}, t_{k,8}), (t_{i,2}, t_{k,10})\}$. Соответствующее расстояние между адресами пары $(t_{i,2}, t_{k,8})$ первого сочетания принимает значение $D(t_{i,2}, t_{k,8}) = \min(|2-8|, 16-|2-8|) = \min(6, 10) = 6$, а второй пары – значение $D(t_{i,6}, t_{k,10}) = \min(|6-10|, 16-|6-10|) = \min(4, 12) = 4$. Для второго сочетания пар адресов 010 получим: $D(t_{i,6}, t_{k,8}) = \min(|6-8|, 16-|6-8|) = \min(2, 14) = 2$; $D(t_{i,2}, t_{k,10}) = \min(|2-10|, 16-|2-10|) = \min(8, 8) = 8$. Далее для каждого сочетания в соответствии с выражением (3) вычисляется сумма полученных расстояний, которая для первого сочетания адресов 010 равняется $D(t_{i,2}, t_{k,8}) + D(t_{i,6}, t_{k,10}) = 6 + 4 = 10$, а для второго – $D(t_{i,6}, t_{k,8}) + D(t_{i,2}, t_{k,10}) = 2 + 8 = 10$. Окончательно для адреса 010 согласно (3) получим значение $D(010) = 10$. Для адреса 000 существуют два сочетания $\{(t_{i,0}, t_{k,5}), (t_{i,4}, t_{k,9})\}$ и $\{(t_{i,4}, t_{k,5}), (t_{i,0}, t_{k,9})\}$ (см. табл. 2). Соответствующие интервалы для каждой из пар двух сочетаний принимают значения $D(t_{i,0}, t_{k,5}) = 5$, $D(t_{i,4}, t_{k,9}) = 5$, $D(t_{i,4}, t_{k,5}) = 1$, $D(t_{i,0}, t_{k,9}) = 7$, тогда $D(000) = 8$. Для рассмотренного выше примера $AD(T_i, T_k) = D(000) + D(001) + D(010) + D(011) + D(100) + D(101) + D(110) + D(111) = 8 + 4 + 10 + 8 + 5 + 10 + 5 + 6 = 56$.

Из приведенного примера видно, что в общем случае для наборов адресов T_i и T_k , соответствующих определению 3, метрика расстояния $AD(T_i, T_k)$, показывающая степень их различия, определяется как минимальная сумма сумм минимальных расстояний $D(A)$ (3) для всех значений адресов $\{0, 1, \dots, 2^m - 1\}$. Расстояние $AD(T_i, T_k)$ для произвольных последовательностей T_i и T_k , удовлетворяющих условиям определения 3, по аналогии со схожей метрикой, рассмотренной в работах [19, 20], принимает значения в диапазоне $0 \leq AD(T_i, T_k) \leq 2^{2m}$. Минимальное значение $\min AD(T_i, T_k) = 0$ достигается для $T_i = T_k$. Оценка максимального значения $AD(T_i, T_k)$ основана на анализе расстояния между адресами $t_{ij} = t_{i,l} = t_{k,g} = t_{k,r}$. Очевидно, что $D(t_{ij}, t_{i,l})$ и $D(t_{k,r}, t_{k,g})$ для произвольных тождественных адресов t_{ij} и $t_{i,l}$ в T_i , а также $t_{k,r}$ и $t_{k,g}$ в T_k не превышают 2^m . Задача максимизации согласно выражению (3), по сути, заключается во взаимном расположении $t_{ij} = t_{i,l} = t_{k,g} = t_{k,r}$ в пространстве значений индексов j, l, g и r . Предположив, что $j < g < l < r$, максимизация значения (3) достигается для $l - j = r - g = 2^m$ и $g - j = r - l = 2^{m-1}$. Таким образом, для каждого адреса $\{0, 1, \dots, 2^m - 1\}$ максимальное значение расстояния (3) равняется 2^m . Учитывая, что для тестовых наборов адресов, удовлетворяющих определению 3 для $p = 2$, их количество равняется 2^m , окончательно получим $\max AD(T_i, T_k) = 2^{2m}$.

Для рассмотренного выше примера $AD(T_i, T_k) = 56$, а $\max AD(T_i, T_k) = 2^{2 \cdot 3} = 64$, что свидетельствует о большой степени различия адресных тестовых последовательностей T_i и T_k , представленных в табл. 2.

2. Мера различия $AD(T_i, T_k)$ для двоичного случая. Более широкая трактовка тестового набора по сравнению с тестовыми наборами адресов существенно усложняет процедуру вычисления меры различия, определенную для общего случая. Даже для двоичных тестовых данных комбинаторный характер определения этой меры значительно увеличивает временную сложность ее вычисления. В качестве примера, иллюстрирующего данное утверждение, рассмотрим случай двоичных тестовых наборов $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,n-1}$ и $T_k = t_{k,0}, t_{k,1}, \dots, t_{k,n-1}$, $i, k \in \{0, 1, \dots, q-1\}$, $t_{i,j}, t_{k,r} \in \{0, 1\}$, приведенных в табл. 3 для величины $n = 8$.

Таблица 3
Пример двоичных тестовых последовательностей

Table 3
Example of binary test sequences

T_i	j	0	1	2	3	4	5	6	7
	$t_{i,j}$	0	0	1	1	1	1	1	1
T_k	r	0	1	2	3	4	5	6	7
	$t_{k,r}$	1	0	1	1	1	0	1	1

В силу того что наборы T_i и T_k состоят из двоичных данных, в общем случае существует большое количество сочетаний пар тождественных данных. Для примера, приведенного в табл. 3, имеются два сочетания данных $\{(t_{i,0}, t_{k,1}), (t_{i,1}, t_{k,5})\}$ и $\{(t_{i,0}, t_{k,5}), (t_{i,1}, t_{k,1})\}$ для нулевого их значения $t_{i,j} = t_{k,r} = 0$. Для данных $t_{i,j} = t_{k,r} = 1$ число сочетаний пар единичных величин в наборах T_i и T_k равняется $6!$, где 6 представляет собой число единиц в T_i и T_k . Далее из всевозможных сочетаний пар (как нулевых значений данных, так и единичных) необходимо выбрать по одному сочетанию. Для этих пар сумма их минимальных положительных разностей индексов (интервалов) также должна быть минимальной, как следует из определения 1. Используя ранее введенную метрику расстояния $D(t_{i,j}, t_{k,r}) = \min(|j - r|, n - |j - r|)$ между $t_{i,j}$ и $t_{k,r}$, для сочетания пар $\{(t_{i,0}, t_{k,1}), (t_{i,1}, t_{k,5})\}$ получим $D(t_{i,0}, t_{k,1}) = \min(|0 - 1|, 8 - |0 - 1|) = 1$ и $D(t_{i,1}, t_{k,5}) = \min(|1 - 5|, 8 - |1 - 5|) = 4$. Соответственно, $D(t_{i,0}, t_{k,1}) + D(t_{i,1}, t_{k,5}) = 1 + 4 = 5$. В то же время для другого сочетания пар $\{(t_{i,0}, t_{k,5}), (t_{i,1}, t_{k,1})\}$ нулевых значений данных в наборах T_i и T_k получим $D(t_{i,0}, t_{k,5}) + D(t_{i,1}, t_{k,1}) = \min(|0 - 5|, 8 - |0 - 5|) + \min(|1 - 1|, 8 - |1 - 1|) = 3$. Отсюда следует, что при определении $AD(T_i, T_k)$ для примера, приведенного в табл. 3, будет учтено второе сочетание пар $\{(t_{i,0}, t_{k,5}), (t_{i,1}, t_{k,1})\}$, для которого сумма минимальных расстояний $D(t_{i,j}, t_{k,r})$, принимающая значение 3, также минимальна. Минимальная сумма минимальных расстояний для единичных данных в рассматриваемом примере определяется путем анализа сумм расстояний для $6!$ сочетаний пар. Примерами подобных сочетаний могут быть следующие два сочетания: $\{(t_{i,2}, t_{k,2}), (t_{i,3}, t_{k,3}), (t_{i,4}, t_{k,4}), (t_{i,5}, t_{k,6}), (t_{i,6}, t_{k,7}), (t_{i,7}, t_{k,0})\}$ и $\{(t_{i,2}, t_{k,3}), (t_{i,3}, t_{k,4}), (t_{i,4}, t_{k,6}), (t_{i,5}, t_{k,7}), (t_{i,6}, t_{k,0}), (t_{i,7}, t_{k,2})\}$. Для первого сочетания сумма минимальных расстояний равняется $0+0+0+1+1+1=3$, а для второго, соответственно, $1+1+2+2+2+3=11$. Подобным образом анализируются все сочетания пар единичных данных. Оказывается, что минимальное значение минимальных сумм так же, как и для нулевых данных, равняется трем и, соответственно, $AD(T_i, T_k) = 6$.

Рассмотренный пример основан на двоичных тестовых наборах, состоящих из одинакового числа единичных и нулевых значений данных $t_{i,j}$ и $t_{k,r}$, что является маловероятным случаем. Чаще всего эти числа не будут совпадать, что приведет к появлению данных, например, в наборе T_i , для которых будут отсутствовать пары в наборе T_k , и наоборот. Примером, иллюстрирующим подобную ситуацию, могут быть наборы $T_i = 10000000$ и $T_k = 01111111$, для которых можно выделить по одной паре как нулевых, так и единичных данных в выбранном сочетании пар. Следуя определению 1, значение меры различия $AD(T_i, T_k)$ определяется суммой $D(t_{i,0}, t_{k,1}) + D(t_{i,1}, t_{k,0}) + D(t_{i,2}, -) + D(t_{i,3}, -) + D(t_{i,4}, -) + D(t_{i,5}, -) + D(t_{i,6}, -) + D(t_{i,7}, -) + D(t_{k,2}, -) + D(t_{k,3}, -) + D(t_{k,4}, -) + D(t_{k,5}, -) + D(t_{k,6}, -) + D(t_{k,7}, -)$, где значения расстояния $D(t_{i,j}, t_{k,r})$ для данных, не входящих в выбранные сочетания пар, принимаются равными $\min(n_i, n_k) = 8$ согласно определению 1 [11], что приводит к существенному увеличению меры $AD(T_i, T_k)$. При этом значимость расстояния между парами идентичных данных будет иметь меньшее влияние на конечное значение суммарной характеристики. Поэтому для двоичного случая примем следующие допуще-

ния. Не нарушая общности рассуждений, будем считать, что $n_i = n_k = n$ и его значение четно. В силу равенства размерностей наборов T_i и T_k количество данных $t_{i,j}$ и $t_{k,r}$, не имеющих пары в другом наборе, будет одинаковым. Поэтому в качестве слагаемых для пар данных, не вошедших в пары идентичных данных, т. е. с несовпадающими значениями, примем величину расстояния, равную $n/2$. Отметим, что $n/2$ представляет собой максимально возможное значение расстояния для пары идентичных данных.

Формально значения расстояний $D(t_{i,j}, t_{k,r}) = \min(|j - r|, n - |j - r|)$ между тождественными данными $t_{i,j} = t_{k,r}$ можно представить в виде двухмерной матрицы их величин (табл. 4).

Таблица 4
Значения расстояний $D(t_{i,j}, t_{k,r})$ между $t_{i,j} = t_{k,r}$ для четного n

Table 4
The values of the distances $D(t_{i,j}, t_{k,r})$ between $t_{i,j} = t_{k,r}$ for even n

Набор Pattern	T_i										
T_k	Данные Data	$t_{i,0}$	$t_{i,1}$	$t_{i,2}$...	$t_{i,n/2-1}$	$t_{i,n/2}$	$t_{i,n/2+1}$...	$t_{i,n-2}$	$t_{i,n-1}$
	$t_{k,0}$	0	1	2	...	$n/2-1$	$n/2$	$n/2-1$...	2	1
	$t_{k,1}$	1	0	1	...	$n/2-2$	$n/2-1$	$n/2$...	3	2
	$t_{k,2}$	2	1	0	...	$n/2-3$	$n/2-2$	$n/2-1$...	4	3

	$t_{k,n/2-1}$	$n/2-2$	$n/2-3$	$n/2-4$...	0	1	2	...	$n/2$	$n/2-1$
	$t_{k,n/2}$	$n/2-1$	$n/2-2$	$n/2-3$...	1	0	1	...	$n/2-1$	$n/2$
	$t_{k,n/2+1}$	$n/2$	$n/2-1$	$n/2-2$...	2	1	0	...	$n/2-2$	$n/2-1$

	$t_{k,n-2}$	2	3	4	...	$n/2-1$	$n/2-2$	$n/2-3$...	0	1
	$t_{k,n-1}$	1	2	3	...	$n/2$	$n/2-1$	$n/2-2$...	1	0

Структура данных приведенной матрицы расстояний размерностью $n \times n$ показывает их регулярность, а численное значение на пересечении столбца, помеченного $t_{i,j}$, и строки, соответствующей $t_{k,r}$, представляет собой значение $D(t_{i,j}, t_{k,r}) = \min(|j - r|, n - |j - r|)$ для $t_{i,j} = t_{k,r}$. Из табл. 4 видно, что, например, при $j = r$ $D(t_{i,j}, t_{k,r}) = 0$.

Очевидно, что минимальное значение меры различия $AD(T_i, T_k)$ достигается в случае тождественности наборов T_i и T_k и для принятых допущений равенства $n_i = n_k = n$ равняется нулю, так как $|n_i - n_k|$ будет равняться нулю и, соответственно, $\min AD(T_i, T_k) = |n_i - n_k| \cdot \min(n_i, n_k) = 0$ (см. определение 1). Для получения $\min AD(T_i, T_k) = 0$ выбирается совокупность пар совпадающих данных $t_{i,j} = t_{k,r}$ с $D(t_{i,j}, t_{k,r}) = 0$ (табл. 5) для случая идентичных наборов $T_i = T_k = 11101001$. Диагональные нулевые значения и определяют пары совпадающих данных, принимающих участие в совокупности пар для определения значения $AD(T_i, T_k)$.

Таблица 5
Значения расстояний $D(t_{i,j}, t_{k,r})$ между $t_{i,j} = t_{k,r}$ для $n = 8$

Table 5
The values of the distances $D(t_{i,j}, t_{k,r})$ between $t_{i,j} = t_{k,r}$ for $n = 8$

Набор Pattern	T_i								
T_k	Данные Data	1	1	1	0	1	0	0	1
	1	0	1	2		4			1
	1	1	0	1		3			2
	1	2	1	0		2			3
	0				0		2	3	
	1	4	3	2		0			3
	0				2		0	1	
	0				3		1	0	
	1	1	2	3		3			0

При заполнении матрицы расстояний указываются только расстояния $D(t_{i,j}, t_{k,r})$ для $t_{i,j} = t_{k,r}$, соответствующие значениям для общего случая (см. табл. 4). При их несовпадении, т. е. при $t_{i,j} \neq t_{k,r}$, соответствующий элемент матрицы остается пустым.

В терминах меры различия согласно определению 1 полностью различными, несовпадающими наборами T_i и T_k считаются те, в которых отсутствуют данные, используемые в другом наборе. Для рассматриваемого случая двоичных тестовых наборов T_i и T_k максимально различными являются наборы $T_i = 000\dots 0$ и $T_k = 111\dots 1$ либо, наоборот, $T_i = 111\dots 1$ и $T_k = 000\dots 0$, для которых $AD(T_i, T_k) = \max AD(T_i, T_k) = n(n/2) = n^2/2$. Последнее соотношение получено на основании ранее принятых допущений о максимальной удаленности на $n/2$ позиций данных в парах несовпадающих данных, которые не вошли в сочетание совпадающих данных, и их количество в этом случае равняется n . Сравнивая полученное значение $AD(T_i, T_k) = 6$ с $\min AD(T_i, T_k) = 0$ и $\max AD(T_i, T_k) = n^2/2 = 32$, можно сделать заключение о практической тождественности (равенстве) тестовых двоичных наборов T_i и T_k данных, приведенных в табл. 3.

Количественное совпадение единичных и нулевых значений в наборах T_i и T_k значительно уменьшает величину $AD(T_i, T_k)$, показывая их схожесть. Различие таких наборов зависит от расположения в них данных и соответствующих расстояний между ними. Очевидно, что большее различие между наборами T_i и T_k наблюдается при неравном количестве у них как единичных, так и нулевых данных. В этом случае появляются пары несовпадающих данных, для которых расстояние принимается равным $n/2$. Примером подобных данных могут служить $T_i = 01010101$ и $T_k = 11100011$, значения расстояний для которых даны в табл. 6.

Таблица 6
Значения расстояний $D(t_{i,j}, t_{k,r})$ между $t_{i,j} = t_{k,r}$
для $T_i = 01010101$ и $T_k = 11100011$

Table 6
The values of the distances $D(t_{i,j}, t_{k,r})$ between $t_{i,j} = t_{k,r}$
for $T_i = 01010101$ and $T_k = 11100011$

Набор Pattern	T_i								
	Данные Data	0	1	0	1	0	1	0	1
T_k	1		1		3		3		1
	1		0		2		4		2
	1		1		1		3		3
	0	3		1		1		3	
	0	4		2		0		2	
	0	3		3		1		1	
	1		3		3		1		1
	1		2		4		2		0

В приведенных наборах количество пар совпадающих нулевых значений равняется трем, так как для четвертого нуля набора T_i отсутствует четвертый нуль в наборе T_k , а количество пар совпадающих единичных данных равняется четырем. Следуя определению 1, для нулевых тождественных данных $t_{i,j} = t_{k,r} = 0$ находим сочетание пар, для которых сумма минимальных расстояний $D(t_{i,j}, t_{k,r})$ также минимальна. Это достигается путем выбора сочетания пар $\{(t_{i,2}, t_{k,3}), (t_{i,4}, t_{k,4}), (t_{i,6}, t_{k,5})\}$, для которых сумма $D(t_{i,2}, t_{k,3}) + D(t_{i,4}, t_{k,4}) + D(t_{i,6}, t_{k,5})$ минимальна и равняется двум. Отметим, что общее количество сочетаний пар нулевых данных в указанных наборах равняется 24. Аналогично выбирается сочетание пар единичных значений $\{(t_{i,1}, t_{k,1}), (t_{i,3}, t_{k,2}), (t_{i,5}, t_{k,6}), (t_{i,7}, t_{k,7})\}$, для которых сумма их расстояний принимает минимальное значение, равное двум. Для пары данных $t_{i,0}$ и $t_{k,0}$, которые не приняли участие в выбранных ранее сочетаниях пар, расстояние принимается равным $n/2 = 4$. Окончательно для $T_i = 01010101$ и $T_k = 11100011$ значение $AD(T_i, T_k)$ вычисляется как сумма $2 + 2 + 4 = 8$.

Рассмотренный пример показал, что основной проблемой при вычислении $AD(T_i, T_k)$ в случае двоичных данных является необходимость анализа большого количества сочетаний дан-

ных. Эта задача носит комбинаторный характер, сложность ее реализации определяется количеством сочетаний независимых пар нулевых и единичных значений в двоичных наборах T_i и T_k . В общем случае для наборов T_i и T_k , имеющих одинаковую размерность n , общее количество Q сочетаний независимых пар данных можно представить соотношением

$$Q = \binom{u}{l} \cdot l! + \binom{n-l}{n-u} \cdot (n-u)! = \frac{u! + (n-l)!}{(u-l)!}, \quad l = \min[w(T_i), w(T_k)], \quad u = \max[w(T_i), w(T_k)]. \quad (4)$$

Выражение $w(T_i)$ означает вес (количество единиц) вектора T_i , а $w(T_k)$ – вес вектора T_k . Первое слагаемое в выражении (4) определяет количество сочетаний независимых пар единичных данных $t_{i,j} = t_{k,r} = 1$, общее число которых равняется l , а второе слагаемое – количество сочетаний независимых $(n-u)$ пар нулевых данных. Отметим, что число данных в наборах T_i и T_k , для которых отсутствует пара с одинаковым значением, определяется разностью $u-l$. Для рассмотренного выше примера (см. табл. 6) $w(T_i) = 4$, $w(T_k) = 5$. Соответственно, $l = 4$, $u = 5$. Таким образом, в наборах T_i и T_k выделяются $l = 4$ пары единичных данных и $n-u = 8-5 = 3$ пары нулевых данных. Общее количество Q сочетаний независимых единичных и нулевых пар данных в соответствии с выражением (4) вычисляется как $(u! + (n-l)!)/(u-l)! = (5! + 4!)/(1!) = 144$. Одна пара данных в наборах T_i и T_k имеет несовпадающие значения, так как $u-l = 5-4 = 1$.

Приведенная оценка Q количества сочетаний независимых пар свидетельствует о большом их числе для случаев реальных тестовых наборов и невозможности их полного перебора, необходимого для вычисления $AD(T_i, T_k)$.

3. Сведение задачи вычисления меры различия $AD(T_i, T_k)$ к задаче о назначениях. Приведенная в разд. 2 формулировка задачи вычисления меры различия $AD(T_i, T_k)$ для двоичного случая вытекает из определения 1, сформулированного для общего случая в работе [11]. Важными уточнениями указанной задачи являются алфавит тестовых данных, декларирующий использование только двоичных значений, и одинаковая размерность n тестовых наборов T_i и T_k . Допущение, что каждый из наборов содержит по n двоичных данных $t_{i,j}$ и $t_{k,r}$, позволяет сформировать полное множество сочетаний из n пар. Все их множество будет состоять из пар совпадающих данных $t_{i,j} = t_{k,r}$, удовлетворяющих определению 1 и требующих решения сложной комбинаторной задачи, подробно описанной в разд. 2, и пар несовпадающих данных $t_{i,j} \neq t_{k,r}$. Сложность выбора оптимального сочетания пар, удовлетворяющих определению 1, следует из того факта, что в зависимости от выбора конкретной пары идентичных данных $t_{i,j} = t_{k,r}$ расстояние $D(t_{i,j}, t_{k,r})$ между ними для четных величин n принимает значение от 0 до $n/2$ в соответствии с табл. 4. Выбор оптимального сочетания пар идентичных данных заключается в нахождении такого их сочетания, для которого сумма расстояний будет минимальной. Для пар несовпадающих данных в силу того, что согласно определению 1 для каждого из этих данных задается фиктивное значение расстояния, следует возможность выбора любого их сочетания. Важным является только значение количества несовпадающих пар, для каждой из которых задается расстояние в виде константы $D(t_{i,j} \neq t_{k,r}) = n/2$. Величина приведенного значения принята равной максимальному расстоянию $D(t_{i,j}, t_{k,r})$ для совпадающих данных.

Для нечетных значений отличием является только максимальное значение расстояния $D(t_{i,j}, t_{k,r}) = \lfloor n/2 \rfloor$, а вид матрицы расстояний остается таким же, как и для четных значений n . Для произвольного целого n матрица $n \times n$ расстояний $D(t_{i,j}, t_{k,r})$ представляет собой циркулярную матрицу, пример которой для четных величин n приведен в табл. 4.

В качестве примера для $n = 9$ рассмотрим наборы двоичных данных $T_i = 111010111$ и $T_k = 111101011$, для которых значения расстояний для всевозможных пар данных приведены в табл. 7.

Отметим, что в отличие от примеров, описанных ранее в табл. 5 и 6, в приведенной квадратной матрице размерностью 9×9 определены все ее элементы. Так как $n = 9$ является нечетным числом, то $D(t_{i,j}, t_{k,r}) = \lfloor n/2 \rfloor = 4$. Очевидно, что и для произвольного случая матрицы размерностью $n \times n$ также будут определены все ее элементы. Для совпадающих данных и четного значения n они будут соответствовать величинам, приведенным в табл. 4, а для нечетных n строки

аналогичной матрицы являются циклическими сдвигами первой строки, состоящей из элементов $0, 1, 2, \dots, \lfloor n/2 \rfloor - 1, \lfloor n/2 \rfloor, \lfloor n/2 \rfloor, \lfloor n/2 \rfloor - 1, \dots, 3, 2, 1$. Для всех пар несовпадающих данных соответствующий им элемент матрицы для произвольного n равняется $\lfloor n/2 \rfloor$.

Таблица 7
Значения расстояний $D(t_{ij}, t_{k,r})$ для $T_i = 111010111$ и $T_k = 111101011$

Table 7
The values of the distances $D(t_{ij}, t_{k,r})$ for $T_i = 111010111$ and $T_k = 111101011$

Набор Pattern	T_i									
Данные Data	1	1	1	0	1	0	1	1	1	1
1	0	1	2	4	4	4	3	2	1	
1	1	0	1	4	3	4	4	3	2	
1	2	1	0	4	2	4	4	4	3	
1	3	2	1	4	1	4	3	4	4	
0	4	4	4	1	4	1	4	4	4	
1	4	4	3	4	1	4	1	2	3	
0	4	4	4	3	4	1	4	4	4	
1	2	3	4	4	3	4	1	0	1	
1	1	2	3	4	4	4	2	1	0	

Таким образом, в случае двоичных тестовых наборов T_i и T_k размерностью n бит задача вычисления меры различия $AD(T_i, T_k)$ может быть сформулирована в виде следующего определения.

Определение 4. Мера различия $AD(T_i, T_k)$, соответствующая определению 1, для двоичных тестовых наборов T_i и T_k , каждый из которых состоит из n бит $t_{ij}, t_{k,r} \in \{0, 1\}$, $j, r \in \{0, 1, \dots, n-1\}$, основана на построении циркулярной матрицы расстояний размерностью $n \times n$. Первая строка данной матрицы имеет вид $0, 1, 2, \dots, n/2-1, n/2, n/2-1, \dots, 3, 2, 1$ для четных значений n и $0, 1, 2, \dots, \lfloor n/2 \rfloor - 1, \lfloor n/2 \rfloor, \lfloor n/2 \rfloor, \lfloor n/2 \rfloor - 1, \dots, 3, 2, 1$ – для нечетных. С использованием циркулярной матрицы строится матрица расстояний таким образом, что в случае совпадающих данных $t_{ij} = t_{k,r}$ соответствующий элемент этой матрицы равняется элементу циркулярной матрицы, а в случае их несовпадения принимает значение расстояния, равное $\lfloor n/2 \rfloor$. Численное значение $AD(T_i, T_k)$ определяется минимальной суммой n элементов матрицы расстояний, покрывающих все строки и все столбцы матрицы расстояний.

Согласно определению 4 строится исходная квадратная матрица расстояний, которая аналогична исходной квадратной матрице стоимости, используемой в задаче о назначениях [21]. В обоих случаях решение задачи вычисления меры различия и задачи о назначениях заключается в нахождении n элементов матрицы, которые покрывают все строки и все столбцы, а сумма их значений минимальна. Отметим, что каждый из n найденных элементов определяет пару, состоящую из строки и столбца с соответствующими индексами. При этом каждый столбец и каждая строка входят только в одну пару.

Задачу о назначениях можно описать, если сформулировать ее, используя двудольный граф. Следовательно, задача нахождения меры различия $AD(T_i, T_k)$ тестовых наборов T_i и T_k , каждый из которых состоит из n_i и n_k данных $t_{ij}, j \in \{0, 1, \dots, n_i - 1\}$, и $t_{k,r}, r \in \{0, 1, \dots, n_k - 1\}$, в терминах задачи на графах формулируется следующим образом. Дан полный двудольный граф G с n вершинами, соответствующими данным t_{ij} набора T_i , и n вершинами, соответствующими данным $t_{k,r}$ набора T_k . Стоимость (вес) каждого ребра графа неотрицательна и равняется величине расстояния $D(t_{ij}, t_{k,r})$ между данными t_{ij} и $t_{k,r}$. Требуется найти совершенное, или полное, паросочетание с наименьшей стоимостью, равной сумме расстояний $D(t_{ij}, t_{k,r})$ и определяемой значением меры различия $AD(T_i, T_k)$.

Приведенное определение соответствует задаче нахождения оптимального сочетания пар данных $t_{i,j}$ и $t_{k,r}$ в общем случае для произвольных наборов T_i и T_k . Наиболее близким известным решением подобной задачи является решение задачи о назначениях с помощью венгерского алгоритма [21]. Главная идея данного алгоритма была разработана Х. Куном и нашла широкое применение на практике [22]. С помощью одного из онлайн-приложений (URL: <https://math.semestr.ru/nazn/index.php>) для решения задачи о назначениях в случае определения меры различия $AD(T_i, T_k)$, где $T_i = 111010111$ и $T_k = 111101011$, а матрица расстояний $D(t_{i,j}, t_{k,r})$ приведена в табл. 7, было получено следующее решение:

$$\begin{aligned} Cmin &= 0 + 0 + 0 + 1 + 1 + 1 + 1 + 0 + 0 = 4. \\ \text{Путь: } &(1;1), (2;2), (3;3), (4;5), (5;4), (6;7), (7;6), (8;8), (9;9). \end{aligned} \quad (5)$$

В терминах задачи вычисления меры различия $Cmin$ (5) представляет собой искомое значение $AD(T_i, T_k) = 4$ для $T_i = 111010111$ и $T_k = 111101011$, а множество *Путь* определяет оптимальное сочетание пар $\{(t_{i,1}, t_{k,1}), (t_{i,2}, t_{k,2}), (t_{i,3}, t_{k,3}), (t_{i,4}, t_{k,5}), (t_{i,5}, t_{k,4}), (t_{i,6}, t_{k,7}), (t_{i,7}, t_{k,6}), (t_{i,8}, t_{k,8}), (t_{i,9}, t_{k,9})\}$. Временная сложность оригинального алгоритма о назначениях имеет полиномиальную оценку $O(n^4)$ [21]. Поэтому даже в случае примера, решение для которого приведено в (5) ($n = 9$), требуются существенные временные затраты. Уменьшение временной сложности алгоритма до $O(n^3)$ [22] также не позволяет применять его с целью вычисления меры различия для реальных случаев управляемых вероятностных тестов.

4. Алгоритм вычисления меры различия $AD(T_i, T_k)$. Приведенный выше анализ показал высокую вычислительную сложность определения меры различия $AD(T_i, T_k)$ для произвольных тестовых наборов T_i и T_k . Основная проблема заключается в необходимости рассмотрения большого количества Q сочетаний независимых пар тождественных данных. Только в этом случае возможно получение минимального значения $AD(T_i, T_k)$, которое показывает степень различия сравниваемых наборов. Получить близкое к оптимальному решение в смысле минимальности значения указанной меры можно с использованием более простых с точки зрения вычислительной сложности алгоритмов, в том числе и широко известных на практике, например жадного алгоритма (*Greedy algorithm*), суть которого состоит в нахождении локально оптимальных решений на каждом этапе, допуская, что конечное решение также окажется оптимальным [23].

Сущность предлагаемого алгоритма вычисления меры различия $AD(T_i, T_k)$ заключается в определении количества пар данных $t_{i,j}$ и $t_{k,r}$ тестовых наборов T_i и T_k , имеющих определенное значение расстояния $D(t_{i,j}, t_{k,r})$. Учитывая то, что для $n = \max(n_i, n_k)$ значение $D(t_{i,j}, t_{k,r})$ равняется минимальному значению из двух величин $|j - r|$ и $n - |j - r|$, необходимо рассмотреть только $\lfloor n/2 \rfloor + 1$ возможных значений $D(t_{i,j}, t_{k,r})$. В отличие от общей постановки задачи вычисления меры $AD(T_i, T_k)$ в предлагаемом алгоритме последовательно от 0 до $\lfloor n/2 \rfloor + 1$ определяются количества пар тождественных данных $t_{i,j} = t_{k,r}$, имеющих соответствующее значение расстояния $D(t_{i,j}, t_{k,r}) = 0 \div \lfloor n/2 \rfloor + 1$.

Исходными являются тестовые наборы T_i и T_k , каждый из которых состоит из n_i и n_k данных $t_{i,j}, j \in \{0, 1, \dots, n_i - 1\}$, и $t_{k,r}, r \in \{0, 1, \dots, n_k - 1\}$. Предположив, что $n_i \geq n_k$, расстояния между данными наборов T_i и T_k генерируются путем циклического сдвига T_i вправо и влево с последующим определением количества совпадающих данных, имеющих одинаковые индексы j и r . В этом случае индекс j для каждого данного $t_{i,j}$ набора T_i примет все возможные значения $j \in \{0, 1, \dots, n_i - 1\}$ как результат операций циклического сдвига T_i . Таким образом, каждому данному $t_{k,r}$ набора T_k последовательно будут сопоставлены все данные $t_{i,j}$ и проанализированы на предмет совпадения. Совпадение данных $t_{i,j}$ и $t_{k,r}$ свидетельствует о наличии пары тождественных данных с расстоянием $D(t_{i,j}, t_{k,r})$, равным количеству циклических сдвигов набора T_i . Независимость пар тождественных данных обеспечивается исключением данных выявленной пары из дальнейшего рассмотрения. В результате подобных действий определяются пары тожде-

ственных данных $t_{i,j}$ и $t_{k,r}$ с фиксированными значениями расстояния $D(t_{i,j}, t_{k,r})$, на основании которых и вычисляется значение меры различия $AD(T_i, T_k)$.

Алгоритм определения $AD(T_i, T_k)$ состоит из следующих шагов:

1. В исходных тестовых наборах T_i и T_k сравниваются данные $t_{i,j}$ и $t_{k,r}$ для $j = r$ и определяется количество $q(0)$ пар совпадающих данных $t_{i,r} = t_{k,r}$. В результате идентифицируются данные наборов T_i и T_k , имеющие расстояние $D(t_{i,j}, t_{k,r}) = 0$, и их число $q(0)$. Выделенные данные, входящие в наборы T_i и T_k , исключаются из дальнейшего рассмотрения. При $q(0) = n_k$, т. е. при совпадении первых n_k данных набора T_i с данными второго набора T_k , задается $q(t) = n_i - n_k$. Переход к п. 5.

2. Последовательно для $v = 1, 2, \dots, \lfloor n_i / 2 \rfloor - 1$ формируются циклические сдвиги тестового набора T_i относительно набора T_k на v позиций вправо и влево. Операция сдвига эквивалентна уменьшению и увеличению значения индекса j на величину v по модулю n_i . Далее определяется количество $q(v)$ тождественных пар данных с совпадающими индексами, а именно с модифицированным в результате сдвига $j = (j \pm v) \bmod n_i$ и индексом r . Величина $q(v)$ определяет количество пар данных $t_{i,j}$ и $t_{k,r}$, имеющих расстояние $D(t_{i,j}, t_{k,r}) = v$. При формировании очередного значения сдвига v набора T_i тождественные данные, выделенные при всех предыдущих сдвигах, меньших v , исключаются из рассмотрения с соблюдением последовательности их выделения. Выполнение данного шага прекращается в случае, если количество пар эквивалентных данных, выявленных в наборах T_i и T_k , достигло величины n_k . Переход к п. 5 после задания $q(t) = n_i - n_k$.

3. Если n_i , большее либо равно n_k , является нечетным числом, повторяются действия шага 2 для $v = \lfloor n_i / 2 \rfloor$, а если n_i четно – формируется только один сдвиг на $v = n_i / 2$ вправо либо влево. В обоих случаях подсчитывается количество $q(\lfloor n_i / 2 \rfloor)$ совпадающих пар данных, имеющих расстояние $D(t_{i,j}, t_{k,r}) = \lfloor n_i / 2 \rfloor$.

4. Определяется суммарное число $q(t)$ данных $t_{i,j}$ и $t_{k,r}$ наборов T_i и T_k , которые не участвовали в парах тождественных данных.

5. Вычисляется значение меры различия $AD(T_i, T_k)$ согласно выражению

$$AD(T_i, T_k) = q(t) \cdot \min(n_i, n_k) + \sum_{v=1}^{\lfloor n_i / 2 \rfloor} v \cdot q(v). \quad (6)$$

Полученное значение $AD(T_i, T_k)$ является искомой величиной меры различия наборов T_i и T_k .

Рассмотрим применение данного алгоритма для случая тестовых наборов адресов T_i и T_k , представленных в табл. 1, для которых двоичные значения адресов $t_{i,j}$ и $t_{k,r}$ заменим их десятичными эквивалентами:

1. Сравниваются данные $t_{i,j}$ и $t_{k,r}$ для $j = r$ и определяется количество пар $q(0)$ совпадающих данных $t_{i,r} = t_{k,r}$:

$j = r$	0	1	2	3	4	5	6	7
$t_{i,j}$	<u>0</u>	<u>1</u>	2	3	4	5	6	7
$t_{k,r}$	<u>0</u>	<u>1</u>	3	2	6	7	5	4

Как видно, данные $t_{i,0}$ и $t_{k,0}$ образуют первую пару тождественных значений $t_{i,0} = t_{k,0} = 0$ и $t_{i,1} = t_{k,1} = 1$, что свидетельствует о наличии второй пары совпадающих данных (обозначены подчеркиванием). Соответственно, $q(0) = 2$, а сами данные $t_{i,0}$, $t_{k,0}$, $t_{i,1}$ и $t_{k,1}$ исключаются из дальнейшего рассмотрения путем замены их значений символом X.

2. Так как $n_i = n_k = 8$, последовательно для $v = 1, 2$ и 3 формируются циклические сдвиги тестового набора T_i относительно набора T_k на v позиций вправо и влево. Определяется количество $q(v)$ тождественных пар данных для совпадающих индексов, а именно модифицированно $j = (j \pm v) \bmod n_i$ и индекса r . Последовательность выполнения данного шага алгоритма приведена в табл. 8.

Таблица 8
Последовательность получения значений $q(v)$ для $v = 1, 2$ и 3

Table 8
Sequence of getting $q(v)$ values for $v = 1, 2$ and 3

$j=(j+v)\bmod 8=r$		0	1	2	3	4	5	6	7	$j=(j-v)\bmod 8=r$		0	1	2	3	4	5	6	7	$q(v)$
$v=1$	$t_{i,j}$	7	X	X	<u>2</u>	3	4	<u>5</u>	6	$v=1$	$t_{i,j}$	X	2	<u>3</u>	4	5	6	7	X	$q(1)=3$
	$t_{k,r}$	X	X	3	<u>2</u>	6	7	<u>5</u>	4		$t_{k,r}$	X	X	<u>3</u>	2	6	7	5	4	
$v=2$	$t_{i,j}$	6	7	X	X	X	X	4	X	$v=2$	$t_{i,j}$	X	X	4	X	<u>6</u>	<u>7</u>	X	X	$q(2)=2$
	$t_{k,r}$	X	X	X	X	6	7	X	4		$t_{k,r}$	X	X	X	X	<u>6</u>	<u>7</u>	X	4	
$v=3$	$t_{i,j}$	X	X	X	X	X	X	X	<u>4</u>	$v=3$	$t_{i,j}$	X	4	X	X	X	X	X	X	$q(3)=1$
	$t_{k,r}$	X	X	X	X	X	X	X	<u>4</u>		$t_{k,r}$	X	X	X	X	X	X	X	4	

При формировании очередного значения сдвига v набора T_i тождественные данные, выделенные при всех предыдущих сдвигах, меньших v , помечаются символом X и исключаются из рассмотрения. После реализации сдвигов на $v = 3$ в наборах T_i и T_k все данные заменяются на X. Переход к п. 5 после принятия $q(t) = 0$.

5. В соответствии с выражением (6) вычисляется $AD(T_i, T_k) = 0 + 1 \cdot 3 + 2 \cdot 2 + 3 \cdot 1 + 4 \cdot 0 = 10$, что соответствует ранее полученному результату для тестовых наборов, представленных в табл. 1.

Для двоичного случая наборов данных T_i и T_k , согласно определению 4 и с учетом введенных допущений и ограничений, соотношение для вычисления $AD(T_i, T_k)$ принимает вид

$$AD(T_i, T_k) = q(t) \frac{n}{2} + \sum_{v=1}^{n/2} v \cdot q(v). \quad (7)$$

В выражении (7) в отличие от (6) значение $q(t)$ определяет число пар несовпадающих данных в наборах, так как при равенстве числа данных $t_{i,r}$ и $t_{k,r}$ в наборах T_i и T_k для каждого из них существует пара либо тождественных данных, либо несовпадающих.

В качестве еще одного примера применения предложенного алгоритма рассмотрим случай двоичных тестовых наборов $T_i = 01010101$ и $T_k = 11100011$, матрица весов для которых представлена в табл. 6:

1. Сравниваются данные $t_{i,j}$ и $t_{k,r}$ для $j = r$ и определяется количество пар $q(0)$ совпадающих данных $t_{i,r} = t_{k,r}$:

$j = r$	0	1	2	3	4	5	6	7
$t_{i,j}$	0	<u>1</u>	0	1	<u>0</u>	1	0	<u>1</u>
$t_{k,r}$	1	<u>1</u>	1	0	<u>0</u>	0	1	<u>1</u>

Данные $t_{i,1}$ и $t_{k,1}$ образуют первую пару тождественных значений, $t_{i,4} = t_{k,4} = 0$ – вторую, $t_{i,7} = t_{k,7} = 1$ – третью. Соответственно, $q(0) = 3$, а сами данные исключаются из дальнейшего рассмотрения путем замены их значений символом X.

2. Последовательно, начиная с $v = 1$, формируются циклические сдвиги тестового набора T_i относительно набора T_k на v позиций вправо и влево. Для каждого сдвига v определяется количество $q(v)$ тождественных пар данных.

Таблица 9
Последовательность получения значения $q(1)$

Table 9
Sequence of getting $q(1)$ value

$j=(j-v)\bmod 8=r$		0	1	2	3	4	5	6	7	$j=(j+v)\bmod 8=r$		0	1	2	3	4	5	6	7	$q(v)$
$v=1$	$t_{i,j}$	X	0	<u>1</u>	X	1	<u>0</u>	X	0	$v=1$	$t_{i,j}$	X	0	X	<u>0</u>	1	X	<u>1</u>	0	$q(1)=4$
	$t_{k,r}$	1	X	<u>1</u>	0	X	<u>0</u>	1	X		$t_{k,r}$	1	X	1	<u>0</u>	X	0	<u>1</u>	X	

После выполнения одной итерации второго шага алгоритма (см. табл. 9) находятся $q(1) = 4$ пары эквивалентных данных, которые выделены подчеркиванием и затем помечены символом X. В результате набор T_i принимает вид 0XXXXXXX, а набор T_k – вид 1XXXXXXX. На последующих шагах алгоритма сформируются значения $q(2)$ и $q(3)$, равные нулю. Далее в соответствии с шагом 3 алгоритма определяется $q(4) = 0$. Суммарное число $q(t)$ пар данных $t_{i,j}$ и $t_{k,r}$ наборов T_i и T_k , которые не участвовали в парах тождественных данных, равняется единице. Полученные значения подставляются в соотношение (7), и окончательно формируется значение $AD(T_i, T_k) = 8$, что соответствует ранее полученному результату вычисления $AD(T_i, T_k)$ согласно определению 1 и его уточнениям для двоичного случая. Такое же значение получается и в результате применения онлайн-приложения, реализующего венгерский алгоритм о назначениях (URL: <https://math.semestr.ru/nazn/index.php>). Как видно из выражения (8), значение $Cmin$ тоже равняется восьми:

$$Cmin = 4 + 0 + 1 + 1 + 0 + 1 + 1 + 0 = 8.$$

Путь: (1;1), (2;2), (3;4), (4;3), (5;5), (6;7), (7;6), (8;8). (8)

Таким образом, предложенный алгоритм позволил получить аналогичный результат, который соответствует результату для меры различия, соответствующей определению 1. Однако применив данный алгоритм для случая $T_i = 111010111$ и $T_k = 111101011$, рассмотренного ранее, получим $AD(T_i, T_k) = 6$, что не соответствует оптимальному решению этой задачи (см. (6)) с использованием венгерского алгоритма.

Вычислительную сложность рассмотренного алгоритма в терминах количества операций сравнения данных в наборах T_i и T_k можно оценить величиной $O(n^2)$, где n – размерность анализируемых наборов. Это следует из того, что наборы, состоящие из n данных, анализируются на n итерациях алгоритма, в каждой из которых реализуется один из n сдвигов данных в наборе T_i относительно данных набора T_k .

5. Экспериментальные оценки меры различия. Рассмотренная мера позволяет оценить степень различия двух тестовых наборов T_i и T_k , которые могут быть неразличимыми при использовании других мер различия. В качестве иллюстрации данного утверждения рассмотрим пример двоичных наборов, второй из которых T_k является инверсией первого T_i . Для подобных наборов T_i и $T_k = \bar{T}_i$ расстояние Хэмминга $HD(T_i, \bar{T}_i)$ всегда неизменно и равняется n , в то время как $AD(T_i, \bar{T}_i)$ принимает различные значения. Например, $AD(00000000, 11111111) = 32$, $AD(00001111, 11110000) = 16$ и $AD(00110011, 11001100) = 8$, при том что расстояние Хэмминга во всех трех случаях равняется восьми.

В предыдущих разделах было показано, что точное вычисление $AD(T_i, T_k)$ для произвольных двоичных наборов T_i и T_k возможно только при использовании венгерского алгоритма. Предложенный авторами алгоритм имеет меньшую вычислительную сложность, однако не всегда повторяет значение $AD(T_i, T_k)$, соответствующее определениям 1 и 4. В то же время, как это видно для случая двоичных тестовых наборов T_i и T_k , представляющих собой случайные двоичные последовательности для $n = 32$, отличий значений может и не быть либо они будут незначительными.

В качестве последовательностей T_i и T_k были использованы двоичные представления случайных целых чисел, сгенерированных в диапазоне $[0 \div 4294967295]$. Эксперимент проводился для 10 000 пар последовательностей T_i и T_k , для которых была получена мера различия $AD(T_i, T_k)$ согласно предложенному авторами алгоритму и результат $Cmin$, вычисленный в соответствии с венгерским алгоритмом.

На рисунке изображена диаграмма для $AD(T_i, T_k)$ и $Cmin$ первых 50 пар T_i и T_k . Во многих случаях $AD(T_i, T_k)$ равняется $Cmin$, а при несовпадении значений их разность минимальна, как это видно на примере 15 пар T_i и T_k двоичных тестовых наборов, приведенных в табл. 10.



Диаграмма значений $AD(T_i, T_k)$ и $Cmin$
Value chart $AD(T_i, T_k)$ and $Cmin$

Таблица 10
Результаты вычислений для первых 15 пар тестовых наборов T_i и T_k

Table 10
Calculation results for the first 15 pairs of test cases T_i and T_k

T_i	T_k	$Cmin$	$AD(T_i, T_k)$
011110110011010110111010111001	00111010110101110100100010000010	124	126
10101111110101000001001011000111	01010000000000001101111001010001	148	148
00010111000111101001000101101101	1100100101101101111000010110110	60	66
00100110100111100111100100100001	01010001011001011001101111000011	62	66
00100011110010111100001001101110	00000110010000011001010101001101	82	84
01100111000001100101111010101110	10011010100001110001110111111110	46	46
01011100001111011001000101011001	00000001100101000000111110001101	104	104
01111111100000110110001011111111	00101100100101001001011010101011	118	118
1111101000001010101010100100010	11100000011001101010101011000111	64	64
10001100010110011101001111010100	01010110011001101111111001100100	56	56
10001101100100000111100111000110	01111110010100100001001111001110	68	68
10001001111110000100011111111101	01101010111000010001000100001101	122	123
0101001100001111111110001010101	01011110101111111001111010111110	106	106
11011001010011110100110111110110	11000101010000011100111101111111	56	64
01011010100110100110111100101101	01101101001111100011110001111101	56	57

Для сравнительного анализа вычислительной сложности двух алгоритмов была написана программа на языке Python. Программа находит оптимальное решение задачи о назначениях по венгерскому алгоритму ($Cmin$) и вычисляет меру различия $AD(T_i, T_k)$ для двух последовательностей T_i и T_k по алгоритму, предложенному авторами. Для нахождения $Cmin$ в программе используется модуль munkres (URL: <https://software.clapper.org/munkres/>). В табл. 11 представлены данные о времени выполнения венгерского алгоритма и алгоритма авторов. Для каждой пары последовательностей T_i и T_k было проведено 1000 тестов, в качестве времени выполнения принималось среднее арифметическое измеренных значений.

Результаты тестирования показали, что алгоритм вычисления меры различия $AD(T_i, T_k)$ характеризуется значительно меньшей вычислительной сложностью по сравнению с венгерским алгоритмом и практически не уступает в точности ее определения.

Таблица 11

Время вычисления меры различия по венгерскому алгоритму C_{min} и авторскому $AD(T_i, T_k)$, мс

Table 11

The time for calculating the measure of difference according to the Hungarian algorithm C_{min} and the author's $AD(T_i, T_k)$, ms

n	T_j	T_k	Венгерский алгоритм <i>Hungarian algorithm</i>	Авторский алгоритм <i>Author's algorithm</i>
8	11110000	00001111	0,33235	0,01918
	10101010	01010101	0,11701	0,01711
9	101010101	010101010	0,43508	0,01820
10	1111100000	0000011111	0,70302	0,02541
	1010101010	0101010101	0,15418	0,02301
11	10101010101	01010101010	0,55974	0,02512
12	111111000000	000000111111	1,06907	0,03751
	101010101010	010101010101	0,20669	0,03029
13	1010101010101	0101010101010	0,79060	0,03230
14	11111110000000	00000001111111	1,76948	0,04371
	10101010101010	01010101010101	0,25969	0,03858
15	101010101010101	010101010101010	1,07059	0,04072
16	1111111100000000	0000000011111111	2,39983	0,05518
	1010101010101010	0101010101010101	0,32111	0,04761

Заключение. Исследована мера различия тестовых наборов для управляемого вероятностного тестирования, основанная на вычислении суммы расстояний между совпадающими данными двух тестовых наборов. Показано, что расстояния между тождественными данными можно представить в виде двухмерной матрицы их величин, которая аналогична исходной квадратной матрице стоимости, используемой в классической задаче о назначениях. Отмечено, что наиболее близким известным решением задачи о назначениях является венгерский алгоритм, характеризующийся значительной вычислительной сложностью.

Как альтернатива венгерскому алгоритму предложен авторский алгоритм вычисления меры различия и приведена оценка его вычислительной сложности. Данный алгоритм, имеющий меньшую вычислительную сложность, позволяет получить значения меры различия, аналогичные венгерскому алгоритму, либо отличие значений, полученное с использованием обоих алгоритмов, будет незначительным. Проведенные экспериментальные исследования показали высокую временную эффективность и точность вычислений авторского алгоритма в сравнении с известными решениями.

Дальнейшие исследования целесообразно расширить в части свойств новой меры различия и ее применимости для различных прикладных задач.

Вклад авторов. Ярмолик В. Н. предложил алгоритм вычисления меры различия для тестовых наборов. Петровская В. В. провела экспериментальные исследования. Мрозек И. принял участие в обобщении и анализе полученных результатов.

Список использованных источников

1. An orchestrated survey on automated software test case generation / S. Anand [et al.] // J. of Systems and Software. – 2014. – Vol. C-39, no. 4. – P. 582–586.
2. Malaiya, Y. K. The coverage problem for random testing / Y. K. Malaiya, S. Yang // Proc. of ITC, Philadelphia, PA, USA, Oct. 1984. – Philadelphia, 1984. – P. 237–242.
3. A survey on adaptive random testing / R. Huang [et al.] // IEEE Transactions on Software Engineering. – 2021. – Vol. 47, no. 10. – P. 2052–2083.
4. Using the information: Incorporating positive feedback information into the testing process / K. P. Chan [et al.] // Proc. of the 7th Annual Intern. Workshop on Software Technology and Engineering Practice (STEP'03), Amsterdam, Netherlands, 19–21 Sept. 2003. – Amsterdam, 2003. – P. 71–76.

5. A preliminary study of adaptive random testing techniques / M. S. Roslina [et al.] // *Intern. J. of Information Technology & Computer Science*. – 2015. – Vol. 19, no. 1. – P. 116–127.
6. An empirical comparison of combinatorial testing, random testing and adaptive random testing / H. Wu [et al.] // *IEEE Transactions on Software Engineering*. – 2020. – Vol. 46, no. 3. – P. 302–320.
7. Ярмолик, В. Н. Контроль и диагностика вычислительных систем / В. Н. Ярмолик. – Минск : Бест-принт, 2019. – 387 с.
8. Ярмолик, В. Н. Многократные управляемые вероятностные тесты / В. Н. Ярмолик, В. А. Леванцевич, И. Мрозек // *Информатика*. – 2015. – № 2(46). – С. 63–76.
9. Mrozek, I. Multiple controlled random testing / I. Mrozek, V. N. Yarmolik // *Fundamenta Informaticae*. – 2016. – Vol. 144, no. 1. – P. 23–43.
10. Садовский, М. Г. О сравнении символьных последовательностей / М. Г. Садовский // *Вычислительные технологии*. – 2005. – № 3(10). – С. 106–116.
11. Ярмолик, В. Н. Мера различия для управляемых вероятностных тестов / В. Н. Ярмолик, Н. А. Шевченко, В. В. Петровская // *Доклады БГУИР*. – 2022. – № 6(20). – С. 52–60.
12. A modified similarity metric for unit testing of object-oriented software based on adaptive random testing / J. Chen [et al.] // *Intern. J. of Software Engineering and Knowledge Engineering*. – 2019. – Vol. 29, no. 4. – P. 577–606.
13. Chen, T, Y. Adaptive random testing based on distribution metrics / T. Y. Chen, F-C. Kuo, H. Liu // *J. of Systems and Software*. – 2009. – Vol. 82, no. 9. – P. 1419–1433.
14. Bard, G. V. Spelling-error tolerant, order-independent pass-phrases via the Damerau – Levenshtein string-edit distance metric / G. V. Bard // *Proc. of the Fifth Australasian Symp. on ACSW Frontiers*, Ballarat, Australia, 30 Jan. – 2 Feb. 2007. – Ballarat, 2007. – P. 117–124.
15. Tannga, M. J. Comparative analysis of Levenshtein distance algorithm and Jaro Winkler for text plagiarism detection application / M. J. Tannga, S. Rahman, Hasniati // *J. of Technology Research in Information System and Engineering*. – 2017. – Vol. 4, no. 2. – P. 44–54.
16. О мерах сходства расположения компонентов в массивах естественно упорядоченных данных / А. С. Гуменюк [и др.] // *Тр. СПИИРАН*. – 2019. – Vol. 18, no. 2. – P. 471–503.
17. Ярмолик, В. Н. Адресные последовательности для многократных маршевых тестов / В. Н. Ярмолик, С. В. Ярмолик // *Автоматика и вычислительная техника*. – 2006. – № 5. – С. 59–68.
18. Savage, C. A survey of combinatorial Gray code / C. Savage // *SIAM Review*. – 1997. – Vol. 3, no. 4. – P. 605–629.
19. Mrozek, I. Transparent memory tests based on the double address sequences / I. Mrozek, V. N. Yarmolik // *Entropy*. – 2021. – No. 23. – P. 894.
20. Неразрушающие тесты с четным повторением адресов для запоминающих устройств / В. Н. Ярмолик [и др.] // *Информатика*. – 2021. – Т. 18, № 3. – С. 18–35.
21. Kuhn, H. W. Variants of the Hungarian method for assignment problems / H. W. Kuhn // *Naval Research Logistics Quarterly*. – 1956. – Vol. 3, no. 4. – P. 253–258.
22. Munkres, J. Algorithms for the assignment and transportation problems / J. Munkres // *J. of the Society for Industrial and Applied Mathematics*. – 1957. – Vol. 5, no. 1. – P. 32–38.
23. Blum, C. Metaheuristics in combinatorial optimization: Overview and conceptual comparison / C. Blum, A. Roli // *ACM Computing Surveys*. – 2003. – Vol. 35, no. 3. – P. 268–308.

References

1. Anand S., Burke E. K., Chen T. Y., Clark J., Cohen M. B., ..., Zhu H. An orchestrated survey on automated software test case generation. *Journal of Systems and Software*, 2014, vol. C-39, no 4, pp. 582–586.
2. Malaiya Y. K., Yang S. The coverage problem for random testing. *Proceedings of the International Test Conference, Philadelphia, PA, USA, October 1984*. Philadelphia, 1984, pp. 237–242.
3. Huang R., Sun W., Xu Y., Chen H., Towey D., Xia X. A survey on adaptive random testing. *IEEE Transactions on Software Engineering*, 2021, vol. 47, no 10, pp. 2052–2083.
4. Chan K. P., Towey D., Chen T. Y., Kuo F.-C., Merkel R. G. Using the information: Incorporating positive feedback information into the testing process. *Proceedings of the 7th Annual International Workshop on Software Technology and Engineering Practice (STEP'03), Amsterdam, Netherlands, 19–21 September 2003*. Amsterdam, 2003, pp. 71–76.
5. Roslina M. S., Ghani A. A. A., Baharom S., Zulzazil H. A preliminary study of adaptive random testing techniques. *International Journal of Information Technology & Computer Science*, 2015, vol. 19, no. 1, pp. 116–127.

6. Wu H., Nie C., Petke J., Jia Y., Harman M. An empirical comparison of combinatorial testing, random testing and adaptive random testing. *IEEE Transactions on software engineering*, 2020, vol. 46, no. 3, pp. 302–320.
7. Yarmolik V. N. Control' i diagnostika vuchislitel'nuch system. *Computer Systems Testing and Diagnoses*. Minsk, Bestprint, 2019, 387 p. (In Russ.).
8. Yarmolik V. N., Levantsevich B. A., Mrozek I. *Multiple controlled random tests*. Informatika [Informatics], 2015, no. 2, pp. 63–76 (In Russ.).
9. Mrozek I., Yarmolik V. N. Multiple controlled random testing. *Fundamenta Informaticae*, 2016, vol. 144, no. 1, pp. 23–43.
10. Sadovskii M. G. *About symbolical sequences comparigion*. Vuchislitel' nue technologii [Computational Technologise], 2005, no. 3(10), pp. 106–116 (In Russ.).
11. Yarmolik V. N., Shauchenka M. A., Petrovskaya V. V. Distance measure for controlled random tests. *Doklady BGUIR [BSUIR Proceedings]*, 2022, no. 6(20), pp. 52–60 (In Russ.).
12. Chen J., Kudjo P. K., Zhang Z., Su C., Guo Y., ..., Song H. A modified similarity metric for unit testing of object-oriented software based on adaptive random testing. *International Journal of Software Engineering and Knowledge Engineering*, 2019, vol. 29, no. 4, pp. 577–606.
13. Chen T. Y., Kuo F.-C., Liu H. Adaptive random testing based on distribution metrics. *Journal of Systems and Software*, 2009, vol. 82, no. 9, pp. 1419–1433.
14. Bard G. V. Spelling-error tolerant, order-independent pass-phrases via the Damerau – Levenshtein string-edit distance metric. *Proceedings of the Fifth Australasian Symposium on ACSW Frontiers, Ballarat, Australia, 30 January – 2 February 2007*. Ballarat, 2007, pp. 117–124.
15. Tangga M. J., Rahman S., Hasniati. Comparative analysis of Levenshnein distance algorithm and Jaro Winkler for text plagiarism detection application. *Journal of Technology Research in Information System and Engineering*, 2017, vol. 4, no. 2, pp. 44–54.
16. Gumenyuk A. S., Skiba A. A., Pozdnichenko N. N., Shpynov S. N. *About similarity measures of components of naturally ordered data arrays*. Trudy SPIIRAN [SPIIRAS Proceedings], 2019, vol. 18, no. 2, pp. 471–503 (In Russ.).
17. Yarmolik V. N., Yarmolik S. V. *Address sequences for multiple march tests*. Avtomatika i vuchislitel'naya tehnika [Automation and Computer Science], 2006, no 5, pp. 59–68 (In Russ.).
18. Savage C. A survey of combinatorial Gray code. *SIAM Review*, 1997, vol. 3, no. 4, pp. 605–629.
19. Mrozek I., Yarmolik V. N. Transparent memory tests based on the double address sequences. *Entropy*, 2021, no. 23, pp. 894.
20. Yarmolik V. N., Mrozek I., Levantsevich V. A., Demenkovets D. V. *Transparent memory tests with even repeating addresses for storage devices*. Informatika [Informatics], 2021, vol. 18, no. 3, pp. 18–35 (In Russ.).
21. Kuhn H. W. Variants of the Hungarian method for assignment problems. *Naval Research Logistics Quarterly*, 1956, vol. 3, no. 4, pp. 253–258.
22. Munkres J. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 1957, vol. 5, no. 1, pp. 32–38.
23. Blum C., Roli A. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys*, 2003, vol. 35, no. 3, pp. 268–308.

Информация об авторах

Ярмолик Вячеслав Николаевич, доктор технических наук, профессор, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: yarmolik10ru@yahoo.com

Петровская Вита Владленовна, магистр технических наук, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: vita.petrovskaya@gmail.com

Мрожек Иренеуш, доктор, адъюнкт, Белостокский технический университет.
E-mail: i.mrozek@pb.edu.pl

Information about the authors

Vyacheslav N. Yarmolik, D. Sc. (Eng.), Professor, Belarusian State University of Informatics and Radioelectronics.
E-mail: yarmolik10ru@yahoo.com

Vita V. Petrovskaya, M. Sc. (Eng.), Belarusian State University of Informatics and Radioelectronics.
E-mail: vita.petrovskaya@gmail.com

Ireneusz Mrozek, D. Sc., Lecture, Bialystok University of Technology.
E-mail: i.mrozek@pb.edu.pl

ЗАЩИТА ИНФОРМАЦИИ И НАДЕЖНОСТЬ СИСТЕМ INFORMATION PROTECTION AND SYSTEM RELIABILITY



УДК 681.324
<https://doi.org/10.37661/1816-0301-2022-19-4-27-41>

Оригинальная статья
Original Paper

Построение и балансировка путей физически неклонируемой функции типа арбитр на FPGA

А. Ю. Шамына[✉], А. А. Иванюк

Белорусский государственный университет
информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь
[✉]E-mail: shamyuna@bsuir.by

Аннотация

Цели. Решается задача построения новой структуры путей физически неклонируемой функции типа арбитр (АФНФ) на FPGA (Field programable gate array), основанных на полном использовании внутренних ресурсов LUT-блоков (англ. Look up table), которые функционально являются повторителями. Актуальность исследования связана с бурным развитием средств физической криптографии. Также преследуется цель разработки способа устранения асимметрии путей АФНФ, связанной с особенностью синтеза подобных схем на FPGA.

Методы. Используются методы синтеза цифровых устройств, их параметрического моделирования и реализации на платах быстрого прототипирования. Для измерения внутренних задержек распространения сигналов через пути АФНФ применяется схема кольцевого осциллятора.

Результаты. Предложена новая структура базового элемента путей АФНФ с использованием двух функциональных повторителей. Продемонстрирована необходимость балансировки задержек путей АФНФ. Разработан способ устранения асимметрии распространения сигналов через пути АФНФ на базе управляемых линий задержки. Показаны недостатки использования в качестве схемы арбитра АФНФ классических подходов и необходимость их модификации.

Заключение. Предложенный подход к построению путей АФНФ показал свою состоятельность и перспективность. Экспериментально подтверждается улучшение характеристик АФНФ, построенных по предложенным способом, а также снижение аппаратных затрат при их реализации по сравнению с классическими схемами АФНФ. Представляется перспективным дальнейшее развитие описанного подхода АФНФ, связанное прежде всего с усовершенствованием структуры арбитра.

Ключевые слова: физическая криптография, физически неклонируемые функции типа арбитр, симметричные пути, управляемые линии задержки, кольцевой осциллятор

Благодарности. Авторы выражают благодарность резиденту Парка высоких технологий компании SK Hynix Memory Solutions Eastern Europe за предоставленное оборудование для проведения экспериментальных исследований в рамках работы совместной учебной лаборатории с Белорусским государственным университетом информатики и радиоэлектроники.

Для цитирования. Шамына, А. Ю. Построение и балансировка путей физически неклонированной функции типа арбитр на FPGA / А. Ю. Шамына, А. А. Иванюк // Информатика. – 2022. – Т. 19, № 4. – С. 27–41. <https://doi.org/10.37661/1816-0301-2022-19-4-27-41>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 22.07.2022

Подписана в печать | Accepted 14.09.2022

Опубликована | Published 29.12.2022

Creating and balancing the paths of arbiter-based physically unclonable functions on FPGA

Artsiom Yu. Shamyna[✉], Alexander A. Ivaniuk

*Belarusian State University of Informatics and Radioelectronics,
st. P. Brovki, 6, Minsk, 220013, Belarus*

[✉]E-mail: shamyna@bsuir.by

Abstract

Objectives. The problem of constructing a new structure of paths of physically unclonable function of the arbiter type (APUF) on the FPGA is being solved, based on the full use of internal resources of LUT-blocks, which are functionally repeaters. The relevance of the study is associated with the rapid development of physical cryptography tools. Another goal is the developing a methodology for eliminating the asymmetry of the APUF paths associated with the peculiarity of the synthesis of such circuits on the FPGA.

Methods. The methods of synthesis of digital devices, their parametric modeling and implementation on rapid prototyping boards are used. A ring oscillator circuit is used to measure the internal propagation delays of signals through the APUF paths.

Results. A new structure of the basic element of APUF paths with the use of two functional repeaters is proposed. The necessity of balancing the delays of APUF paths is demonstrated. A technique has been developed to eliminate the asymmetry of signal propagation through APUF paths based on controlled delay lines. The disadvantages of classical approaches as an APUF arbitrator and the need for their modification are shown.

Conclusion. The proposed approach to build APUF paths has shown its viability and promise. An improvement in the characteristics of APUF constructed according to the proposed method, as well as a reduction in hardware costs during their implementation compared to classical APUF schemes, is experimentally confirmed. It seems promising to develop the described methodology for constructing the APUF to improve the structure of the arbiter.

Keywords: physical cryptography, arbiter-based physically unclonable functions, symmetrical paths, propagation delay line, ring oscillator

Acknowledgments. The authors express gratitude to the HTP resident of the "SK Hynix Memory Solutions Eastern Europe" company for the equipment provided for carrying out experiments within the framework of the joint laboratory with the Belarusian State University of Informatics and Radioelectronics.

For citation. Shamyna A. Yu., Ivaniuk A. A. *Creating and balancing the paths of arbiter-based physically unclonable functions on FPGA*. Informatika [Informatics], 2022, vol. 19, no. 4, pp. 27–41 (In Russ.). <https://doi.org/10.37661/1816-0301-2022-19-4-27-41>

Conflict of interest. The authors declare of no conflict of interest.

Введение. В настоящее время все большая роль отводится средствам физической криптографии, где одним из наиболее популярных направлений является изучение физически неклонированных функций (ФНФ) [1]. Основополагающая идея ФНФ заключается в извлечении харак-

теристик, свойственных конкретной физической системе и являющихся уникальными и неповторяемыми, но при этом достаточно стабильными и удовлетворяющими определенным критериям при их многократном извлечении. Большинство ФНФ, реализованных в составе цифровых устройств, которые идентичны с точки зрения проектного описания и технологии изготовления интегральных схем, основаны на вариативности задержек распространения сигналов по фиксированным путям. Это свойство обусловлено естественными флуктуациями в материалах, используемых при производстве данных устройств, а также некоторым несовершенством производственного процесса.

Популярным схемотехническим решением, позволяющим на основе уникальности задержек распространения сигналов по топологически симметричным путям различных экземпляров одного устройства генерировать битовую последовательность для некоторого множества фиксированных запросов, является использование так называемых физически неклонированных функций типа арбитр [2, 3].

Классическая схема АФНФ. В классической схеме АФНФ подразумевается наличие генератора тестового сигнала (ГТС), блока симметричных путей (БСП) и арбитра (АРБ). Схема арбитра позволяет определить очередность прохождения фронтов тестового импульса через блок симметричных путей и выработать на этой базе ответ R (рис. 1). В свою очередь, БСП представляет собой последовательно соединенные звенья (ЗВ), которые, как правило, строятся с помощью двух мультиплексоров с конфигурацией 2×1 и обеспечивают прямую либо перекрестную передачу двух тестовых сигналов в зависимости от значения разряда запроса.

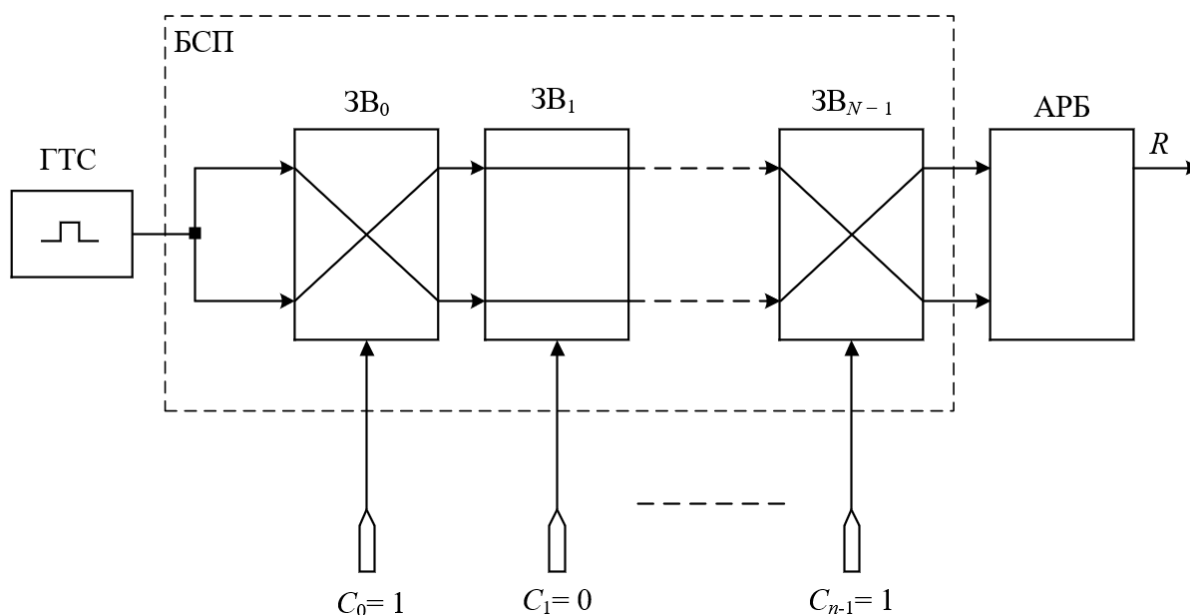


Рис. 1. Классическая схема АФНФ

Fig. 1. Classic APUF scheme

Общим подходом для построения звена $ЗВ_j$ является схема, состоящая из двух мультиплексоров (рис. 2, а). При реализации звена АФНФ на перепрограммируемых логических интегральных схемах типа FPGA будут использованы два блока LUT3 (рис. 2, б).

Вместе с тем при реализации АФНФ на современных FPGA, таких как Artix 7 фирмы Xilinx, возникает ситуация неполного использования ресурсов LUT-компонентов. Так, для реализации звена пути классической АФНФ требуются два LUT3, хотя фактически используются два технологических LUT6 и значительная часть их ресурсов остается незадействованной. Потенциально более полное использование ресурсов предоставляемых технологических компонентов FPGA может значительно сократить совокупные аппаратные затраты при реализации АФНФ и улучшить их характеристики [4].

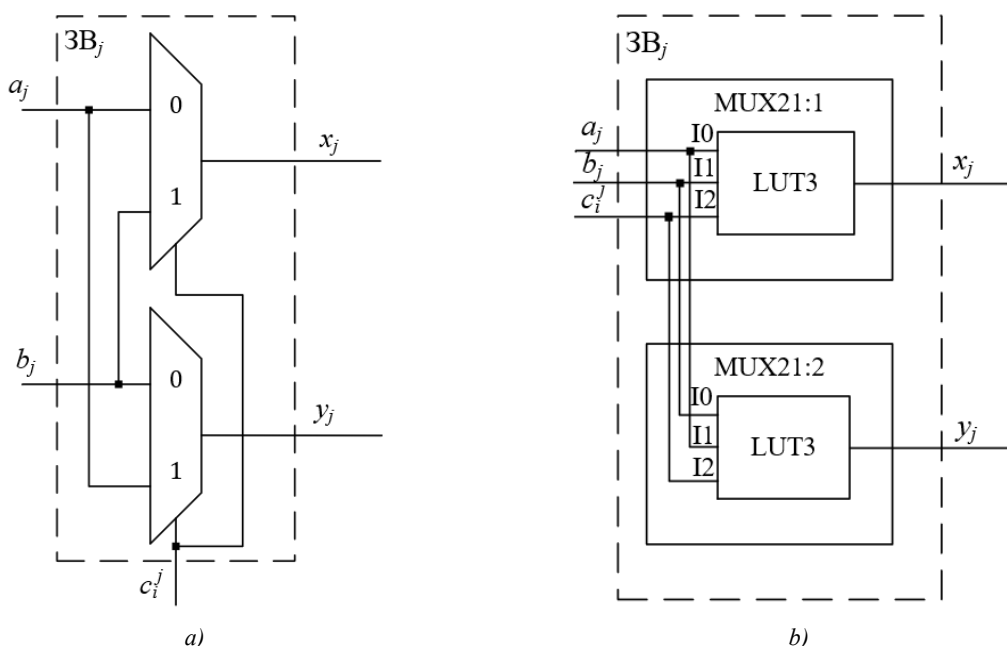


Рис. 2. Реализация звена БСП классических АФНФ: а) RTL-схема; б) технологическая схема

Fig. 2. Implementation of the BSP link of classical APUF: a) RTL schematic; b) technology schematic

Схематично компонент LUT представляет собой память конфигурации и каскад мультиплексоров, обеспечивающий трансляцию единственного выбранного значения из этой памяти на выход схемы в зависимости от значений сигналов на адресных входах. На рис. 3 представлена реализация структурной схемы блока LUT4 FPGA фирмы Xilinx серии Spartan-3E, сконфигурированного для реализации одного мультиплексора классической АФНФ.

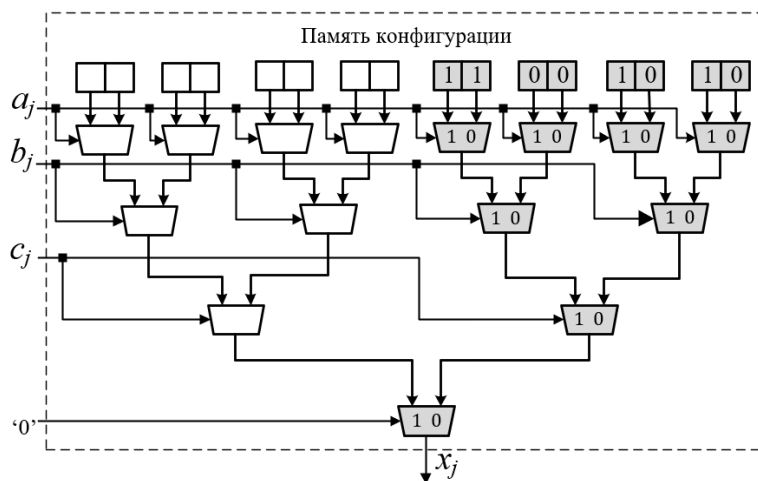


Рис. 3. Реализация мультиплексора звена БСП на LUT4

Fig. 3. Implementation of the BSP link multiplexer on LUT4

На рис. 3 видно, что ресурсы LUT4 задействованы лишь наполовину. Как правило, в стратегиях синтеза цифровых устройств не предусмотрено применение тех LUT-компонентов, которые уже были использованы, несмотря на то что значительная часть их ресурсов остается незадействованной, а при применении LUT-компонентов большей разрядности доля неиспользуемых ресурсов LUT-блоков становится еще более значительной. Так, в случае применения физического компонента LUT6 при реализации классической АФНФ для одного мультиплексора звена БСП будет израсходована лишь 1/8 доступных ресурсов.

Кроме того, как это было отмечено в работе [5], построение АФНФ на базе двух независимых путей может быть более предпочтительным с точки зрения статистических свойств, чем построение с помощью классического подхода, что обусловлено меньшей взаимной зависимостью задержек распространения сигналов через звенья пути.

Синтез предлагаемой архитектуры симметричных путей. Следует отметить, что при классическом подходе к построению АФНФ осуществляется неполное применение внутренних ресурсов LUT-блоков.

Идея более полного расхода доступных ресурсов LUT при реализации БСП АФНФ рассмотрена в работе [6]. В ней описан подход к построению БСП с использованием блоков функциональных мультиплексоров, копии которого размещены на одном LUT, а часть запроса, соответствующая одному звену БСП, отвечает за выбор не только прямой либо перекрестной передачи сигналов через звено, но и одной из двух копий мультиплексора, ее осуществляющей.

В настоящей работе в качестве звена БСП (рис. 4) предлагается задействовать схему двух функциональных повторителей, которые будут полностью расходовать ресурсы LUTN и обеспечивать 2^{N-1} уникальные трансляции в зависимости от $(N-1)$ -разрядного запроса. Уникальность задержки сигнала при этом объясняется отличием пути прохождения сигнала непосредственно внутри самого LUT-блока в зависимости от значений сигналов на его адресных входах. Таким образом, при реализации, например, 128-разрядной АФНФ с применением компонентов LUT4 предложенная архитектура БСП позволяет сократить использование технологических LUT-компонентов в три раза при неизменной мощности множества запросов и ответов. Конфигурация LUT4, синтезированная по данной схеме, изображена на рис. 5.

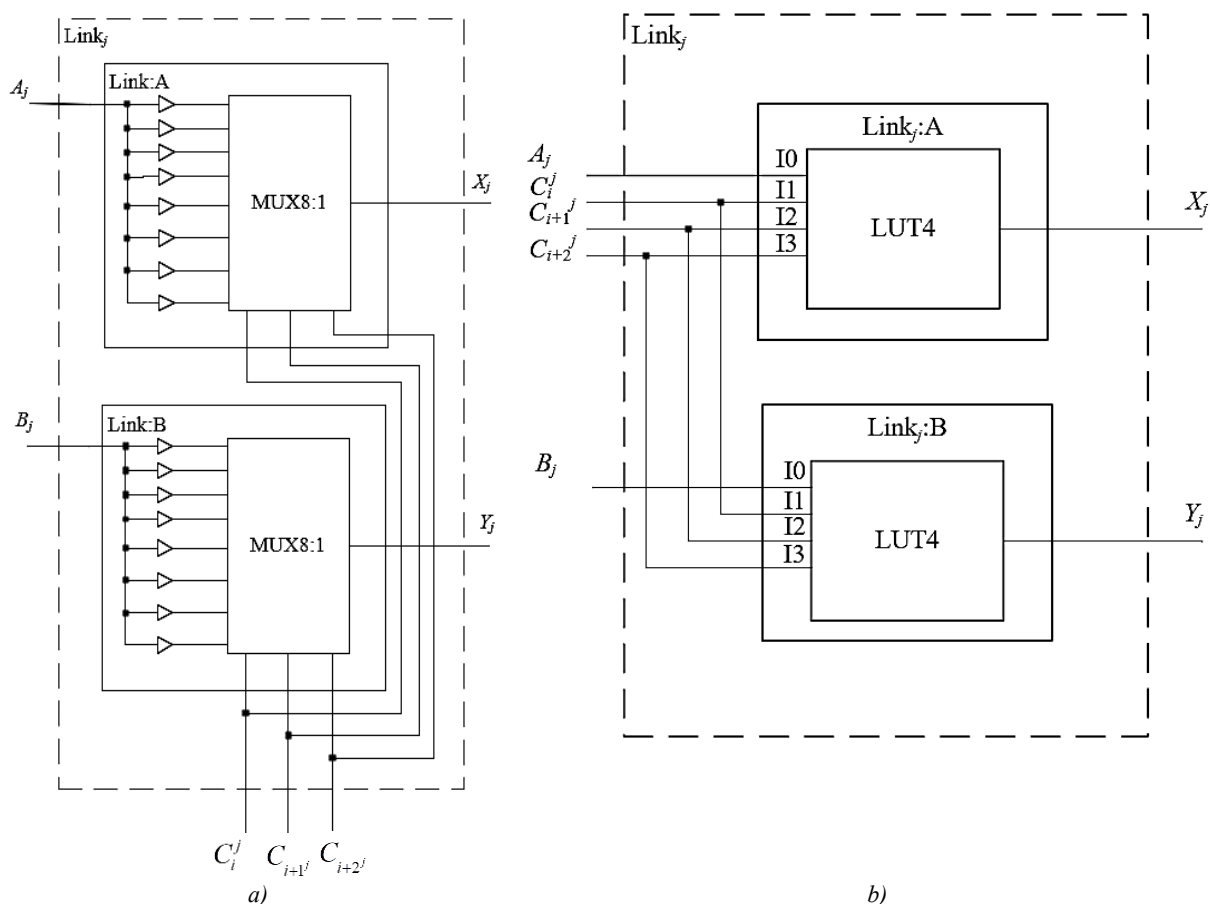


Рис. 4. Реализация одного звена БСП предложенной структуры: а) RTL-синтез; б) технологическая схема
 Fig. 4. Implementation of one link of BSP of proposed structure: a) RTL schematic; b) technology schematic

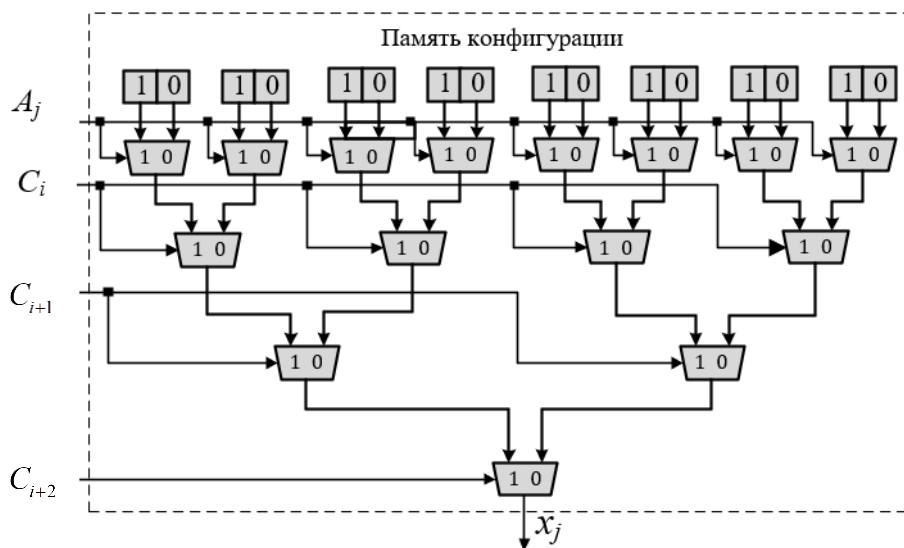


Рис. 5. Конфигурация LUT4 по предложенной структуре звена БСП

Fig. 5. LUT4 configuration according to proposed BSP link structure

Рассматриваемая схема полностью использует доступные ресурсы блока LUT и легко масштабируется под любую размерность LUT. Ее потенциальным недостатком является малое взаимное отличие проходимых сигналом путей при некоторых запросах. Так, при значениях векторов запроса $C_0C_1C_2 = 000$ и $C_0C_1C_2 = 100$ различие проходимых путей будет лишь в одном мультиплексоре. Данный факт требует дополнительных исследований.

Построение экспериментальной установки. Следует отметить важность построения симметричных путей АФНФ как источника для извлечения характеристик задержек распространения сигналов через пары путей, которые несут уникальный и случайный характер. В связи с асимметрией путей АФНФ могут значительно ухудшаться характеристики случайности и уникальности. Связано это прежде всего с тем, что ответ АФНФ в таком случае будет определен не уникальными свойствами временных задержек своего экземпляра, а асимметрией двух путей, обусловленной асимметричными межсоединениями их звеньев. Это может привести к тому, что в конечном итоге для всех изготовленных по идентичному проекту АФНФ при фиксированном запросе будет преобладать одинаковый ответ. Данная асимметрия межсоединений, как правило, не может быть перекрыта разницей задержек между конкретными экземплярами АФНФ на различных кристаллах и особенно для сравнительно небольшого количества звеньев путей АФНФ.

Ввиду неуправляемости автоматизированного синтеза на ПЛИС типа FPGA достаточно проблематично построить идеально симметричные пути с ее использованием. Получаемую при этом асимметрию можно условно разделить на асимметрию внутри конфигурационных блоков, которая в целом может быть нивелирована ручным размещением технологических компонентов, и на асимметрию межсоединений. Для минимизации асимметрии внутри конфигурационных блоков LUT-компоненты, которые являются составными звеньями БСП, были размещены особым образом [7]. Для этого использовались команды ограничения САПР Vivado LOC и BEL.

Для оценки временных характеристик распространения сигналов через симметричные пути АФНФ, построенные по предложенной структуре, ввиду невозможности временных измерений задержек внутри кристалла ПЛИС был использован подход на базе кольцевого осциллятора (КО), который подразумевает охват исследуемого пути отрицательной обратной связью. Данный подход требует наличия схемы управления режимом осцилляции, а также регистрации формируемой схемой импульсной последовательности (рис. 6).

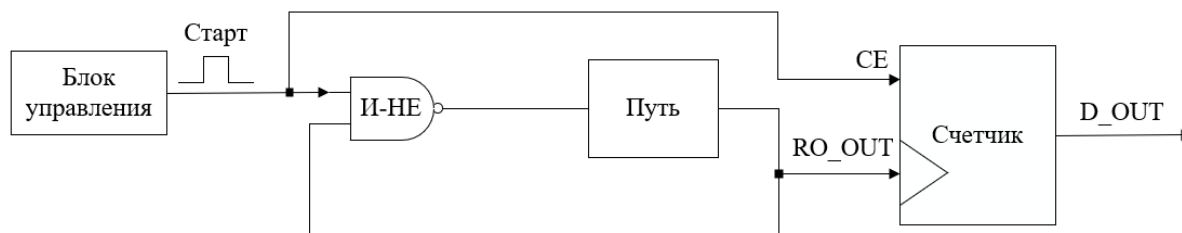


Рис. 6. Схема измерения задержек распространения сигналов на основе КО

Fig. 6. Scheme for measuring propagation delays based on QoS

Для достижения большей достоверности измерений было принято решение применить одну цепь обратной связи для измерения задержек двух симметричных путей. Для этого в последнем SLICE-блоке БСП был использован компонент F8MUX, выход которого соединен с цепью обратной связи, а два входа – с последними звеньями двух симметричных путей БСП. Таким образом, благодаря топологической и структурной схожести соединений внутри SLICE удалось свести к минимуму различия в измеряемых путях, непосредственно не относящиеся к ним.

Проектное описание экспериментальной установки, в которой помимо самостоятельно описанных модулей широко использовались IP-ядра и софт-процессор Microblaze, было создано в САПР Vivado 2018.2 с помощью языка VHDL. Эксперимент проводился на пяти идентичных платах быстрого прототипирования Digilent Nexys 4 с FPGA Artix 7. Общая схема эксперимента аналогична используемой в работе [8].

Экспериментальные исследования. Предложенная структура звена базируется на предположении об уникальности распространения задержек сигнала от входа до выхода LUT-блока в зависимости от значения запроса. Для реализации АФНФ на основе новой структуры звена БСП требуется обеспечить уникальность задержек в рамках как одного LUT-блока, так и нескольких LUT-блоков при соответствующих запросах. Для проверки предположения изначально был проведен следующий эксперимент. В матрице FPGA Artix 7 был выбран единственный SLICE-блок, в котором все доступные четыре LUT6-блока были сконфигурированы предложенным способом. Учитывая, что для измерения задержек была задействована схема КО, для увеличения достоверности измерений использовались фиксированная цепь обратной связи и схема измерений. Выбранный SLICE-блок был подключен к цепи измерений через выход внутреннего неконфигурируемого мультиплексора F8MUX. Также для коммутации LUT-блоков внутри SLICE были использованы два мультиплексора F7MUX. Данный способ измерения задержек для LUT-блоков, размещенных в одном SLICE-блоке, обоснован структурным подобием внутренних соединений внутри SLICE. Полученные таким способом экспериментальные данные должны максимально достоверно отражать уникальные задержки LUT6-блоков (рис. 7). Пять младших разрядов запроса C использовались как значения для адресных входов LUT-блоков, а два старших – как селектирующий сигнал для мультиплексоров, значения которых позволяют коммутировать выход одного из четырех LUT-блоков на выход SLICE.

Измеренные значения задержек распространения сигналов для всех LUT-компонентов SLICEL (X33Y90) представлены в табл. 1. Согласно данным табл. 1 задержки для каждого запроса являются уникальными. Значения среднеквадратического отклонения составили: $\sigma_{A6} = 0,004912384$, $\sigma_{B6} = 0,006597065$, $\sigma_{C6} = 0,001980557$, $\sigma_{D6} = 0,00171887$.

Затем эксперимент для выбранной конфигурации был повторен 100 раз на пяти кристаллах. Результаты продемонстрировали, что все полученные измерения задержек являются уникальными.

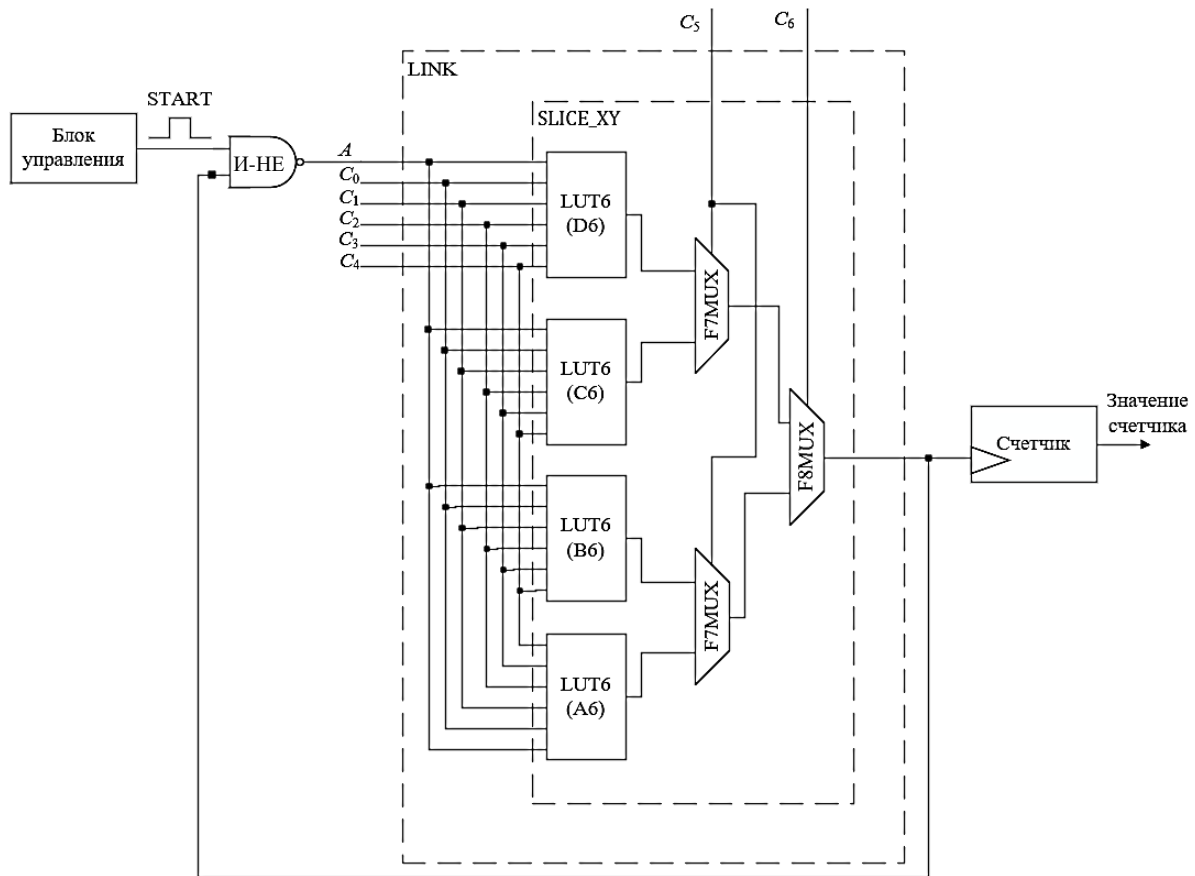


Рис. 7. Конфигурация SLICE-блока для эксперимента

Fig. 7. SLICE configuration for the experiment

Таблица 1
Значения задержек распространения сигналов через LUT-компоненты SLICEL, псTable 1
Propagation delay measurements through SLICEL LUTs, ps

Запрос Challenge	A6LUT	B6LUT	C6LUT	D6LUT
00000	1602,60	1086,39	1249,22	1548,14
00001	1601,53	1085,79	1249,08	1547,87
00010	1599,99	1084,92	1248,77	1548,07
00011	1603,18	1087,72	1248,64	1547,97
00100	1602,73	1087,74	1243,53	1549,38
00101	1601,51	1087,01	1243,35	1549,18
00110	1599,80	1087,12	1243,66	1552,14
00111	1602,46	1089,62	1243,42	1552,08
01000	1608,77	1081,38	1248,39	1550,58
01001	1607,54	1083,79	1248,48	1550,69
01010	1607,21	1084,76	1247,48	1551,10
01011	1611,55	1087,28	1247,43	1551,01
01100	1605,92	1088,65	1248,38	1550,23
01101	1609,22	1088,42	1248,17	1550,28
01110	1607,68	1083,32	1250,88	1551,86
01111	1607,48	1083,24	1250,65	1551,51
10000	1597,86	1069,96	1247,83	1547,60
10001	1597,86	1072,68	1247,79	1547,52

Окончание табл. 1

End of table 1

Запрос Challenge	A6LUT	B6LUT	C6LUT	D6LUT
10010	1594,46	1069,59	1247,36	1545,83
10011	1598,61	1073,16	1247,19	1545,58
10100	1595,90	1074,93	1247,22	1547,71
10101	1598,68	1075,67	1246,97	1547,70
10110	1597,31	1075,30	1244,97	1550,63
10111	1599,71	1078,72	1245,05	1550,72
11000	1592,80	1074,12	1249,15	1550,38
11001	1594,74	1073,70	1248,91	1550,49
11010	1595,77	1072,92	1247,42	1548,84
11011	1599,64	1073,27	1247,44	1548,54
11100	1599,53	1074,72	1247,25	1549,27
11101	1597,07	1074,80	1247,42	1549,07
11110	1596,70	1073,40	1249,47	1548,43
11111	1597,17	1074,21	1249,17	1548,23

Эксперимент был повторен и для других SLICE. Графики измеренных задержек для двух SLICE(X0Y198, X0Y199) представлены на рис. 8.

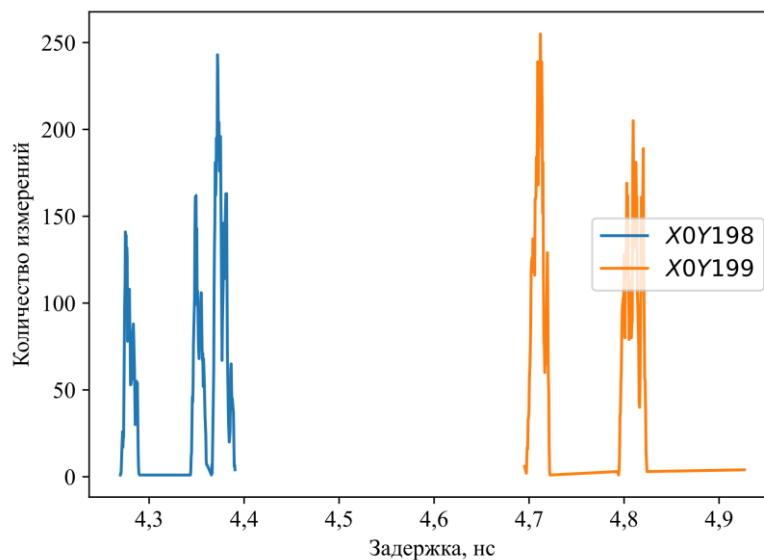


Рис. 8. Временное распределение задержек для двух SLICE

Fig. 8. Time distribution of delays for two SLICES

Явный взаимный временной сдвиг графиков задержек для нескольких SLICE может говорить об отличиях в задержках как по линии отрицательной обратной связи схемы измерения, так и по линии распространения сигналов непосредственно в SLICE-блоках.

Полученные результаты демонстрируют целесообразность исследования задержек распространения сигналов через симметричные пути, построенные на базе звеньев предложенной структуры. Для этих целей было создано соответствующее VHDL-описание двух симметричных путей (рис. 9) для разрядности запроса $N = 64$. Для достижения максимальной симметричности путей в проектном описании LUT-блоки одного звена размещались в одном SLICE в соответствии с подходом, описанным в работе [7]. Генерирование запросов осуществлялось с помощью генератора M-последовательности на базе LFSR, всего было сгенерировано

$C = 10^6$ запросов. Следует также отметить, что, как и в предыдущих экспериментах, для измерений применялась фиксированная схема для двух путей. Результаты измерения временных задержек через симметричные пути для данного эксперимента представлены на рис. 10.

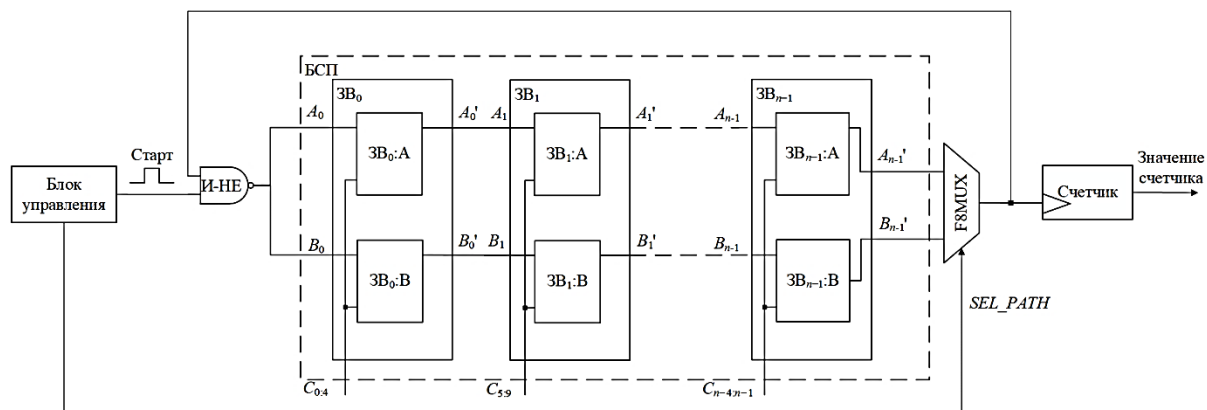


Рис. 9. Схема эксперимента

Fig. 9. Experiment scheme

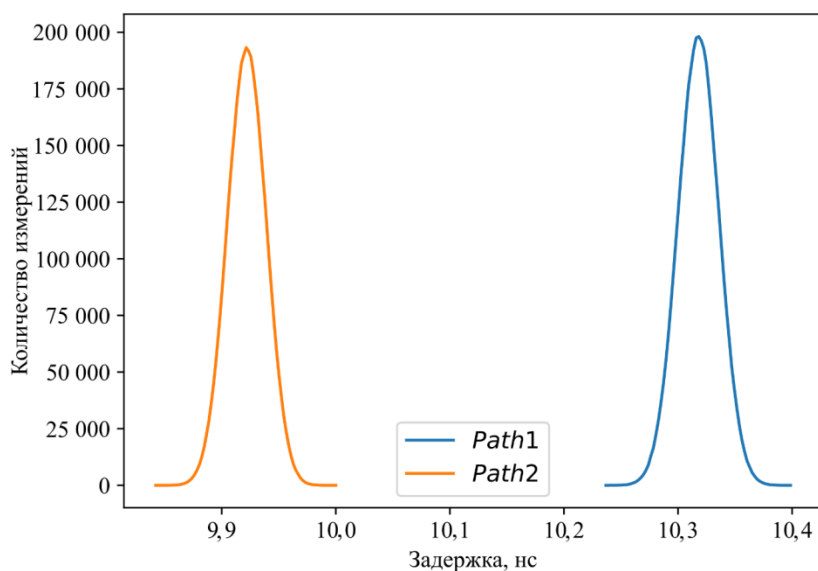


Рис. 10. Временное распределение задержек для двух симметричных путей при $N = 64$

Fig. 10. Time distribution of delays for two symmetrical paths with $N = 64$

Полученные результаты свидетельствуют о сдвиге по временной оси графиков задержек для двух симметричных путей и обусловлены прежде всего заведомой неуправляемостью автоматизированного построения межсоединений SLICE-блоков, а также уникальностью и неповторимостью значений задержек распространения сигналов через технологические компоненты FPGA на кристалле. Для построения АФНФ на их основе необходимо, чтобы временные задержки двух путей лежали в одном временном интервале. Достичь этого для текущей конфигурации можно при помощи управляемых линий задержек, базирующихся на последовательно соединенных мультиплексорах и встроенных между последним звеном БСП и потенциальным арбитром. Предлагаемое решение схемы балансировки представлено на рис. 11 и 12.

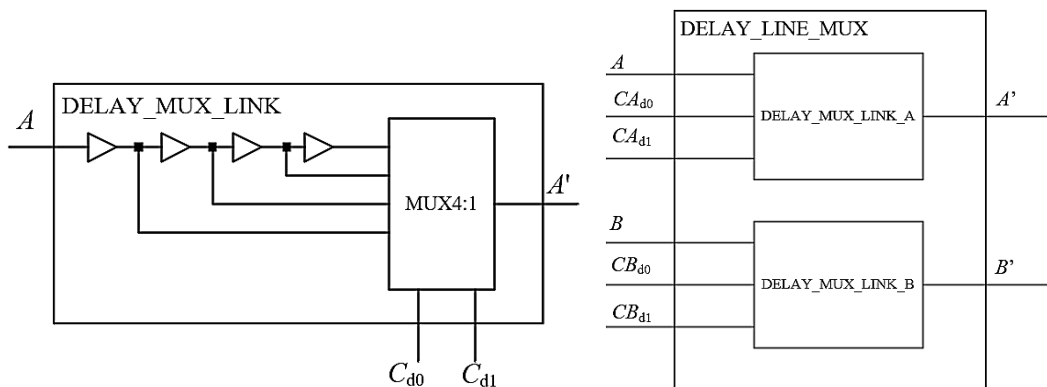


Рис. 11. Структура звена управляемой линии задержки

Fig 11. Propagation delay line link structure

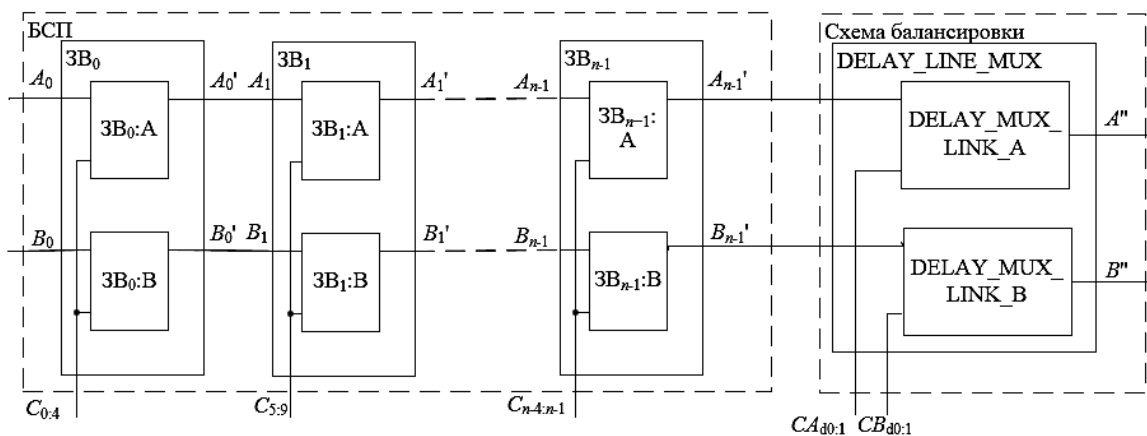


Рис. 12. Схема пары путей АФНФ с управляемой линией задержки

Fig. 12. Scheme of the pair of APUF paths with a propagation delay line

После построения схемы были подобраны такие значения управляющих сигналов $CA_{d0:1}$ и $CB_{d0:1}$ для двух линий задержек, при которых наблюдается минимальная разница между средними значениями задержек для двух построенных путей. Затем полученные значения управляющих сигналов были зафиксированы и с ними был повторен предыдущий эксперимент. Полученные результаты представлены на рис. 13.

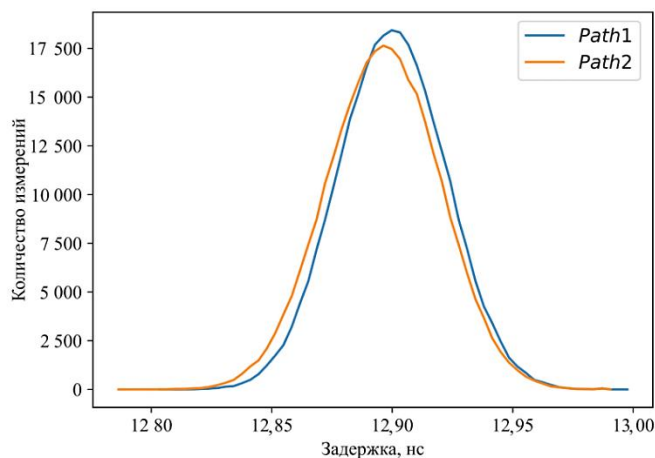


Рис. 13. Временное распределение задержек для двух симметричных путей при $N = 64$ с управляемой линией задержки

Fig. 13. Time distribution of delays for two symmetrical paths with $N = 64$ with controlled delay line

Вместе с тем, как показали исследования, крайне проблематично подобрать такие значения управляющих сигналов на линиях задержек выбранной конфигурации, чтобы измерения для двух путей лежали в одном временном интервале. Более того, смещения по временной оси измерений двух путей на каждом кристалле также различны, что еще более усложняет задачу сведения временных измерений в один интервал. Для решения этой задачи в схему балансировки была внедрена еще одна линия задержки DELAY_LINE_LUT, построенная на звеньях конфигурации, которая аналогична используемой в БСП для более тонкой подстройки фиксированных задержек. Из данных табл. 1 видно, что разброс значений задержек (разница между минимальным и максимальным значениями) для каждого звена составляет ~ 10 пс, также видна зависимость данной задержки от значения запроса. При этом межсоединения между SLICE остаются фиксированными и не зависят от запроса, как в линиях задержки DELAY_LINE_MUX, в которых различия значений задержек имеют больше, чем внутри LUT. В общем виде предлагаемая конфигурация АФНФ показана на рис. 14.

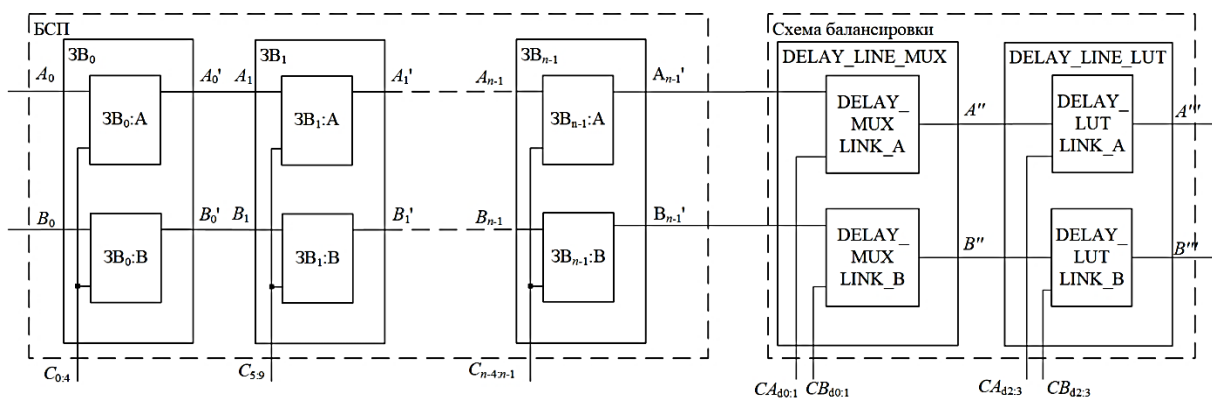


Рис. 14. Схема пары путей АФНФ с управляемой линией задержки

Fig. 14. Scheme of the pair of APUF paths with a propagation delay line

Стратегия подбора фиксированных значений управляющих сигналов для линий задержек состоит из нескольких шагов. Изначально подбирается такая пара управляющих сигналов для линий задержек, построенная на последовательно соединенных повторителях и мультиплексорах (DELAY_LINE_MUX), для которой разница между средними значениями для измерений двух путей минимальна. Затем данная пара управляющих значений фиксируется и выполняется подбор значений управляющих сигналов для линий задержек, построенных на базе LUT-компонентов, которые сконфигурированы аналогично звену БСП (DELAY_LINE_LUT). Критерием для определения лучшей комбинации управляющих сигналов служит разница между средними значениями. Процедура подбора значений для управления линиями фиксированных задержек выполняется индивидуально для каждого кристалла. После этого наиболее подходящие значения линий задержек фиксируются для проведения дальнейших измерений.

Распределения временных задержек для БСП размерности запроса $N = 64$ и график разниц измерений задержек для двух путей одного запроса показаны на рис. 15 и 16 соответственно.

Из полученных результатов следует, что в основном разницы задержек $\Delta(\text{Delay}_1, \text{Delay}_2)$ лежат в интервале, совпадающем с интервалом метастабильности для D-триггера, который чаще всего используется в качестве схемы арбитра АФНФ. Так, согласно документации (URL: https://china.xilinx.com/content/dam/xilinx/support/documents/data_sheets/ds181_Artix_7_Data_Sheet.pdf) для кристалла xc7a100tcs324-1 время предустановки t_{setup} и время удержания t_{hold} составляют 0,07 и 0,12 нс соответственно для триггеров, расположенных в SLICEL. Таким образом, интервал метастабильности составляет $\Delta \in [-0,07; 0,12]$ нс. Данный факт требует дополнительной проработки схемы арбитра АФНФ для повышения стабильности ответов.

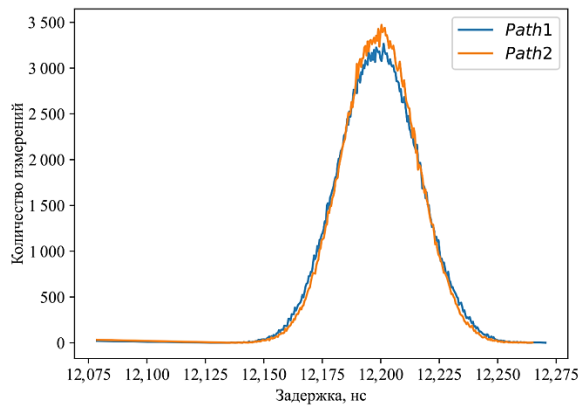


Рис. 15. Временное распределение задержек для конфигурации $N = 64$

Fig. 15. Time distribution of delays for two symmetrical paths with $N = 64$

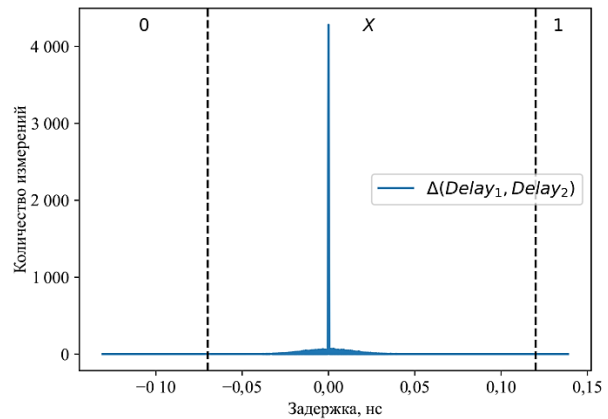


Рис. 16. Временное распределение разниц измерений задержек двух путей конфигурации $N = 64$

Fig. 16. Time distribution of delay measurement differences of two configuration paths $N = 64$

Также были рассчитаны важнейшие характеристики ФНФ стабильности S и межкристальной уникальности U_{cmp} [8] (табл. 2). Характеристика ФНФ стабильности S отражает свойство ФНФ сохранять идентичность ответа на фиксированный запрос при многократной его подаче. Под характеристикой межкристальной уникальности ФНФ U_{cmp} принято понимать долю таких запросов, при которых ответы для всех реализаций ФНФ на различных кристаллах будут уникальны. Аппаратурные затраты приведены в табл. 3.

Таблица 2
 Характеристики АФНФ
 Table 2
 Characteristics of APUF

N	S	U_{cmp}	Доля ответов Response rate		
			X	0	1
Предлагаемая схема АФНФ Proposed scheme APUF					
32	0,88	0,5	0,02	0,46	0,52
64	0,94	0,48	0,01	0,48	0,51
128	0,96	0,53	0,01	0,49	0,5
Классическая АФНФ Classical APUF					
32	0,97	0,02	0,02	0,45	0,53
64	0,98	0,01	0,02	0,48	0,5
128	0,98	0,01	0,03	0,47	0,5

Таблица 3
 Аппаратурные затраты, количество LUT
 Table 3
 Hardware utilization, LUTs

N	БСП BSP	Линии задержки Delay line	Общие затраты Full utilization	Классическая АФНФ Classical APUF
32	12	24	36	64
64	24	24	48	128
128	48	28	76	256

Из представленных данных видно, что, несмотря на определенные затраты на реализацию линий задержек, совокупные аппаратурные затраты значительно ниже, чем затраты на реализацию классической АФНФ с такой же разрядностью запросов. Кроме того, потенциально АФНФ, построенная по предложенной схеме, будет обладать значительно более высокими характеристиками уникальности и случайности по сравнению с классической схемой [6].

С увеличением числа звеньев БСП наблюдается улучшение стабильности, а также снижение количества измерений с равным значением задержки для одного запроса.

Заключение. В статье представлена новая архитектура звеньев БСП, базирующаяся на полном использовании ресурсов LUT-компонентов, которое подразумевает применение LUT-компонентов как функциональных повторителей. Описан подход снижения асимметрии двух путей БСП, основанный на ручном расположении LUT-компонентов. Для устранения асимметрии межсоединений SLICE-блоков предложен способ использования управляемых линий задержек. Однако значения различий временных задержек двух путей для фиксированного запроса демонстрируют невозможность применения в качестве арбитра для предложенной конфигурации симметричных путей D-триггера, который присутствует в классической схеме АФНФ, из-за нарушений условий предустановки и удержания входных сигналов.

В дальнейшем авторы планируют продолжить исследования по разработке конфигурации и архитектуры схемы арбитра, которая могла бы использоваться совместно с рассматриваемой в данной статье архитектурой БСП, а также исследования по изучению характеристик различных конфигураций управляемых линий задержек.

Вклад авторов. А. Ю. Шамына предложил идею построения и балансировки путей ФНФ с управляемой задержкой сигналов и провел экспериментальные исследования. А. А. Иванюк принял участие в обобщении и анализе полученных результатов.

Список использованных источников

1. Pappu, R. *Physical One-Way Functions: PhD Thesis in Media Arts and Sciences* / R. Pappu. – Cambridge : Massachusetts Institute of Technology, 2001. – 154 p.
2. Ярмолик, В. Н. Физически неклонированные функции / В. Н. Ярмолик, Ю. Г. Вашинго // Информатика. – 2011. – № 2(30). – С. 92–103.
3. Иванюк, А. А. Физическая криптография и защита цифровых устройств / А. А. Иванюк, С. С. Заливако // Доклады БГУИР. – 2019. – № 2(120). – С. 50–58.
4. Yang, J. A low resource consumption Arbiter PUF improved switch component design for FPGA / J. Yang, X. Yu, R. Wei // *J. of Physics: Conference Series*. – 2022. – Vol. 2221. – P. 012011.
5. Ярмолик, В. Н. Физически неклонированные функции типа арбитр с заведомо асимметричными параметрами путей / В. Н. Ярмолик, А. А. Иванюк // Доклады БГУИР. – 2022. – № 20(4). – С. 71–79.
6. Иванюк, А. А. Синтез симметричных путей физически неклонированной функции типа арбитр на FPGA / А. А. Иванюк // Информатика. – 2019. – Т. 16, № 2. – С. 99–108.
7. Secure lightweight obfuscated delay-based physical unclonable function design on FPGA / M. H. Ishak [et al.] // *Bulletin of Electrical Engineering and Informatics*. – 2022. – Vol. 11, no. 2. – P. 1075–1083. <https://doi.org/10.11591/eei.v11i2.3265>
8. Шамына, А. Ю. Исследование временных параметров физически неклонированной функции типа арбитр с использованием кольцевого осциллятора / А. Ю. Шамына, А. А. Иванюк // Цифровая трансформация. – 2022. – № 1(28). – С. 27–38.

References

1. Pappu, R. *Physical One-Way Functions: PhD Thesis in Media Arts and Sciences*. Cambridge, Massachusetts Institute of Technology, 2001, 154 p.
2. Yarmolik V. N., Vashinko Y. G. *Physical unclonable functions*. Informatika [Informatics], 2011, no. 2(30), pp. 92–103 (In Russ.).
3. Ivaniuk A. A., Zalivaka S. S. *Physical cryptography and security of digital devices*. Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki [Reports of the Belarusian State University of Informatics and Radioelectronics], 2019, no. 2(120), pp. 50–58 (In Russ.).

4. Yang J. , Yu X., Wei R. A low resource consumption Arbiter PUF improved switch component design for FPGA. *Journal of Physics: Conference Series*, 2022, vol. 2221, p. 012011.
5. Yarmolik V. N., Ivaniuk A. A. *Arbiter physical unclonable functions with asymmetric pairs of paths*. Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki [*Reports of the Belarusian State University of Informatics and Radioelectronics*], 2022, no. 20(4), pp. 71–79 (In Russ.).
6. Ivaniuk A. A. *Synthesis of symmetric paths of arbiter physically unclonable function on FPGA*. *Informatika [Informatics]*, 2019, vol. 16, no. 2, pp. 99–108 (In Russ.).
7. Ishak M. H., Mispan M. S., Chiew W. Ya, Kamaruddin M. R., Korobkov M. A. Secure lightweight obfuscated delay-based physical unclonable function design on FPGA. *Bulletin of Electrical Engineering and Informatics*, 2022, vol. 11, no. 2, pp. 1075–1083. <https://doi.org/10.11591/eei.v11i2.3265>
8. Shamyna A. Yu., Ivaniuk A. A. *Investigation of the timing parameters of the arbiter-based physically unclonable function using a ring oscillator*. *Cifrovaya transformaciya [Digital Transformation]*, 2022, no. 1(28), pp. 27–38 (In Russ.).

Информация об авторах

Шамына Артем Юрьевич, магистр технических наук, старший преподаватель, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: shamyna@bsuir.by

Иваниук Александр Александрович, доктор технических наук, доцент, профессор кафедры информатики, заведующий совместной учебной лабораторией «СК хайникс мемори солюшнс Восточная Европа», Белорусский государственный университет информатики и радиоэлектроники.
E-mail: ivaniuk@bsuir.by

Information about the authors

Artsiom Yu. Shamyna, M. Sc. (Eng.), Senior Lecturer, Belarusian State University of Informatics and Radioelectronics.
E-mail: shamyna@bsuir.by

Alexander A. Ivaniuk, D. Sc. (Eng.), Assoc. Prof., Prof. of Computer Science Department, Head of the Joint Educational Laboratory "SK Hynix Memory Solutions Eastern Europe", Belarusian State University of Informatics and Radioelectronics.
E-mail: ivaniuk@bsuir.by



УДК 004.422
<https://doi.org/10.37661/1816-0301-2022-19-4-42-52>

Оригинальная статья
Original Paper

Методика и программное средство для проведения аудита систем менеджмента информационной безопасности

В. А. Бойправ^{1✉}, Л. Л. Утин²

¹Национальный центр современных искусств Республики Беларусь,
ул. Некрасова, 3, Минск, 220040, Беларусь
✉E-mail: name_abs@rambler.ru

²Белорусский государственный университет
информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь

Аннотация

Цели. Исследование проводилось с целью классификации показателей безопасности информационных систем (ИС) и создания на основе полученных результатов методики усовершенствования ранее разработанного программного средства для проведения аудита систем менеджмента информационной безопасности в организациях Республики Беларусь.

Методы. В ходе разработки и усовершенствования программного средства с помощью метода системно-информационного анализа были определены подходы к его реализации с использованием следующих возможностей: организации анкетирования руководителей подразделений и служб, специалисты которых работают с ИС, предназначенными для обработки не отнесенной к государственным секретам информации, распространение и (или) предоставление которой ограничено; оценки уровня соответствия системы защиты информации ИС организации требованиям, установленным законодательством Республики Беларусь и другими национальными нормативными правовыми актами; систематизации рекомендаций по повышению уровня соответствия системы защиты информации ИС организации установленным требованиям.

Результаты. По результатам апробации разработанного и усовершенствованного программного средства установлено, что его использование позволяет на 20–30 % сократить финансирование затрат на реализацию процесса проведения аудита систем менеджмента информационной безопасности организации.

Заключение. Разработанное и усовершенствованное программное средство по сравнению с аналогами характеризуется пониженной стоимостью ввиду следующих его свойств: простоты запуска и настройки; независимости от типа операционной системы; возможности организации как локального, так и удаленного доступа к нему. Разработанное и усовершенствованное программное средство было апробировано в филиале «Междугородная связь» РУП «Белтелеком».

Ключевые слова: аудит системы менеджмента, информационная безопасность, информационная система, показатели безопасности, программное средство

Для цитирования. Бойправ, В. А. Методика и программное средство для проведения аудита систем менеджмента информационной безопасности / В. А. Бойправ, Л. Л. Утин // Информатика. – 2022. – Т. 19, № 4. – С. 42–52. <https://doi.org/10.37661/1816-0301-2022-19-4-42-52>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 26.10.2022
Подписана в печать | Accepted 18.11.2022
Опубликована | Published 29.12.2022

Methodology and software development for auditing information security management systems

Vladimir A. Boiprav^{1✉}, Leonid L. Utin²

¹*National Center for Contemporary Arts of the Republic of Belarus,
st. Nekrasova, 3, Minsk, 220040, Belarus*

✉E-mail: name_abs@rambler.ru

²*Belarusian State University of Informatics and Radioelectronics,
st. P. Brovki, 6, Minsk, 220013, Belarus*

Abstract

Objectives. Classification of information systems (IS) security indicators and the creation of the method of improved software tool based on its results (in comparison with similar software tool developed earlier by the authors) for auditing information security management systems of organizations in the Republic of Belarus.

Methods. During the development and improvement of the software tool using the method of system-information analysis and the approaches to its implementation were identified based on following capabilities: organization of questionnaires of heads of departments and services whose specialists work with IS designed to information processing not classified as state secrets or IS with limited dissemination; assessment of the level of compliance of the organization's IS information protection system with the requirements established by the legislation of the Republic of Belarus and other national regulatory legal acts; systematization of recommendations for improving the level of compliance of the organization's IS protection system with the established requirements.

Results. Based on the results of the developed improved software tool approbation, it was found that the use of this tool makes it possible to reduce by 20–30 % the financing of costs for the implementation of the auditing the information security management systems of an organization.

Conclusion. The developed improved software tool, compared to analogues, is characterized by reduced cost due to the following properties: ease of launch and configuration; independence from the type of operating system; the possibility of organizing both local and remote access. The developed improved software tool was tested in the branch "Long-Distance Communication Branch" of RUE "Beletelecom".

Keywords: management system audit, information security, information system, security indicator, software tool

For citation. Boiprav V. A., Utin L. L. *Methodology and software development for auditing information security management systems*. Informatika [Informatics], 2022, vol. 19, no. 4, pp. 42–52 (In Russ.). <https://doi.org/10.37661/1816-0301-2022-19-4-42-52>

Conflict of interest. The authors declare of no conflict of interest.

Введение. Обеспечение информационной безопасности в организациях должно осуществляться согласно Концепции информационной безопасности Республики Беларусь, утвержденной постановлением Совета Безопасности Республики Беларусь от 18 марта 2019 г. № 1. Одной из ключевых задач в достижении этой цели является проведение мониторинга, анализа и оценки состояния информационной безопасности, т. е. регулярный аудит систем менеджмента информационной безопасности (СМИБ).

Для сокращения временных и материальных затрат на проведение аудита СМИБ, а также для автоматизации этого процесса авторами настоящей статьи в 2018 г. было разработано программное средство [1], которое позволяет оценить уровень соответствия системы защиты информации (СЗИ) ИС организации требованиям, изложенным в приказе Оперативно-аналитического центра при Президенте Республики Беларусь от 30 августа 2013 г. № 62 «О некоторых вопросах технической и криптографической защиты информации», и систематизировать рекомендации по улучшению СЗИ. Данный приказ утратил значение в силу приказа Оперативно-аналитического центра при Президенте Республики Беларусь от 20 февраля 2020 г. № 66 «О мерах по реализации Указа Президента Республики Беларусь от 9 декабря 2019 г. № 449» с изменениями и дополнениями, утвержденными приказом Оперативно-аналитического центра при Президенте Республики Беларусь от 12 ноября 2021 г. № 195, что сделало неактуальным ранее разработанное программное средство.

Цель исследования, результаты которого представлены в настоящей статье, состояла в классификации показателей безопасности ИС, утвержденных приказом Оперативно-аналитического центра при Президенте Республики Беларусь от 20 февраля 2020 г. № 66, и создании на основе результатов такой классификации методики разработки усовершенствованного программного средства (по сравнению с программным средством, представленным в работе [1]) для проведения аудита СМИБ организаций Республики Беларусь.

Методика разработки программного средства. Усовершенствованное программное средство должно обеспечивать следующие возможности для аудитора СМИБ:

– проведение анкетирования руководителей подразделений и служб [2, 3], специалисты которых работают с ИС, предназначенными для обработки не отнесенной к государственным секретам информации, распространение и(или) предоставление которой ограничено (далее – руководители подразделений и служб);

– оценку уровня соответствия СЗИ ИС организации требованиям, изложенным в Положении о порядке технической и криптографической защиты информации в ИС, предназначенных для обработки информации, распространение и(или) предоставление которой ограничено, утвержденном приказом Оперативно-аналитического центра при Президенте Республики Беларусь от 20 февраля 2020 г. № 66 с изменениями и дополнениями, утвержденными приказом Оперативно-аналитического центра при Президенте Республики Беларусь от 12 ноября 2021 г. № 195 (далее – Положение);

– систематизацию рекомендаций по повышению уровня соответствия СЗИ ИС организации требованиям, изложенным в Положении.

Реализация возможности проведения анкетирования руководителей подразделений и служб. Контрольные листы, предназначенные для анкетирования руководителей подразделений и служб, должны включать в себя вопросы, сформулированные на основе представленного в Положении перечня требований к СЗИ, подлежащих включению в частное техническое задание или задание по безопасности на ИС. Эти требования, по сути, определяют показатели безопасности информации, обрабатываемой с помощью таких ИС.

Так как анкетирование руководителей подразделений и служб направлено на установление факта выполнения (невыполнения) изложенных в Положении требований, то количество вопросов, ответы на которые необходимо получить от указанной категории лиц, составляет более 50 (по количеству требований). В целях оптимизации и структурирования данного процесса в ходе разработки усовершенствованного программного средства была выполнена следующая последовательность действий:

1. Показатели безопасности условно разделены на два класса: общие и частные.

При этом были использованы следующие наименования общих показателей:

- «реализация организационных мер по защите информации» (показатель 1);
- «использование средств технической и криптографической защиты информации» (показатель 2);
- «обеспечение защиты информации в виртуальной инфраструктуре» (показатель 3);
- «обеспечение защиты информации, передаваемой по каналам связи» (показатель 4);
- «обеспечение защиты системы защиты информации» (показатель 5).

Каждый из общих показателей безопасности информации, обрабатываемой в ИС, соответствует определенному виду мероприятий по обеспечению безопасности информации, реализуемых в рамках такой системы.

Информация, обрабатываемая в ИС аудируемой организации, может характеризоваться как всеми, так и некоторыми из указанных общих показателей безопасности в зависимости от класса такой системы. В частности, в ИС классов 4-ин, 4-спец и 4-бг обрабатываемая информация характеризуется показателями 1, 2 и 5; классов 4-юл, 4-дсп – показателями 1, 2, 3, 5; классов 3-ин, 3-спец, 3-бг, 3-юл, 3-дсп – показателями 1–5.

2. Каждому из общих показателей безопасности поставлен в соответствие уникальный набор частных показателей безопасности. С помощью последних может быть установлена степень полноты выполнения требований, определяющих общие показатели безопасности информации, обрабатываемой в ИС аудируемой организации.

3. Разработаны восемь контрольных листов для анкетирования руководителей подразделений и служб.

Контрольный лист 1 включает в себя один вопрос и перечень вариантов ответов на него. По результатам ответов можно установить класс ИС, с которым работают организации электро-связи.

Контрольный лист 2 включает в себя пункты для выбора, по результатам которого можно выполнить следующие действия:

- установить, какие из мероприятий, соответствующих общим показателям безопасности информации, реализуются в рамках ИС аудируемой организации;
- сгенерировать на основе контрольных листов 3–7 пункты для выбора в виде перечня частных показателей безопасности информации, соответствующих выбранным общим показателям безопасности информации.

Содержание контрольного листа 2 зависит от ответа на вопрос из контрольного листа 1. Предложены три вида контрольного листа 2 (табл. 1). На рис. 1 представлен алгоритм проведения анкетирования с помощью контрольного листа 1, который отражает взаимосвязь между ответом на вопрос из контрольного листа 1 и выбором вида контрольного листа 2.

Разработанные авторами контрольные листы 3–8 включают в себя пункты для выбора, по результатам которого можно установить уровень соответствия ИС аудируемой организации требованиям, определяющим частные показатели безопасности информации, обрабатываемой с помощью этой системы. Содержание рассматриваемых контрольных листов представляет собой совокупность модулей, номенклатура которых зависит от выбранного ответа на вопрос из контрольного листа 1, т. е. от класса ИС, а также от номенклатуры выбранных пунктов в контрольном листе 2.

Обобщенный алгоритм проведения анкетирования, отражающий взаимосвязь между выбором пунктов из контрольного листа 2 и выбором модулей для формирования контрольных листов 3–8, показан на рис. 2, где используются следующие обозначения: N – номер вида контрольного листа ($N \in \{1, 2, 3\}$); M – номер пункта для выбора в контрольном листе ($M \in \{1, 2, 3\} | N=1; M \in \{1, 2, 3, 4\} | N=2; M \in \{1, 2, 3, 4, 5\} | N=3$); L – номер формируемого контрольного листа ($L \in \{3, 4, 5, 6, 7, 8\}$), который зависит от выбранного варианта ответа на вопрос из контрольного листа 1.

Таблица 1
Содержание трех видов контрольного листа 2

Table 1
Contents of the three kinds of checklist 2

Вид контрольного листа <i>Checklist kind</i>	Содержание <i>Content</i>
Вид 1	Выберите реализуемые на Вашем предприятии разновидности мероприятий по защите информации, обрабатываемой в информационной системе. Реализация организационных мер по защите информации (вариант 1.1). Использование средств технической и криптографической защиты информации (вариант 1.2). Обеспечение защиты системы защиты информации (вариант 1.3)
Вид 2	Выберите реализуемые на Вашем предприятии разновидности мероприятий по защите информации, обрабатываемой в информационной системе. Реализация организационных мер по защите информации (вариант 2.1). Использование средств технической и криптографической защиты информации (вариант 2.2). Обеспечение защиты информации в виртуальной инфраструктуре (вариант 2.3). Обеспечение защиты системы защиты информации (вариант 2.4)
Вид 3	Выберите реализуемые на Вашем предприятии разновидности мероприятий по защите информации, обрабатываемой в информационной системе. Реализация организационных мер по защите информации (вариант 3.1). Использование средств технической и криптографической защиты информации (вариант 3.2). Обеспечение защиты информации в виртуальной инфраструктуре (вариант 3.3). Обеспечение защиты информации, передаваемой по каналам связи (вариант 3.4). Обеспечение защиты системы защиты информации (вариант 3.5)

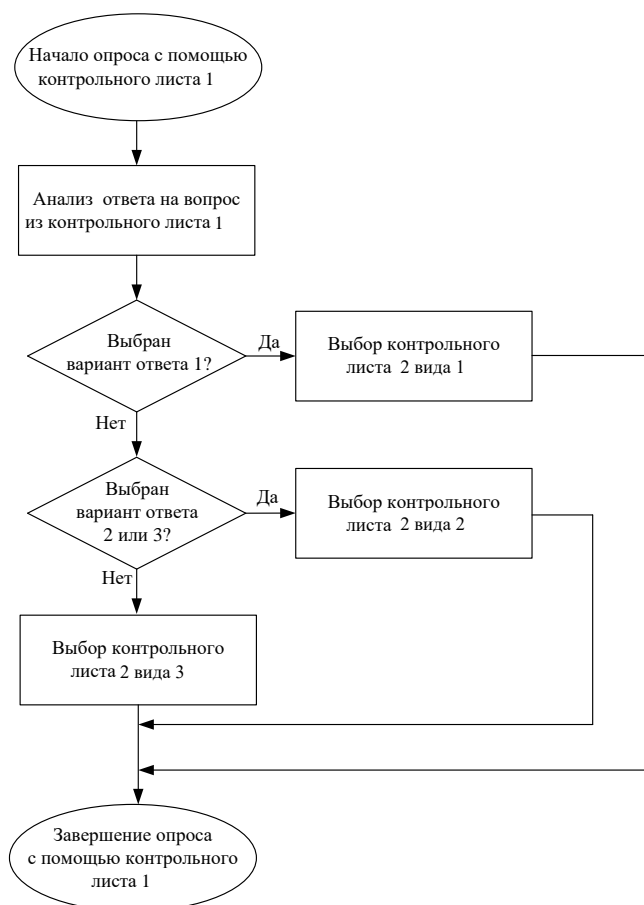


Рис. 1. Схема проведения анкетирования согласно контрольному листу 1
Fig. 1. Scheme of conducting a survey according to checklist 1

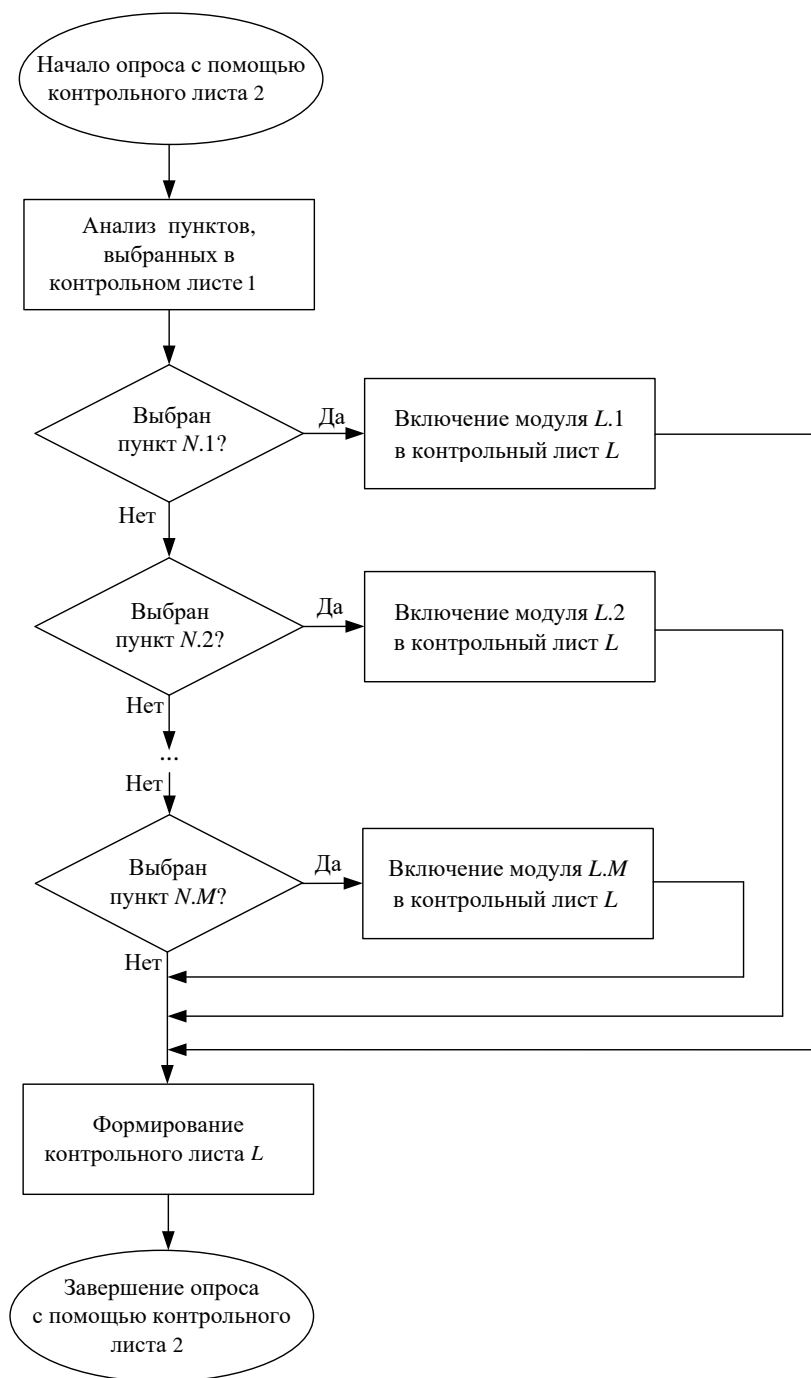


Рис. 2. Обобщенный алгоритм проведения анкетирования с помощью контрольного листа 2
 Fig. 2. Generalized algorithm for the questionnaire conducting using checklist 2

Таким образом, если в качестве ответа на вопрос из контрольного листа 1 выбран вариант 1, то ему соответствует контрольный лист 3. Если в качестве ответа на вопрос из контрольного листа 1 выбран один из вариантов 2–6, то номер формируемого контрольного листа будет соответственно 4, 5, 6, 7 или 8.

Реализация возможности оценки уровня соответствия СЗИ ИС организации требованиям, изложенным в Положении. В ходе разработки усовершенствованного программного средства авторами предложено ранжировать по пяти уровням соответствие СЗИ ИС требованиям:

- полное несоответствие;
- низкий уровень соответствия;
- средний уровень соответствия;
- высокий уровень соответствия;
- полное соответствие.

Уровни соответствия СЗИ ИС вышеизложенным требованиям устанавливаются по каждому из общих показателей безопасности системы. Например, СЗИ ИС может характеризоваться средним уровнем соответствия требованиям Положения в части реализации организационных мер по защите информации, но при этом полностью не соответствовать требованиям в части реализации технической и криптографической защиты информации, защиты информации в виртуальной инфраструктуре, защиты информации, передаваемой по каналам связи, защиты СЗИ.

Уровень соответствия СЗИ ИС общему показателю безопасности зависит от количества выбранных анкетируемым сотрудником частных показателей безопасности, на которых основаны контрольные листы 3–8.

Взаимосвязь между уровнем соответствия общего показателя безопасности СЗИ ИС и количеством выбранных анкетируемым сотрудником частных показателей безопасности представлена в табл. 2, где MAX – количество частных показателей безопасности, на основе которых составлен контрольный лист для анкетирования и которые соответствуют определенному общему показателю безопасности; MV – медианное значение среди множества значений, отражающих количество частных показателей безопасности, на основе которых составлены все контрольные листы для анкетирования и каждый из которых соответствует определенному общему показателю безопасности.

Таблица 2

Уровни соответствия общих показателей безопасности СЗИ ИС количеству выбранных анкетируемым сотрудником частных показателей безопасности (PM)

Table 2

Compliance degrees of the general security indicators of the information security system of the IS to the number of private security indicators (PM) selected by the surveyed employee

Уровень соответствия <i>Compliance degree</i>	Значение PM <i>PM value</i>
Полное несоответствие	0
Низкий	$\begin{cases} 0 < PM < \frac{MAX - 1}{2}, & \text{если } MAX - \text{нечетное число и } MAX \leq MV; \\ 0 < PM < \frac{MAX}{2}, & \text{если } MAX - \text{четное число}; \\ 0 < PM < \frac{MAX + 1}{2}, & \text{если } MAX - \text{нечетное число и } MAX > MV \end{cases}$
Средний	$\begin{cases} \frac{MAX - 1}{2}, & \text{если } MAX - \text{нечетное число и } MAX \leq MV; \\ \frac{MAX}{2}, & \text{если } MAX - \text{четное число}; \\ \frac{MAX + 1}{2}, & \text{если } MAX - \text{нечетное число и } MAX > MV \end{cases}$
Высокий	$\begin{cases} MAX - 1, & \text{если } MAX \leq MV; \\ MAX - 2, & \text{если } MAX > MV \end{cases}$
Полное соответствие	MAX

Значение MV определяется на основе совокупности следующих условий:

$$\left\{ \begin{array}{l} MV \in \{GV_1, GV_2, \dots, GV_m, \dots, GV_r\}; \\ MV = GV_m; \\ m = \frac{r}{2}, \text{ если } r - \text{ четное число}; \\ m = \frac{r+1}{2}, \text{ если } r - \text{ нечетное число}, \end{array} \right.$$

где $GV_1, GV_2, \dots, GV_m, \dots, GV_r$ – проранжированные по возрастанию значения количества частных показателей безопасности, каждый из которых соответствует определенному общему показателю безопасности; GV_m – значение, которое является медианным в множестве $GV_1, GV_2, \dots, GV_m, \dots, GV_r$; m – порядковый номер значения, которое является медианным в множестве $GV_1, GV_2, \dots, GV_m, \dots, GV_r$; r – суммарное количество общих показателей безопасности в множестве $GV_1, GV_2, \dots, GV_m, \dots, GV_r$.

Для разработки программного средства использован язык программирования JavaScript.

Результаты исследования. Разработанное и усовершенствованное программное средство представляет собой совокупность файлов формата html, в связи с чем (в отличие от аналогов [4–8]) оно не требует установки и может быть запущено и использовано на всех рабочих станциях, оснащенных интернет-браузером.

На рис. 3–7 показаны фрагменты диалоговых окон файлов программного средства.

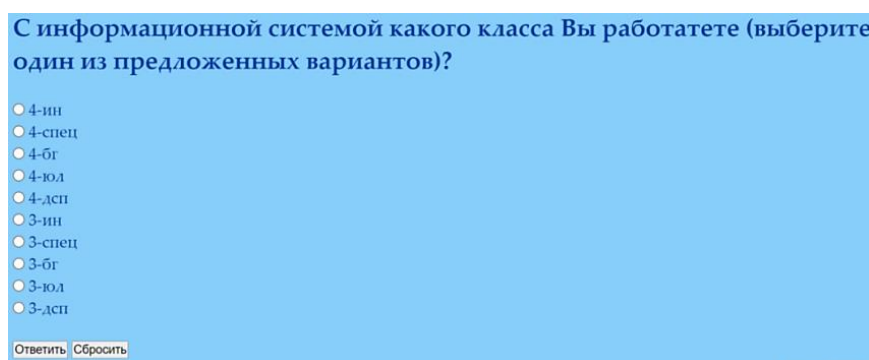


Рис. 3. Фрагмент диалогового окна программного средства для выбора класса ИС

Fig. 3. Fragment of the dialog box of the software tool for selecting the IP class

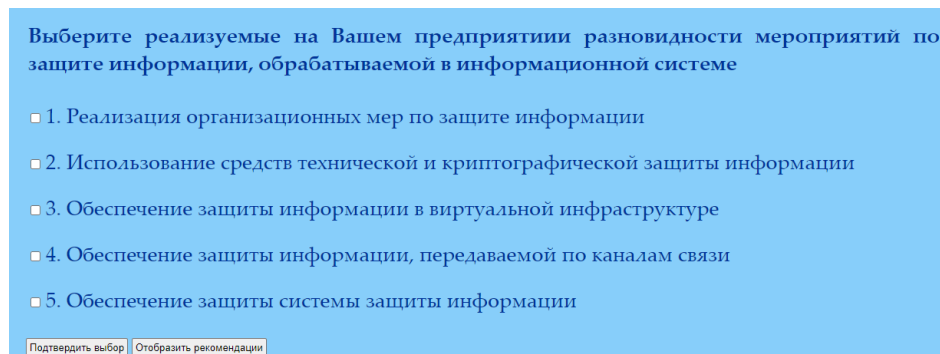


Рис. 4. Фрагмент диалогового окна программного средства для выбора реализуемых на предприятии разновидностей мероприятий по защите информации

Fig. 4. Fragment of the dialog box of the selection tool types of information security measures implemented at the enterprise

2.1 Выберите реализуемые на Вашем предприятии мероприятия разновидности 2

- Обеспечение идентификации и аутентификации пользователей информационной системы
- Обеспечение защиты обратной связи при вводе аутентификационной информации
- Обеспечение полномочного управления (создание, активация, блокировка и уничтожение) учетными записями пользователей информационной системы
- Обеспечение блокировки доступа к объектам информационной системы после истечения установленного времени бездействия (неактивности) пользователя информационной системы или по его запросу
- Обеспечение резервирования информации, подлежащей резервированию
- Обеспечение защиты средств вычислительной техники от вредоносных программ
- Обеспечение подлинности и контроля целостности электронных документов в информационной системе (средства выработки электронной цифровой подписи, средства проверки электронной цифровой подписи, средства выработки личного ключа или открытого ключа электронной цифровой подписи)

Подтвердить выбор Отобразить рекомендации

Рис. 5. Фрагмент диалогового окна программного средства после выбора реализуемых на предприятии разновидностей мероприятий по защите информации и нажатия на кнопку «Подтвердить выбор»

Fig. 5. A fragment of the dialog box of the software tool after selecting the types of information protection measures implemented at the enterprise and clicking on the "Confirm selection" button

Низкая степень соответствия требованиям приказа ОАЦ № 66 в части реализации организационных мер по защите информации
Полное несоответствие требованиям приказа ОАЦ № 66 в части реализации технической и криптографической защиты информации
Полное несоответствие требованиям приказа ОАЦ № 66 в части реализации защиты информации в виртуальной инфраструктуре
Полное несоответствие требованиям приказа ОАЦ № 66 в части реализации защиты информации, передаваемой по каналам связи
Полное несоответствие требованиям приказа ОАЦ № 66 в части реализации защиты системы защиты информации

Рис. 6. Текст диалогового окна программного средства после нажатия на кнопку «Подтвердить выбор» при условии сделанного выбора пунктов

Fig. 6. The text of the dialog box of the software tool after clicking on the "Confirm selection" button, provided that the items were selected

Для поддержания соответствия в части реализации организационных мер по защите информации необходимо: обновление программного обеспечения объектов информационной системы и контроль за своевременность такого обновления
Для поддержания соответствия в части реализации технической и криптографической защиты информации необходимо: обеспечение подлинности и контроля целостности электронных документов в информационной системе (средства выработки электронной цифровой подписи, средства проверки электронной цифровой подписи, средства выработки личного ключа или открытого ключа электронной цифровой подписи)
Для обеспечения соответствия в части реализации защиты системы защиты информации необходимо: обеспечение изменения атрибутов безопасности сетевого оборудования, системного программного обеспечения и средств защиты информации, установленных по умолчанию

Рис. 7. Текст диалогового окна программного средства после нажатия на кнопку «Отобразить рекомендации»

Fig. 7. The text of the dialog box of the tool after clicking on the button "Display recommendations"

В разработанном и усовершенствованном программном средстве предусмотрена возможность редактирования вопросов, что создает условия для его использования при проведении аудита СЗИ ИС любых организаций независимо от вида их деятельности, формы собственности и ведомственной подчиненности.

Заключение. Разработанное и усовершенствованное программное средство характеризуется пониженной (по сравнению с аналогами) стоимостью ввиду простоты запуска и настройки, независимости от типа операционной системы, а также возможности организации как локального, так и удаленного доступа к нему. Указанные свойства избавляют аудитора или сотрудников

аудируемой организации от необходимости выполнения ручной установки, а руководителя этой организации – от необходимости предоставления отдельного помещения для проведения аудита и закупки дополнительного оборудования для этих целей. В разработанном программном средстве имеется возможность редактирования вопросов для проведения аудита. Применение разработанного программного средства позволит существенно сократить время и материальные затраты на проведение аудита.

Представленное в настоящей статье программное средство было апробировано в филиале «Междугородная связь» РУП «Белтелеком». Установлено, что использование этого средства позволяет сократить на 20–30 % финансирование затрат на реализацию процесса проведения аудита СМИБ организации.

Программное средство зарегистрировано в установленном порядке на государственном предприятии «Национальный центр интеллектуальной собственности» (свидетельство о регистрации № 1447 от 14.10.2021).

Вклад авторов. В. А. Бойправ создал методику разработки и усовершенствования программного средства, выполнил его практическую реализацию. Л. Л. Утин предложил подходы по оптимизации процесса практической реализации программного средства.

Список использованных источников

1. Бойправ, В. А. Программное средство для проведения аудита системы защиты информации организации / В. А. Бойправ, В. В. Ковалев, Л. Л. Утин // Доклады БГУИР. – 2018. – № 5(115). – С. 44–49.
2. Pandey, S. K. A comparative study of risk assessment methodologies for information systems / S. K. Pandey, K. Mustafa // Bulletin of Electrical Engineering and Informatics. – 2012. – Vol. 1, no. 2. – P. 111–122.
3. Сагитова, В. В. Применение метода экспертных оценок для автоматизации аудита информационных систем персональных данных / В. В. Сагитова, В. И. Васильев // Вестник УГАТУ. – 2017. – Т. 21, № 3(73). – С. 105–112.
4. Якимова, З. В. Динамика уровня вовлеченности персонала в зависимости от стажа работы в организации / З. В. Якимова, А. С. Пушкина // АНИ: экономика и управление. – 2018. – № 1(22). – С. 283–286.
5. Information security risk assessment / I. Kuzminykh [et al.] // Encyclopedia. – 2021. – Vol. 1(3). – P. 602–617.
6. Nest, D. P. van der. The use of generalised audit software by internal audit functions in a developing country: a maturity level assessment / D. P. van der Nest, L. Smidt, D. Lubbe // Risk Governance and Control: Financial Markets & Institutions. – 2017. – Vol. 7(4–2). – P. 189–202.
7. Lehmann C. M. Integrating generalized audit software and teaching fraud detection in information systems auditing courses / C. M. Lehmann // J. of Forensic & Investigative Accounting. – 2012. – Vol. 4, iss. 1. – P. 319–368.
8. SECURQUAL: An instrument for evaluating the effectiveness of enterprise information security programs / P. J. Steinbart [et al.] // J. of Information Systems. – 2015. – Vol. 30(1). – P. 71–92.

References

1. Boiprav V. A., Kovalev V. V., Utin L. L. Software for audit of information protection system of the organization. Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki [Reports of the Belarusian State University of Informatics and Radioelectronics], 2018, № 5(115), pp. 44–49 (In Russ.).
2. Pandey S. K., Mustafa K. A comparative study of risk assessment methodologies for information systems. Bulletin of Electrical Engineering and Informatics, 2012, vol. 1, no. 2, pp. 111–122.
3. Sagitova V. V., Vasil'ev V. I. Application of the method of expert assessments to automate the audit of personal data information systems. Vestnik Ufimskogo gosudarstvennogo aviacionnogo tehniceskogo universiteta [Bulletin of the Ufa State Aviation Technical University], 2017, vol. 21, no. 3(73), pp. 105–112 (In Russ.).

4. Yakimova Z. V., Pushkina A. S. *Dynamics of the level of personnel involvement depending on the length of service in the organization*. *Azimuth nauchnyh issledovanij: jekonomika i upravlenie [Azimuth of Scientific Research: Economics and Management]*, 2018, no. 1(22), pp. 283–286 (In Russ.).

5. Kuzminykh I., Ghita B., Sokolov V., Bakhshi T. Information security risk assessment. *Encyclopedia*, 2021, vol. 1(3), pp. 602–617.

6. Nest D. P. van der, Smidt L., Lubbe D. The use of generalised audit software by internal audit functions in a developing country: a maturity level assessment. *Risk Governance and Control: Financial Markets & Institutions*, 2017, vol. 7(4–2), pp. 189–202.

7. Lehmann C. M. Integrating generalized audit software and teaching fraud detection in information systems auditing courses. *Journal of Forensic & Investigative Accounting*, 2012, vol. 4, iss. 1, pp. 319–368.

8. Steinbart P. J., Gal G., Dilla W. N., Raschke R. L. SECURQUAL: An instrument for evaluating the effectiveness of enterprise information security programs. *Journal of Information Systems*, 2015, vol. 30(1), pp. 71–92.

Информация об авторах

Бойправ Владимир Андреевич, заместитель директора по общим вопросам, Национальный центр современных искусств Республики Беларусь.

E-mail: name_abs@rambler.ru

Утин Леонид Львович, кандидат технических наук, доцент, заместитель начальника военного факультета по учебной и научной работе – первый заместитель начальника, Белорусский государственный университет информатики и радиоэлектроники.

E-mail: utin@bsuir.by

Information about the authors

Vladimir A. Boiprav, Deputy Director for General Affairs, National Center for Contemporary Arts of the Republic of Belarus.

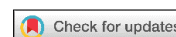
E-mail: name_abs@rambler.ru

Leonid L. Utin, Ph. D. (Eng.), Associate Professor, Deputy Head of the Military Faculty for Educational and Scientific Work – First Deputy Head, Belarusian State University of Informatics and Radioelectronics.

E-mail: utin@bsuir.by

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING
AND PATTERN RECOGNITION



УДК 004.81
<https://doi.org/10.37661/1816-0301-2022-19-4-53-68>

Оригинальная статья
Original Paper

Разработка алгоритма распознавания эмоций человека с использованием сверточной нейронной сети на основе аудиоданных

В. В. Семенюк, М. В. Складчиков[✉]

*Донецкий техникум промышленной автоматике
имени А. В. Захарченко,
ул. Горького, 163, Донецк, 83000, Украина
✉E-mail: maxsklad19981@yandex.ru*

Аннотация

Цели. Приведено описание и рассмотрен опыт создания алгоритма распознавания эмоционального состояния субъекта.

Методы. Используются методы обработки изображений.

Результаты. Предложенный алгоритм позволяет распознавать эмоциональные состояния субъекта на основании звукового набора данных. Благодаря проведенному исследованию удалось улучшить точность работы алгоритма путем изменения подаваемого на вход нейронной сети набора данных.

Описаны этапы обучения сверточной нейронной сети на заранее заготовленном наборе звуковых данных, а также структура алгоритма. Для валидации нейронной сети был отобран иной, не участвующий в тренировке, набор аудиоданных. В результате проведения исследования построены графики, демонстрирующие точность работы предлагаемого метода.

После получения первоначальных данных сделан анализ возможностей улучшения алгоритма с точки зрения эргономики и точности его работы. Разработана стратегия, позволяющая добиться лучшего результата и получить более точный алгоритм. На основании заключений, изложенных в статье, приводится обоснование выбора представления набора данных и программного комплекса, необходимого для реализации программной части алгоритма.

Заключение. Предложенный алгоритм обладает высокой точностью и не требует больших вычислительных затрат.

Ключевые слова: нейронная сеть, распознавание эмоций человека, сверточная нейронная сеть, дактилоскопия звука, программная библиотека TensorFlow, нейросетевая библиотека Keras, пакет программ Matlab

Для цитирования. Семенюк, В. В. Разработка алгоритма распознавания эмоций человека с использованием сверточной нейронной сети на основе аудиоданных / В. В. Семенюк, М. В. Складчиков // Информатика. – 2022. – Т. 19, № 4. – С. 53–68. <https://doi.org/10.37661/1816-0301-2022-19-4-53-68>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 08.08.2022

Подписана в печать | Accepted 08.09.2022

Опубликована | Published 29.12.2022

Algorithm development for recognizing human emotions using a convolutional neural network based on audio data

Viktoriya V. Semenuk, Maxim V. Skladchikov[✉]

Donetsk Technical School of Industrial Automation

after A. V. Zakharchenko,

st. Gorkogo, 163, Donetsk, 83000, Ukraine

[✉]*E-mail: maxsklad19981@yandex.ru*

Abstract

Objectives. This article provides a description and experience of creating the algorithm for recognizing the emotional state of the subject.

Methods. Image processing methods are used.

Results. The proposed algorithm makes it possible to recognize the emotional states of the subject on the basis of an audio data set. It was possible to improve the accuracy of the algorithm by changing the data set supplied to the input of the neural network.

The stages of training convolutional neural network on a pre-prepared set of audio data are described, and the structure of the algorithm is described. To validate the neural network different set of audio data, not participating in the training, was selected. As a result of the study, graphs were constructed demonstrating the accuracy of the proposed method.

After receiving the initial data of the study, the analysis of the possibilities for improving the algorithm in terms of ergonomics and accuracy of operation was also carried out. The strategy was developed to achieve a better result and obtain a more accurate algorithm. Based on the conclusions presented in the article, the rationale for choosing the representation of the data set and the software package necessary for the implementation of the software part of the algorithm is given.

Conclusion. The proposed algorithm has a high accuracy of operation and does not require large computational costs.

Keywords: neural network, human emotion recognition, convolutional neural network, sound fingerprinting, TensorFlow software library, Keras neural network library, Matlab software package

For citation. Semenuk V. V., Skladchikov M. V. *Algorithm development for recognizing human emotions using a convolutional neural network based on audio data*. Informatika [Informatics], 2022, vol. 19, no. 4, pp. 53–68 (In Russ.). <https://doi.org/10.37661/1816-0301-2022-19-4-53-68>

Conflict of interest. The authors declare of no conflict of interest.

Введение. Новизна предлагаемого метода заключается в высокой точности работы описываемого алгоритма по сравнению с имеющимися алгоритмами идентификации эмоций. Для достижения поставленной цели в качестве архитектуры была выбрана сверточная нейронная сеть. Использование разработанной структуры нейронной сети обусловлено высокой точностью и простотой распознавания изображений. Следует отметить, что большинство алгоритмов, идентифицирующих эмоции, формируют результат на основании видеоданных. Такой подход в первую очередь требует высококачественного регистрирующего оборудования. Кроме того, для работы необходимо производить сложные вычисления для регистрации активности лицевых мышц.

Мотивацией к разработке предложенного алгоритма и архитектуры нейронной сети стали исследование системы голосового управления и в целом классификация различных систем [1]. Как утверждают авторы доклада, точность распознавания звуковых данных при создании идеальных условий составила около 92 %.

Сверточные нейронные сети зачастую применяются для анализа изображений. Поэтому записанные данные звука преобразовывались в изображение с использованием технологии «дактилоскопия звука» [2]. Благодаря этому удалось снизить временные затраты, необходимые для обработки входных данных, что в свою очередь повысило эргономические свойства предлагаемого метода.

Каждый человек выражает эмоции при возникновении определенных внешних или внутренних возбудителей. Ввиду индивидуальности каждого субъекта, а также его психологического состояния, которое в разные периоды жизни может меняться, довольно сложно выделить единый способ оценки эмоций [3–7]. Это приводит к появлению большого количества подходов к их идентификации и классификации [8–24].

В современном мире существует большое количество алгоритмов и систем, способных анализировать эмоциональный окрас человека [13, 15]. В век цифровизации и развития искусственного интеллекта в когнитивистике наиболее актуальным подходом к анализу эмоций человека является технология SER (Speech Emotion Recognition). Этот подход основывается на обработке и анализе звуковых сигналов. Голос как набор данных для идентификации отдельных типов эмоций является наиболее информативным, что позволяет качественно классифицировать эмоции человека по сравнению с другими подходами к оценке эмоционального окраса человека.

Проанализировав существующие решения в области распознавания эмоционального состояния человека по голосу, можно выделить ряд проблем [9, 10, 12–15, 17, 18, 20, 21, 24]:

1. Жесткая взаимосвязь точности и количества идентифицируемых эмоций снижает сферу применения алгоритмов.

2. Для увеличения точности работы может использоваться более громоздкий математический аппарат, который не всегда обоснован с точки зрения эргономики работы алгоритма.

3. Для сокращения затрачиваемых ресурсов анализируется малый участок идентифицируемого набора данных, на основании которого делается вывод об адекватности эмоционального состояния. Это в свою очередь увеличивает риск ошибочного заключения на ином участке распознавания.

4. Использование исключительно статистических методов машинного обучения для распознавания эмоций имеет ряд существенных недостатков и не позволяет с необходимой точностью обеспечивать управление информационными потоками, развитие, перестроение и увеличение набора данных. Также подобные системы тяжело синтезировать и обеспечивать их взаимодействие на различных уровнях между собой.

Сказанное выше обуславливает цель исследования, которая заключается в построении системы, удовлетворяющей решению выявленных проблем.

Методики оценки эмоционального состояния. Современные модели оценки эмоционального состояния человека можно разделить на два основных вида: дискретные и многомерные.

Дискретная модель подразумевает упрощенную оценку эмоций человека. При использовании такой модели результат оценки эмоционального состояния субъекта основывается лишь на базовых (первичных) факторах. Это приводит к уменьшению точности работы системы, что сильно сказывается на объективности вывода алгоритма.

Многомерная модель предполагает более глубокий анализ эмоционального состояния субъекта. С ее помощью оцениваются не только базовые параметры, влияющие на состояние человека, но и косвенные (пульс, изменение цвета лица и т. д.). Многомерная модель позволяет приближенно имитировать работу мозга при оценке эмоционального состояния человека. Применение такой модели дает возможность получить более точный результат, однако при этом возрастают сложность и громоздкость разрабатываемого алгоритма.

Вариантом использования описанных моделей являются аффективные вычисления. Данная область знаний базируется на множестве научных дисциплин (информатике, когнитивистике,

психологии и т. д.). Ее основная задача – анализ и интерпретация эмоций человека. В зависимости от сферы применения аффективные вычисления могут использоваться для идентификации эмоций человека (методики распознавания эмоций), а также для создания или симулирования его эмоциональных состояний (робототехника). Для реализации систем первого типа имеются различные человеко-машинные интерфейсы. При идентификации эмоций в «статике» можно использовать изображение. Такие системы просты в построении и имеют достаточную точность, необходимую для получения качественного результата работы алгоритма, однако они не позволяют идентифицировать интенсивность эмоционального состояния субъекта.

Целью настоящего исследования является разработка алгоритма идентификации эмоций человека на основании набора звуковых данных. Для классификации и распознавания эмоций человека применялись аудиозаписи, которые содержали характерные признаки той или иной эмоции. Для реализации процесса распознавания вместо привычной рекуррентной нейронной сети использовалась сверточная нейронная сеть, которая была оптимизирована специально для работы с аудиозаписями и текстом в качестве эксперимента, так как возможно, что за счет использования сверточной нейронной сети можно получить более быстрый результат. Ввиду нестандартного выбора типа нейронной сети, применяемой в исследовании, необходимо было преобразовать входной информационный поток в изображение. Для этого использовалась технология *audio fingerprint*, основная задача которой – преобразование входного аудиопотока данных в спектрограмму.

Проведение сравнительного эксперимента. Для проведения сравнительного эксперимента и оценки зависимости точности работы алгоритма распознавания эмоций человека от количества идентифицируемых эмоций было принято решение разделить обучение нейронной сети на две независимые модели: на три класса эмоций (позитивные, нейтральные и негативные) и с целью более точной классификации – на восемь классов (счастье, агрессия, спокойствие, отвращение, удивление, нейтральное состояние, печаль, страх). Выбор конкретного набора эмоциональных состояний обусловлен возможностями используемой нейросетевой библиотеки Keras.

На рис. 1 изображено дерево разбиения эмоционального состояния на классы.



Рис. 1. Иерархия классов эмоций

Fig. 1. Hierarchy of emotion classes

Для обучения нейронной сети в соответствии с иерархией классов (рис. 1) был отобран набор аудиофайлов. Для каждого эмоционального состояния нейронная сеть выделяла признаки на определенном количестве аудиофайлов, отображающих одну и ту же эмоцию. В табл. 1–4 представлены данные о количестве аудиофайлов для каждого класса.

Таблица 1
 Количество аудиофайлов для каждого класса обучающей выборки

Table 1
Number of audiofiles for each class of training sample

Класс <i>Class</i>	Количество аудиофайлов <i>Number of images</i>
Агрессия	665
Спокойствие	299
Отвращение	519
Страх	665
Счастье	665
Нейтральные	244
Печаль	346
Удивление	200
Всего	3603

Таблица 2
 Количество аудиофайлов для каждого класса тестовой выборки

Table 2
Number of audiofiles for each class of test set

Класс <i>Class</i>	Количество аудиофайлов <i>Number of images</i>
Агрессия	167
Спокойствие	75
Отвращение	130
Страх	167
Счастье	167
Нейтральные	113
Печаль	87
Удивление	50
Всего	956

Таблица 3
 Количество аудиофайлов для обобщенных классов обучающей выборки

Table 3
The number of audiofiles for generalized classes of training sample

Класс <i>Class</i>	Количество аудиофайлов <i>Number of images</i>
Позитивные	865
Нейтральные	543
Негативные	2195
Всего	3603

Таблица 4
 Количество аудиофайлов для обобщенных классов тестовой выборки

Table 4
Number of audiofiles for generalized classes of test sample

Класс <i>Class</i>	Количество аудиофайлов <i>Number of images</i>
Позитивные	217
Нейтральные	188
Негативные	551
Всего	956

В табл. 5 приведены результаты сравнительного анализа систем распознавания эмоций и их характеристики [25].

Для разрабатываемого метода идентификации эмоций был использован специальный алгоритм, принимающий на вход набор аудиофайлов. В результате работы данного алгоритма сформировался соответствующий набор спектрограмм.

Задачей классификации эмоций занимаются уже давно. Данный факт обусловлен низкой точностью работы применяемых алгоритмов, что требует дальнейших исследований в этой области. Большая часть работ, которые были изучены для разработки стратегии исследования, базировалась на классификации эмоций с помощью видеопотока данных. В качестве аттракторов использовались опорные точки лица. Основная задача нейронной сети – построение карты точек лица. В результате удастся сформировать данные, необходимые для тренировки. В связи с необходимостью сложных вычислений для такой задачи используются сверточные нейронные сети. На вход нейронной сети подается набор анализируемых изображений и на основании геометрических параметров лица происходит сегментация отдельных его зон. Далее полученная информация используется для выделения ключевых признаков, на основании которых

и происходит в дальнейшем классификация эмоций. С помощью специализированных данных DataSet, создаваемых в идеальных условиях и применяемых в задачах обучения нейронных сетей, удается разработать оптимальный алгоритм для распознавания [26, 27].

Таблица 5
Программные пакеты для распознавания эмоций на основании видеоданных

Table 5
Software packages for emotion recognition based on video data

Программный пакет <i>Software package</i>	Количество эмоций <i>Number of emotions</i>	Способы поиска решения <i>Ways to find a solution</i>	Методы классификации <i>Classification methods</i>
Compound emotion	7	Распознавание эмоций с помощью фильтра Габора	k -ближайших соседей. Дискриминантный анализ Кернела
EmotioNet	23	Вычисление евклидова расстояния между нормализованными ориентирами. Вычисление угла между ключевыми точками. Использование фильтра Габора	Дискриминантный анализ Кернела
Real-time mobile	7	Активная форма модели. Смещение между ориентирами	Метод опорных векторов
Local region specific feature	7	Функция извлечения локального двоичного шаблона (LBP). Геометрическая нормализация центра	Метод опорных векторов

Ввиду того что работы по распознаванию эмоций зачастую базируются на видеоданных, было решено разработать модель нейронной сети, позволяющую классифицировать эмоциональные состояния на основании речевого набора данных. На вход алгоритма преобразования поступал аудиофайл, характеризующий определенную эмоцию (рис. 2, *a*). Сигнал анализировался, и в соответствии со спектральной плотностью мощности формировалось изображение на выходе (рис. 2, *b*).

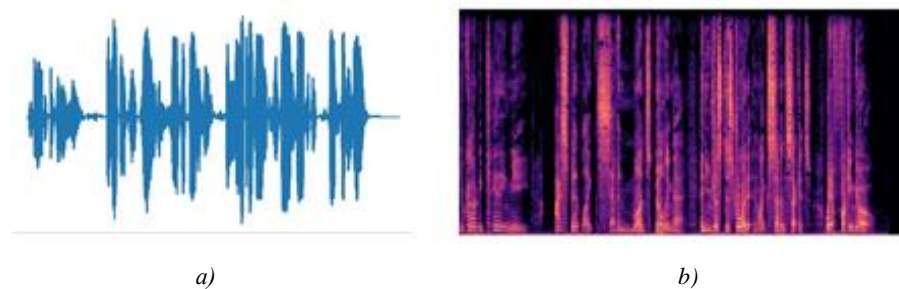


Рис. 2. Частотная синусоида (*a*) и спектрограмма (*b*)
Fig. 2. Frequency sinusoid (*a*) and spectrogram (*b*)

Для реализации разрабатываемого алгоритма было решено создать структуру нейронной сети, включающую следующие слои:

- Conv2D – три сверточных слоя;
- MaxPooling2D – слой выделения признаков;
- Dropout – два слоя коррекции (вносят случайную величину в веса нейронов);
- Flatten – конвертер из сверточной структуры в многослойную;
- Dense – два слоя многоосной нейронной сети.

В соответствии со структурой нейронной сети построена ее концептуальная модель (рис. 3).

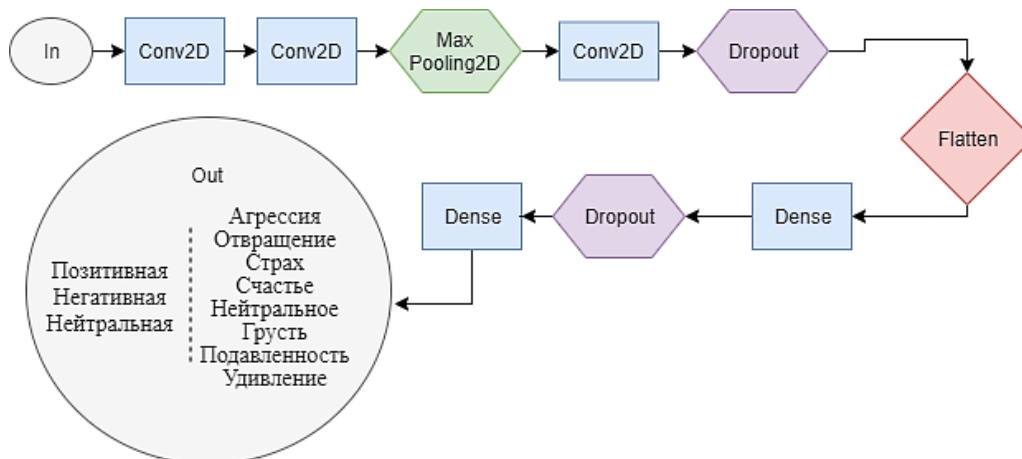


Рис. 3. Концептуальная модель нейронной сети

Fig. 3. Conceptual model of neural network

В итоге созданы две модели для каждого из подходов с различным количеством выходных нейронов. Модели для обоих подходов имеют одинаковую структуру.

Для разработанного проекта были созданы две модели сверточной нейронной сети: для восьми (подробных) и трех (общих) классов. На рис. 4 показана модель для общих классов эмоций. Обе модели имеют полностью одинаковые структуры за исключением выхода, количество нейронов на котором должно быть равно количеству классов.

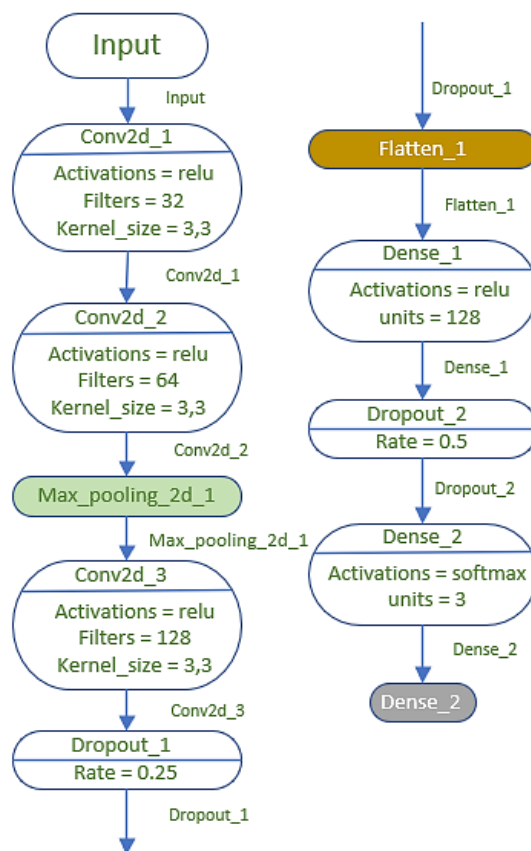


Рис. 4. Модель для трех классов

Fig. 4. Model for three classes

Эксперимент был реализован с помощью следующего набора инструментов:
 TensorFlow – открытой программной библиотеки для машинного обучения;
 Keras – нейросетевой библиотеки;
 Librosa – аудиобиблиотеки для анализа звуковых сигналов;
 PyAudio – модуля с кроссплатформенной библиотекой PortAudio, позволяющего проигрывать и записывать звуки;
 Pillow – библиотеки для работы с изображениями;
 NumPy – библиотеки языка программирования Python для реализации вычислительных алгоритмов, оптимизированной для работы с многомерными массивами (для ускорения вычислительных процессов);
 SciPy – библиотеки языка программирования Python для выполнения научных и инженерных расчетов.

Результаты первоначального исследования. Графики точности и ошибки для восьми классов эмоций показаны на рис. 5, *a*, для трех классов – на рис. 5, *b*.

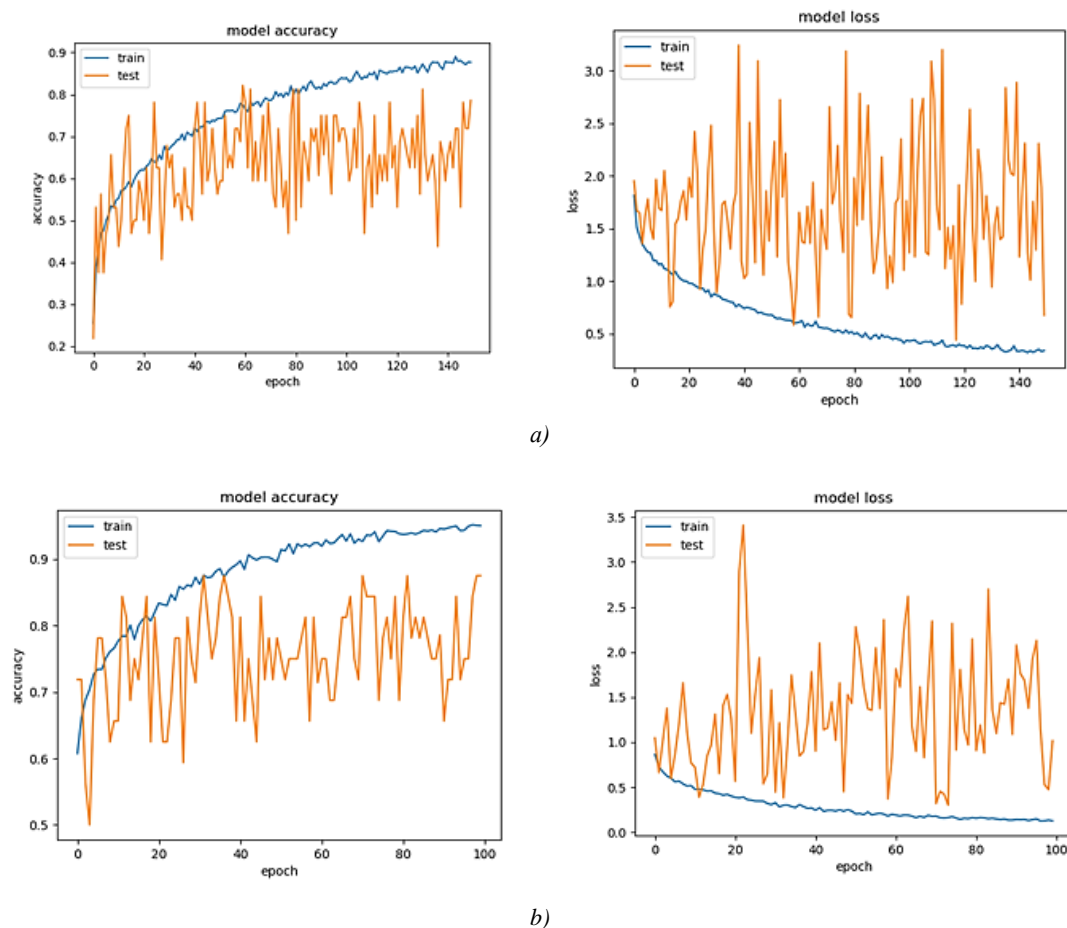


Рис. 5. График точности и график ошибок для восьми классов эмоций (*a*), для трех классов (*b*)

Fig. 5. Accuracy plot and error plot for eight emotion classes (a), for three classes (b)

Модели имеют идентичные структуры и состоят из практически одинаковых наборов файлов: `train` – обеспечивает процесс обучения нейронной сети, `test` – необходим для выполнения тестирования на наборе данных, `single_test` – представляет собой рабочий исполняемый файл, используемый для анализа эмоционального состояния. Принципиальное различие моделей заключается только в выходах, количество нейронов на выходе соответствует количеству классов.

Тестирование выполнялось в несколько этапов: классификация групп (тестирование набора данных) по классам, процентная оценка и тестирование каждого файла в отдельности.

Тестирование набора данных осуществлялось с помощью тестов по 100 файлов в каждой категории. На рис. 6, *a* показаны результаты тестирования для восьми классов эмоций, на рис. 6, *b* – для трех классов.

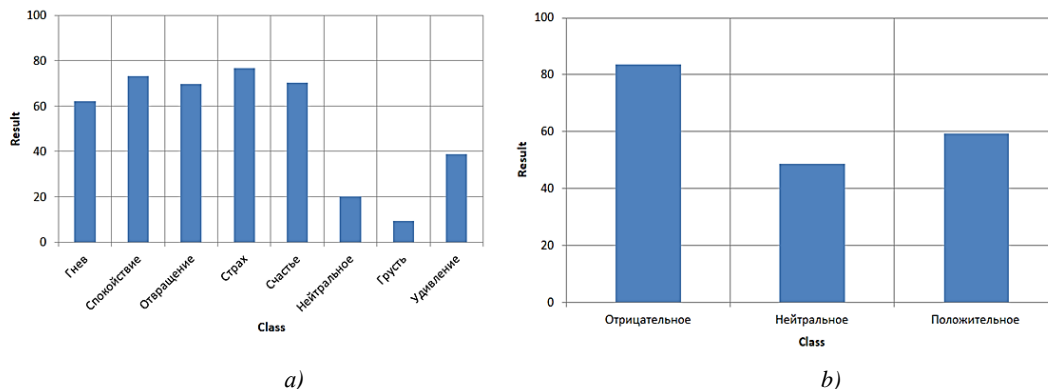


Рис. 6. Точность классификации для восьми классов эмоций (*a*), для трех классов (*b*)

Fig. 6. Classification accuracy for eight emotion classes (*a*), for three classes (*b*)

При тестировании отдельных образцов для трех классов в качестве теста были взяты следующие эмоции: агрессия, счастье, отвращение, страх, нейтральное состояние, грусть и удивление. Результаты исследований показали, что корректно распознаются не все эмоции. Нейронная сеть хорошо распознает агрессию, счастье, отвращение, страх и грусть. Удивление и нейтральное состояние распознаются некорректно, нейронная сеть относит их к негативным эмоциям.

При тестировании отдельных образцов для восьми классов в качестве теста были использованы такие же параметры. Результаты исследований показали, что 100%-го попадания в класс не было. Нейронная сеть хорошо распознает счастье, отвращение, грусть и удивление, вместо класса «агрессия» получен результат «страх», вместо «нейтральное состояние» – «счастье», а вместо «страх» – «отвращение».

Поиск решения для повышения точности алгоритма. На рис. 6 показано, что тестирование не является однозначным, так как человеческие эмоции определяются сложно. Набор данных не может показать определенный результат, поэтому был выполнен отдельный анализ для каждого тестового образца. Определение обобщенных классов дало более точный результат, чем определение конкретных классов. Стоит заметить, что из-за меньшего количества классов скорость работы программы значительно больше при классификации трех классов, чем при классификации восьми классов, поэтому программа будет требовать для работы меньше процессорного времени.

Согласно рис. 6 получены следующие показатели точности, %:

1. Распознавание трех эмоций:

отрицательная – 83,5;

нейтральная – 48,7;

положительная – 59,3.

2. Распознавание восьми эмоций:

гнев – 62,1;

спокойствие – 73,3;

отвращение – 69,8;

страх – 76,8;

счастье – 70,3;

нейтральное – 20,05;

грусть – 9,4;

удивление – 38,7.

Видно, что точность мала, алгоритм с полученными результатами будет иметь высокую погрешность. Низкая точность может быть обусловлена ошибками первого и второго рода при распознавании.

Следующим этапом исследования стал поиск путей для увеличения точности работы алгоритма. Для этого сначала потребовалась проверка работоспособности готовых нейронных сетей с имеющимся набором данных. Для тестирования алгоритма было решено перейти в среду для разработки Matlab. Были выбраны нейронные сети GoogleNet и Rasnet-50. Они зарекомендовали себя как одни из самых лучших структур, применяемых при классификации изображений. Для проведения исследования эти структуры нейронных сетей были переобучены на имеющийся набор данных. По итогам обучения обе сети дали очень похожие результаты (рис. 7).

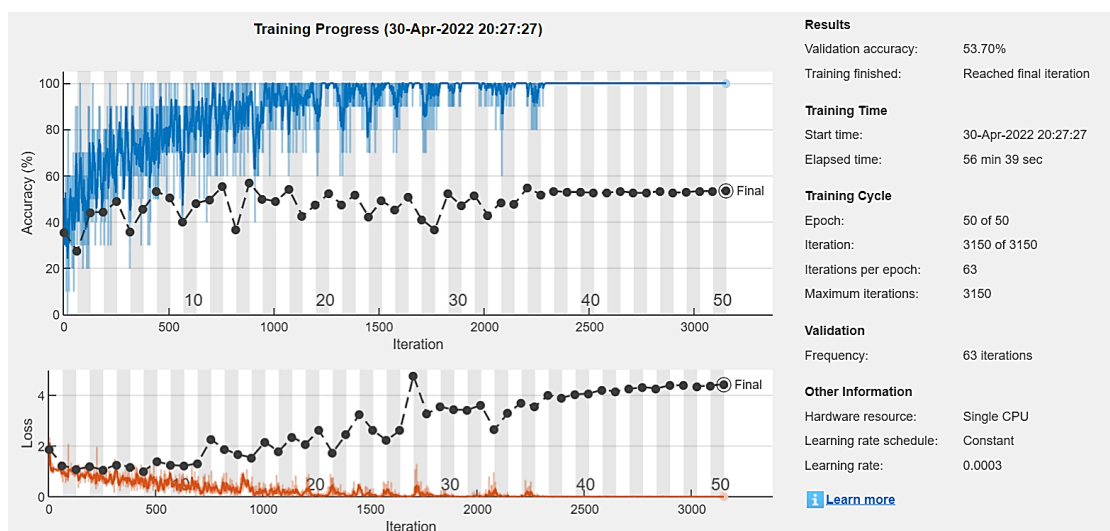


Рис. 7. Результаты обучения готовых структур нейронных сетей

Fig. 7. Results of training of ready-made structures of neural networks

В результате исследования была установлена жесткая зависимость точности работы алгоритма и плохого качества входных данных. Следовательно, необходимо было изменить входной набор данных. Для повышения точности было решено использовать mel-спектрограммы. На рис. 8 показан аудиосигнал и соответствующая ему mel-спектрограмма.

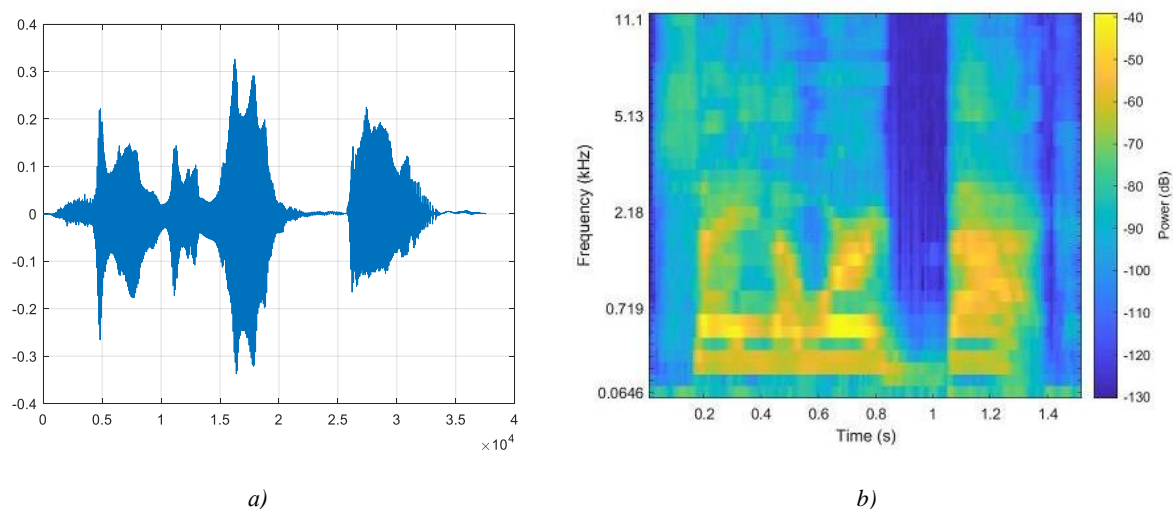
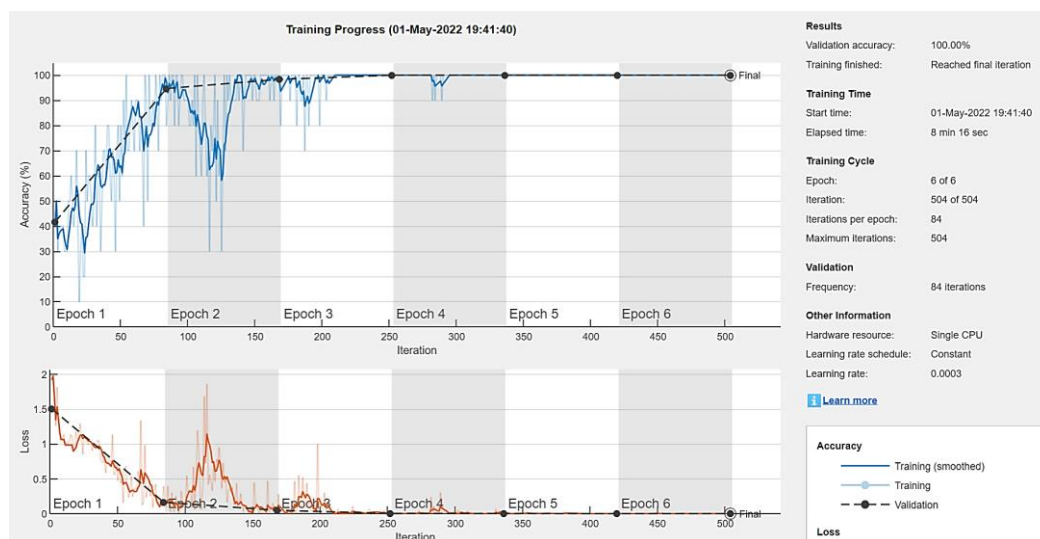


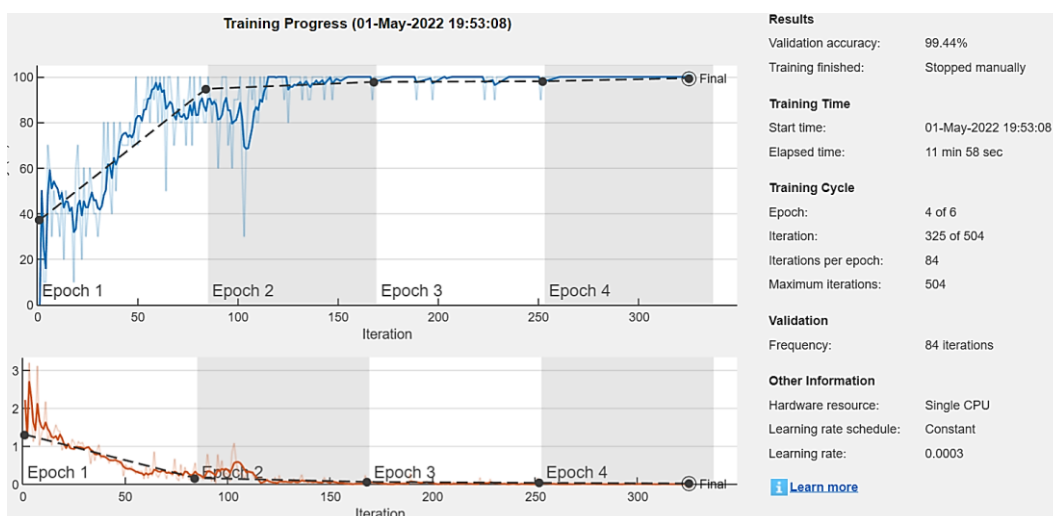
Рис. 8. График аудиосигнала (a), MFCC-спектрограмма (b)

Fig. 8. Audio signal plot (a), MFCC spectrogram (b)

Все входные данные, полностью преобразованные к виду MFCC-спектрограммы, были поданы вновь на готовые нейронные сети. В результате удалось достичь 100 %-й точности распознавания трех эмоций на тренировочном наборе данных при использовании RasNet-50 (рис. 9, *a*) и 99,44 %-й точности при использовании GoogleNet (рис. 9, *b*).



a)



b)

Рис. 9. Графики обучения нейронных сетей Rasnet-50 (*a*) и GoogleNet (*b*)

Fig. 9. Training graphs for Rasnet-50 (a) and GoogleNet (b) neural networks

Разработка усовершенствованной модели распознавания и анализ полученных данных. После окончания текущего исследования было решено разработать новую структуру нейронной сети, состоящую из 24 слоев, с использованием программного пакета Matlab (рис. 10).

На вход разработанной нейронной сети подавались MFCC-спектрограммы. Полученные данные были разделены следующим образом: 60 % для обучения, 20 % для валидации, 20 % для проверки точности работы (эти данные не использовались при обучении). На рис. 11, *a* приведены графики обучения нейронной сети для распознавания трех эмоций (точность на обучающем наборе данных составила 100 %), на рис. 11, *b* – для распознавания семи эмоций (точность на обучающем наборе данных составила 99,82 %).

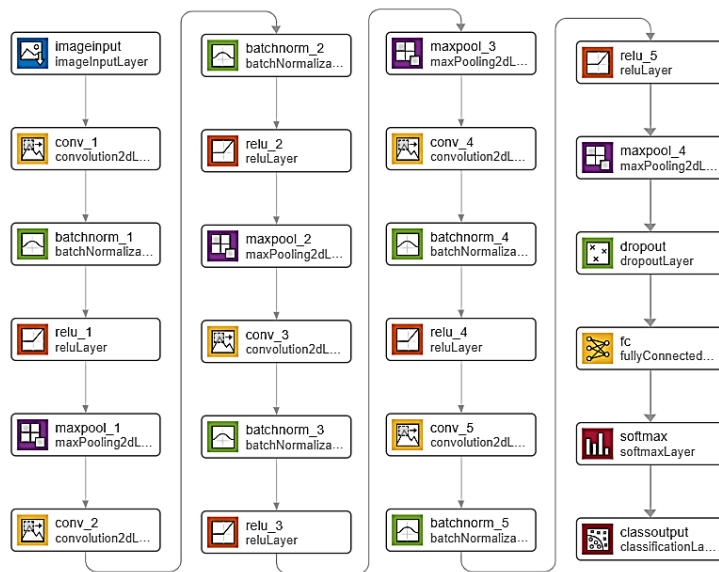
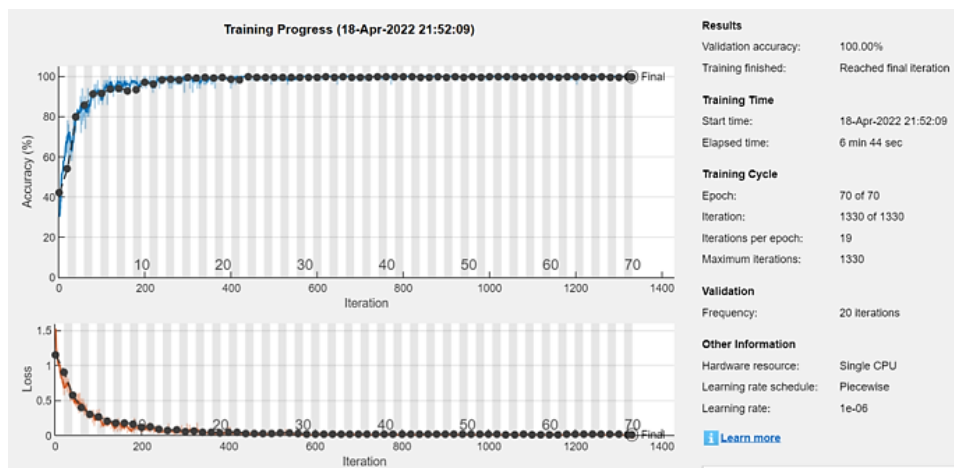
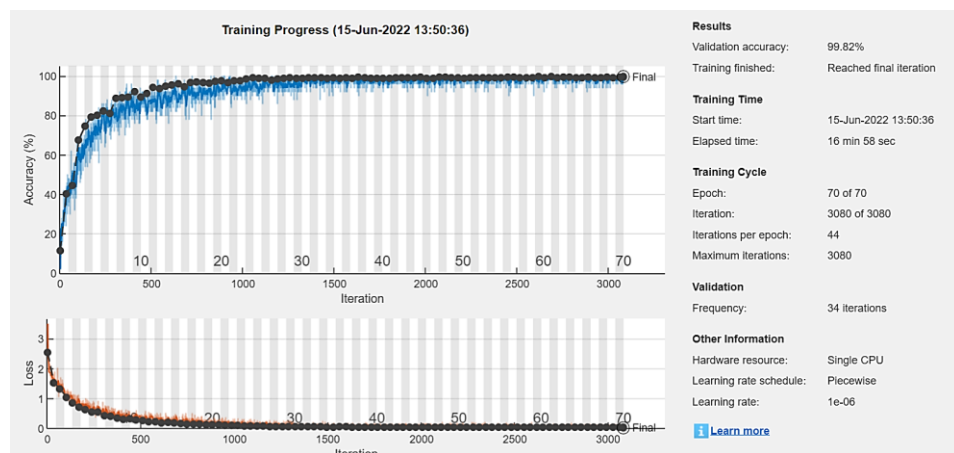


Рис. 10. Структура нейронной сети
Fig. 10. Neural network structure



a)



b)

Рис. 11. Результат обучения нейронной сети для трех эмоций (a), для семи эмоций (b)
Fig. 11. The result of neural network training for three emotions (a), for seven emotions (b)

После обучения для проверки работоспособности на вход нейронной сети подавались данные, не участвовавшие в обучении. Для трех эмоций (рис. 12, а) точность составила 98,33 %, а для семи эмоций (рис. 12, б) – 92,87 %.

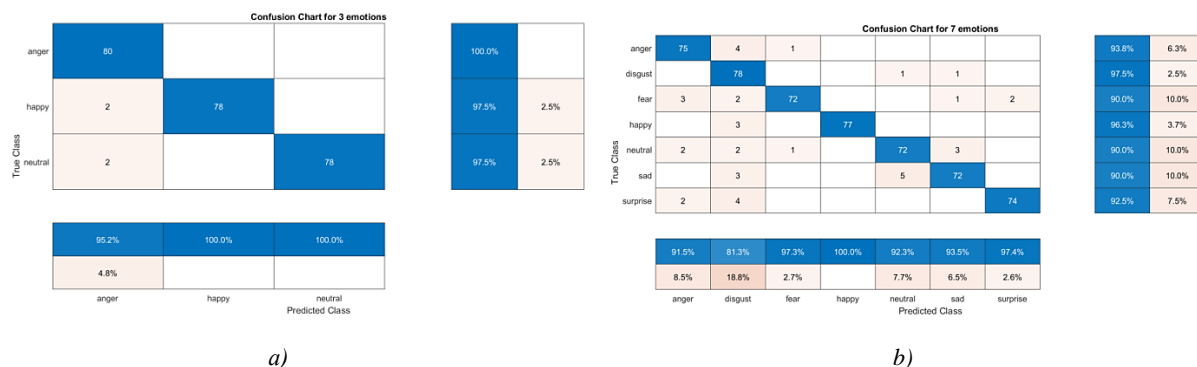


Рис. 12. Точность работы алгоритма на наборе, не участвовавшем в обучении: а) для трех эмоций; б) для семи эмоций

Fig. 12. The accuracy of the algorithm on the set that did not participate in training: a) for three emotions; b) for seven emotions

Закключение. Авторами был разработан алгоритм, на вход которого подавались спектрограммы, полученные в результате оконного преобразования Фурье. Однако точность полученного алгоритма была слишком мала, чтобы завершить на этом процесс исследования. Для повышения точности алгоритма было решено провести исследование, позволяющее определить зависимость результата от входных параметров или структуры нейронной сети. Изменение структуры нейронной сети при неизменных входных данных не повысило точность алгоритма, что навело на мысль о необходимости изменения вида входных данных. Для этого входной набор данных был преобразован к виду MFCC, что в последующем показало зависимость точности работы алгоритма от входного набора данных.

Для разработки новой структуры нейронной сети был использован программный пакет Matlab, с помощью которого удалось получить высокие эргономические параметры исследуемой области. В качестве актуальности данной тематики можно отметить высокую точность работы предложенного алгоритма по сравнению с имеющимися на текущий момент вариантами.

Вклад авторов. В. В. Семенюк осуществила постановку задачи, определила направление и цель исследования, разработала структуру нейронной сети с использованием языка программирования Python. М. В. Складчиков проанализировал полученные данные, разработал концепцию, позволяющую улучшить точность алгоритма путем выявления факторов, влияющих на ее показатели, разработал архитектуру нейронной сети в программе Matlab и провел соответствующие эксперименты, осуществил научное редактирование статьи.

Список использованных источников

1. Mesaros, A. Acoustic scene classification: Overviews of DCASE 2017 challenge entries / A. Mesaros, T. Heittola, T. Virtanen // 16th Intern. Workshop on Acoustic Signal Enhancement (IWAENC 2018), Tokyo, Japan, 17–20 Sept. 2018. – Tokyo, 2018. – P. 411–415.
2. Haitsma, J. A highly robust audio fingerprinting system / J. Haitsma, T. Kalker // 3rd Intern. Conf. on Music Information Retrieval, Paris, France, 13–17 Oct. 2002. – Paris, 2002. – P. 107–115.
3. Ильин, Е. П. Эмоции и чувства / Е. П. Ильин. – СПб. : Питер, 2001. – 752 с.
4. Изард, К. Э. Психология эмоций / К. Э. Изард. – СПб. : Питер, 2012. – 464 с.
5. Карелина, И. О. Развитие понимания эмоций в период дошкольного детства: психологический ракурс : монография / И. О. Карелина. – Прага : Vědecko vydavatelské centrum «Sociosféra-CZ», 2017. – 178 с.
6. Орехова, О. А. Цветовая диагностика эмоций. Типология развития : монография / О. А. Орехова. – СПб. : Речь; М. : Сфера, 2008. – 176 с.

7. Шаповал, Ж. Я. Распознавание эмоций человека по изображению как часть автоматизированного переводчика языка жестов / Ж. Я. Шаповал // Молодежный научно-технический вестник. – 2017. – № 7. – С. 55.
8. Голубинский, А. Н. Выявление эмоционального состояния человека по речевому сигналу на основе вейвлет-анализа / А. Н. Голубинский // Вестник Воронежского института МВД России. – 2011. – № 3. – С. 144–153.
9. Сидоров, К. И. Автоматическое распознавание эмоций человека на основе реконструкций аттракторов образцов речи / К. И. Сидоров, Н. Н. Филатова // Программные системы и вычислительные методы. – 2012. – № 1. – С. 67–79.
10. Галичий, Д. А. Распознавание эмоций человека при помощи современных методов глубокого обучения / Д. А. Галичий, Г. И. Афанасьев, Ю. Г. Нестеров // E-SCIO. – 2021. – Т. 5, № 56. – С. 316–329.
11. Бредихин, А. И. Применение вейвлетов в задаче распознавания эмоций человека по его речи / А. И. Бредихин // Сборник избранных статей научной сессии ТУСУР. – 2018. – № 1–3. – С. 115–119.
12. Рюмина, Е. В. Аналитический обзор методов распознавания эмоций по выражениям лица человека / Е. В. Рюмина, А. А. Карпов // Научно-технический вестник информационных технологий, механики и оптики. – 2020. – Т. 20, № 2. – С. 163–176.
13. Dvoynikova, A. Emotion recognition and sentiment analysis of extemporaneous speech transcriptions in Russian / A. Dvoynikova, O. Verkholyak, A. Karpov // Lectures notes in computer science. – 2020. – Vol. 12335. – P. 136–144. https://doi.org/10.1007/978-3-030-60276-5_14
14. Devi, J. S. Speaker emotion recognition based on speech features and classification techniques / J. S. Devi, S. Yarrammelle, S. P. Nandyala // Intern. J. of Image, Graphics, and Signal Processing. – 2014. – Vol. 6, no. 7. – P. 61–77. <https://doi.org/10.5815/ijigsp.2014.07.08>
15. Speech emotion recognition based on an improved brain emotion learning model / Z. I. Liu [et al.] // Neurocomputing. – 2018. – Vol. 309. – P. 145–156. <https://doi.org/10.1016/j.neucom.2018.05.005>
16. Shirami, A. Speech emotion recognition based on SVM as both features selector and classifier / A. Shirami, A. R. N. Nilchi // Intern. J. of Image, Graphics, and Signal Processing. – 2016. – Vol. 8, no. 4. – P. 39–45. <https://doi.org/10.5815/ijigsp.2016.04.05>
17. Assuncao, G. Intermediary fuzzyfication in speech emotion recognition / G. Assuncao, P. Menezes // IEEE Intern. Conf. on Fuzzy System, Glasgow, United Kingdom, 19–24 July 2020. – Glasgow, 2020. – P. 9177699. <https://doi.org/10.1109/FUZZ48607.2020.9177699>
18. Zisad, S. N. Speech emotion recognition in neurological disorders using convolutional neural network / S. N. Zisad, M. S. Hossain, K. Andersson // Lecture Notes in Computer Science. – 2020. – Vol. 12241. – P. 287–296. https://doi.org/10.1007/978-3-030-59277-6_26
19. Werner, S. Speech emotion recognition: humans vs machines / S. Werner, G. K. Petrenko // Discourse. – 2019. – Vol. 5, no. 5. – P. 136–152. <https://doi.org/10.32603/2412-8562-2019-5-5-136-152>
20. Muppidi, A. Speech emotion recognition using quaternion convolutional neural networks / A. Muppidi, M. Radfar // IEEE Intern. Conf. of Acoustics, Speech and Signal Processing-Proceedings, Toronto, ON, Canada, 6–11 June 2021. – Toronto, 2021. – P. 6309–6313. <https://doi.org/10.1109/ICASSP39728.2021.9414248>
21. Zheng, W. Multi-scale discrepancy adversarial network for crosscorpus speech emotion recognition / W. Zheng, Y. Zong // Virtual Reality and Intelligent Hardware. – 2021. – Vol. 3, no. 1. – P. 65–75. <https://doi.org/10.1016/j.vrih.2020.11.006>
22. Hazjan, V. Context-independent multilingual emotion recognition from speech signals / V. Hazjan, Z. Kacic // Intern. J. of Speech Technology. – 2003. – Vol. 6, no. 3. – P. 311–320.
23. Zhang, C. Autoencoder with emotion embedding for speech emotion recognition / C. Zhang, L. Xue // IEEE Access. – 2021. – Vol. 9. – P. 51231–51241. <https://doi.org/10.1109/ACCESS.2021.3069818>
24. Kanwal, S. Speech emotion recognition using clustering based GA-optimized feature set / S. Kanwal, S. Asghar // IEEE Access. – 2021. – Vol. 9. – P. 125830–125842. <https://doi.org/10.1109/ACCESS.2021.3111659>
25. Byoung, C. K. A brief review of facial emotion recognition based on visual information / C. K. Byoung // Sensors. – 2018. – Vol. 18, iss. 2. – P. 401. <https://doi.org/10.3390/s18020401>
26. Audio-visual emotion recognition using deep transfer learning and multiple temporal models / X. Ouyang [et al.] // ICMI '17 : Proc. of the 19th ACM Intern. Conf. on Multimodal Interaction, Glasgow, United Kingdom, 13–17 November 2017. – Glasgow, 2017. – P. 577–582. <https://doi.org/10.1145/3136755.3143012>
27. Hassani, B. Facial expression recognition using enhanced deep 3D convolutional neural networks / B. Hassani, M. H. Mahoor // 2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017. – Honolulu, 2017. – P. 1955–1962. <https://doi.org/10.1109/CVPRW.2017.282>

References

1. Mesaros A., Heittola T., Virtanen T. Acoustic scene classification: Overviews of DCASE 2017 challenge entries. *16th International Workshop on Acoustic Signal Enhancement (IWAENC 2018), Tokyo, Japan, 17–20 September 2018*. Tokyo, 2018, pp. 411–415.
2. Haitsma J., Kalker T. A highly robust audio fingerprinting system. *3rd International Conference on Music Information Retrieval, Paris, France, 13–17 October 2002*. Paris, 2002, pp. 107–115.
3. Ilin E. P. Jemocii i chuvstva. *Emotions and Feelings*. Saint Petersburg, Piter, 2001, 752 p. (In Russ.).
4. Izard K. E. Psihologija jemocij. *Psychology of Emotions*. Saint Petersburg, Piter, 2012, 464 p. (In Russ.).
5. Karelina I. O. Razvitie ponimaniya jemocij v period doskol'nogo detstva: psihologicheskij rakurs. *Developing an Understanding of Emotions during Preschool Childhood: A Psychological Perspective*, Prague, Vědecko vydavatelské centrum "Sociosféra-CZ", 2017, 178 p. (In Russ.).
6. Orehova O. A. Cvetovaja diagnostika jemocij. Tipologija razvitija. Monografija. *Color Diagnostics of Emotions. Typology of Development. Monograph*. Saint Petersburg, Sphere, 2008, 176 p. (In Russ.).
7. Shapoval J. A. Recognition of Human Emotions by image as part of an automated sign language translator. *Molodezhnyj nauchno-tehnicheskij vestnik [Youth Scientific and Technical Bulletin]*, 2017, no. 7, p. 55 (In Russ.).
8. Golubinskij A. N. Identification of a person's emotional state by a speech signal based on a Wavelet analysis. *Vestnik Voronezhskogo instituta Ministerstva vnutrennih del Rossii [Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia]*, 2011, no. 3, pp. 144–153 (In Russ.).
9. Sidorov K. I., Filatova N. N. Automatic recognition of human emotions based on reconstructions of attractors of speech samples. *Programmnye sistemy i vychislitel'nye metody [Software systems and computational methods]*, 2012, no. 1, pp. 67–79 (In Russ.).
10. Galichij D. A., Afanaciev G. I., Nesterov U. G. Recognition of human emotions using modern methods of deep learning. *E-SCIO*, 2021, vol. 5, no. 56, pp. 316–329 (In Russ.).
11. Bredihin A. I. The use of wavelets in the task of recognizing a person's emotions by his speech. *Sbornik izbrannyh statej nauchnoj sessii Tomskogo gosudarstvennogo universiteta sistem upravlenija i radioelektroniki [Collection of selected articles of the scientific session of Tomsk State University of Control Systems and Radioelectronics]*, 2018, no. 1–3, pp. 115–119 (In Russ.).
12. Rumina E. V., Karpov A. A. Analytical review of emotion recognition methods based on human facial expressions. *Nauchno-tehnicheskij vestnik informacionnyh tekhnologij, mekhaniki i optiki [Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics]*, 2020, vol. 20, no. 2, pp. 163–176 (In Russ.). <https://doi.org/10.17586/2226-1494-2020-20-2-163-176>
13. Dvoynikova A., Verkholyak O., Karpov A. Emotion recognition and sentiment analysis of extemporaneous speech transcriptions in Russian. *Lectures Notes in Computer Science*, 2020, vol. 12335, pp. 136–144. https://doi.org/10.1007/978-3-030-60276-5_14
14. Devi J. S., Yarrammelle S., Nandyala S. P. Speaker emotion recognition based on speech features and classification techniques. *International Journal of Image, Graphics, and Signal Processing*, 2014, vol. 6, no. 7, pp. 61–77. <https://doi.org/10.5815/ijigsp.2014.07.08>
15. Liu Z. I., Xie Q., Wu M., Cao W. H., Mao J. W., Mei Y. Speech emotion recognition based on an improved brain emotion learning model. *Neurocomputing*, 2018, vol. 309, pp. 145–156. <https://doi.org/10.1016/j.neucom.2018.05.005>
16. Shirami A., Nilchi A. R. N. Speech emotion recognition based on SVM as both features selector and classifier. *International Journal of Image, Graphics, and Signal Processing*, 2016, vol. 8, no. 4, pp. 39–45. <https://doi.org/10.5815/ijigsp.2016.04.05>
17. Assuncao G., Menezes P. Intermediary fuzzyfication in speech emotion recognition. *IEEE International Conference on Fuzzy System, Glasgow, United Kingdom, 19–24 July 2020*. Glasgow, 2020, p. 9177699. <https://doi.org/10.1109/FUZZ48607.2020.9177699>
18. Zisad S. N., Hossain M. S., Andersson K. Speech emotion recognition in neurological disorders using convolutional neural network. *Lecture Notes in Computer Science*, 2020, vol. 12241, pp. 287–296. https://doi.org/10.1007/978-3-030-59277-6_26
19. Werner S., Petrenko G. K. Speech emotion recognition: humans vs machines. *Discourse*, 2019, vol. 5, no. 5, pp. 136–152. <https://doi.org/10.32603/2412-8562-2019-5-5-136-152>
20. Muppidi A., Radfar M. Speech emotion recognition using quaternion convolutional neural networks. *IEEE International Conference of Acoustics, Speech and Signal Processing-Proceedings, Toronto, ON, Canada, 6–11 June 2021*. Toronto, 2021, pp. 6309–6313. <https://doi.org/10.1109/ICASSP39728.2021.9414248>

21. Zheng W., Zong Y. Multi-scale discrepancy adversarial network for crosscorpus speech emotion recognition. *Virtual Reality and Intelligent Hardware*, 2021, vol. 3, no. 1, pp. 65–75. <https://doi.org/10.1016/j.vrih.2020.11.006>
22. Hazjan V., Kacic Z. Context-independent multilingual emotion recognition from speech signals. *International Journal of Speech Technology*, 2003, vol. 6, no. 3, pp. 311–320.
23. Zhang C., Xue L. Autoencoder with emotion embedding for speech emotion recognition. *IEEE Access*, 2021, vol. 9, pp. 51231–51241. <https://doi.org/10.1109/ACCESS.2021.3069818>
24. Kanwal S., Asghar S. Speech emotion recognition using clustering based GA-optimized feature set. *IEEE Access*, 2021, vol. 9, pp. 125830–125842. <https://doi.org/10.1109/ACCESS.2021.3111659>
25. Byoung C. K. A brief review of facial emotion recognition based on visual information. *Sensors*, 2018, vol. 18, iss. 2, pp. 401. <https://doi.org/10.3390/s18020401>
26. Ouyang X., Kawai S., Goh E. G. H., Shen S., Ding W., ..., D.-Y. Huang. Audio-visual emotion recognition using deep transfer learning and multiple temporal models. *ICMI '17 : Proceedings of the 19th ACM International Conference on Multimodal Interaction, Glasgow, United Kingdom, 13–17 November 2017*. Glasgow, 2017, pp. 577–582. <https://doi.org/10.1145/3136755>. 3143012
27. Hassani B., Mahoor M. H. Facial expression recognition using enhanced deep 3D convolutional neural networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017*. Honolulu, 2017, pp. 1955–1962. <https://doi.org/10.1109/CVPRW.2017.282>

Информация об авторах

Семенюк Виктория Валерьевна, магистр технических наук, преподаватель специальных дисциплин, Донецкий техникум промышленной автоматизации имени А. В. Захарченко.
E-mail: semenuk.viktoriya@gmail.com

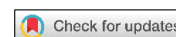
Складчиков Максим Владимирович, магистр технических наук, преподаватель специальных дисциплин, Донецкий техникум промышленной автоматизации имени А. В. Захарченко.
E-mail: maxsklad19981@yandex.ru

Information about the authors

Viktoriya V. Semenuk, M. Sc. (Eng.), Teacher of Special Disciplines, Donetsk Technical School of Industrial Automation after A. V. Zakharchenko.
E-mail: semenuk.viktoriya@gmail.com

Maxim V. Skladchikov, M. Sc. (Eng.), Teacher of Special Disciplines, Donetsk Technical School of Industrial Automation after A. V. Zakharchenko.
E-mail: maxsklad19981@yandex.ru

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ COMPUTER-AIDED DESIGN



УДК 004.65: 621.7/.9
<https://doi.org/10.37661/1816-0301-2022-19-4-69-83>

Оригинальная статья
Original Paper

Информационное обеспечение САПР технологических процессов опытных образцов структурно-сложных изделий

А. В. Петухов

*Гомельский государственный технический университет
имени П. О. Сухого,
пр. Октября, 48, Гомель, 246746, Беларусь
E-mail: Petukhov_2000@gstu.by*

Аннотация

Цели. Целью исследования информационного обеспечения системы автоматизированного проектирования технологических процессов (САПР ТП) опытных образцов структурно-сложных изделий является формализация принятия технологических решений с использованием взаимосвязанных баз данных.

Методы. В работе использовался метод декомпозиции процессов проектирования технологий.

Результаты. При исследовании информационного обеспечения САПР ТП опытных образцов структурно-сложных изделий решена задача определения структуры и последовательности применения такого обеспечения на различных стадиях внедрения системы, разработана универсальная схема формирования операционного описания технологического процесса.

Заключение. Использование результатов проведенного исследования при создании САПР ТП опытных образцов структурно-сложных изделий обеспечило повышение производительности технологического проектирования в 7,8 раза.

Ключевые слова: информационное обеспечение, система автоматизированного проектирования, технологический процесс, образцы структурно-сложных изделий, декомпозиция процессов проектирования, взаимосвязанные базы данных

Для цитирования. Петухов, А. В. Информационное обеспечение САПР технологических процессов опытных образцов структурно-сложных изделий / А. В. Петухов // Информатика. – 2022. – Т. 19, № 4. – С. 69–83. <https://doi.org/10.37661/1816-0301-2022-19-4-69-83>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 09.08.2022
Подписана в печать | Accepted 26.09.2022
Опубликована | Published 29.12.2022

Information support of CAD of technological processes of structurally complex product prototypes

Alexander V. Petukhov

*Sukhoi State Technical University of Gomel,
av. Octiabria, 48, Gomel, 246746, Belarus
E-mail: Petukhov_2000@gstu.by*

Abstract

Objectives. The purpose of the study of information support for computer-aided process planning system of prototypes of structurally complex products is to formalize the adoption of technological decisions using interconnected databases.

Methods. The method of decomposition of technology design processes was used in the work.

Results. In the study of information support for computer-aided process planning system of prototypes of structurally complex products, the problem of determining the structure of information support was solved, the sequence of using information support at various stages of system implementation was determined and universal scheme for the formation of operational description of the technological process was developed.

Conclusion. The use of the results of the study in the creation of computer-aided process planning system for prototypes of structurally complex products could increase the productivity of technological design in 7.8 times.

Keywords: information support, computer-aided design system, technological process, samples of structurally complex products, decomposition of design processes, interconnected databases

For citation. Petukhov A. V. *Information support of CAD of technological processes of structurally complex product prototypes*. Informatika [Informatics], 2022, vol. 19, no. 4, pp. 69–83 (In Russ.).
<https://doi.org/10.37661/1816-0301-2022-19-4-69-83>

Conflict of interest. The author declares of no conflict of interest.

Введение. При жестких сроках проектирования и освоения производства структурно-сложных изделий особое место в их жизненном цикле занимает изготовление опытных образцов. С одной стороны, этот процесс венчает конструкторскую разработку, а с другой – предваряет технологическую подготовку основного производства [1].

Несмотря на важность вопросов разработки научно обоснованной модели технологической подготовки производства (ТПП) опытных образцов, они до сих пор не нашли достойного отражения в научных исследованиях. Отчасти это объясняется тем, что информация по данной теме является закрытой, так как может быть использована в конкурентной борьбе производителей.

Требования к информационному обеспечению САПР ТП опытных образцов структурно-сложных изделий продиктованы спецификой технологического проектирования в условиях экспериментального производства и сводятся к следующему:

1. Перечень деталей и сборочных единиц собственного изготовления САПР ТП опытных образцов структурно-сложных изделий должна получать из автоматизированной системы структурного анализа (АССА).

2. На вход САПР ТП опытных образцов структурно-сложных изделий должна подаваться информация в виде 2D-чертежей деталей или сборочных единиц (на бумажном или машинном носителе). При этом необходимо учитывать, что разработка управляющих программ для программно-управляемого оборудования значительно ускоряется при использовании на входе 3D-моделей деталей.

3. Несмотря на то что изготовление опытного образца осуществляется в условиях единичного производства, САПР ТП опытных образцов структурно-сложных изделий должна обладать информационным обеспечением, позволяющим выбирать степень детализации описания технологического процесса. В экспериментальном производстве работают исполнители высокой квалификации, способные самостоятельно решать некоторые технологические вопросы. По-

этому степень детализации описания технологического процесса может варьироваться от маршрутного до операционного и зависит от сложности изготавливаемой детали. Кроме того, информационное обеспечение должно поддерживать процесс формирования специальной формы описания – маршрутной ведомости. В этом документе мастер зафиксировывает табельный номер исполнителя, а сотрудник отдела технического контроля – количество годных деталей.

4. Информационное обеспечение САПР ТП и АССА должно обеспечить возможность формирования итоговых документов по материальным и трудовым затратам на изготовление определенного опытного образца.

Перечисленные требования определили цель исследования информационного обеспечения САПР ТП опытных образцов структурно-сложных изделий. Она заключается в формализации принятия технологических решений с использованием взаимосвязанных баз данных и метода декомпозиции процессов проектирования технологий.

Для реализации названной цели при исследовании информационного обеспечения САПР ТП опытных образцов структурно-сложных изделий решаются задачи:

- определения структуры информационного обеспечения;
- определения последовательности использования информационного обеспечения на различных стадиях внедрения системы;
- разработки универсальной схемы формирования операционного описания технологического процесса;
- определения порядка применения информационного обеспечения САПР ТП опытных образцов при формировании операционного описания технологического процесса;
- формирования описания средств технологической оснастки с использованием информационного обеспечения САПР ТП опытных образцов структурно-сложных изделий;
- разработки структуры архива технологических процессов.

Структура исследуемого информационного обеспечения. Важным аспектом, который необходимо учитывать при создании САПР ТП опытных образцов структурно-сложных изделий, в частности зерноуборочной и кормоуборочной техники, является форма представления информации о сборочных единицах и деталях, входящих в состав изделия. Учитывая современный уровень развития автоматизации конструкторского проектирования, в Научно-техническом центре комбайностроения ОАО «Гомсельмаш» было принято решение использовать гибридную форму представления такого вида информации – на машинных и бумажных носителях.

Для повышения степени адаптируемости системы к изменяющимся производственным условиям вся справочная информация, отражающая специфические условия опытного производства, должна быть представлена в виде справочников, изменения в которых не приводят к перепрограммированию ядра системы [2].

Информационное обеспечение САПР ТП опытных образцов зерноуборочной и кормоуборочной техники представляет собой в совокупности:

- массивы данных справочной технологической информации;
- языковые средства и процедуры;
- процедуры и средства управления массивами данных.

Языковые средства и процедуры обеспечивают преобразование массивов данных, представленных на естественном (или другом) языке описания объекта, в данные на языке машины.

Таким образом, под информационным обеспечением автоматизированного технологического проектирования следует понимать процесс представления информации по запросам САПР ТП опытных образцов зерноуборочной и кормоуборочной техники. Основной задачей информационного обеспечения автоматизированного технологического проектирования является полное, достоверное, надежное и своевременное поступление всей необходимой информации для решения задач САПР ТП при условии минимального времени ожидания и наименьших затрат как на ее создание, так и на эксплуатацию [3].

Информацию, используемую в САПР ТП, условно можно разделить на входную (исходную) и выходную (производную) [4]. Входной называется информация, существующая до начала автоматизированного проектирования. Она подразделяется на переменную и условно-постоянную.

К переменной относится следующая информация о конкретной детали:

- 1) обозначение и наименование детали (сборочной единицы);
- 2) сведения о материале для изготовления детали (о составе сборочной единицы);
- 3) объем производства и количество деталей в партии;
- 4) сведения о термической обработке и виде покрытия;
- 5) общие параметры детали (габаритные размеры, точность изготовления, шероховатость поверхности и т. д.);
- 6) форма, размеры, точностные и качественные характеристики отдельных поверхностей детали и их взаимное расположение.

Данная информация вводится в оперативное запоминающее устройство каждый раз при проектировании нового технологического процесса на конкретную деталь.

Условно-постоянная (справочная и методическая) информация включает в себя характеристики:

- оборудования опытного производства;
- технологической оснастки;
- нормализованных режущих, вспомогательных и измерительных инструментов;
- видов заготовок [5].

Данная информация является достаточно стабильной и постоянно хранится во внешней памяти компьютера при его автономной работе или на сервере при работе в вычислительной сети.

Последовательность использования исследуемого информационного обеспечения на различных стадиях внедрения системы. На начальной стадии внедрения САПР ТП опытных образцов зерноуборочной и кормоуборочной техники применяется метод синтеза, поскольку в архиве технологических процессов отсутствует аналог. Проектирование начинается с ввода информации, размещаемой в заголовочной части технологических документов, на рис. 1 представлена информационная модель их формирования. К такого рода информации относятся следующие сведения: фамилии разработчика, нормировщика и нормоконтролера, проектирующих данный технологический процесс; обозначение и наименование детали, а также сборочной единицы и изделия, в которые эта деталь входит; размер изготавливаемой партии; параметры исходной заготовки. Вместе с тем на этом этапе проектирования вводятся данные, которые в дальнейшем будут использоваться для поиска в электронном архиве процесса-аналога при проектировании технологии на конструктивно подобную деталь:

- число размеров на поле чертежа;
- качество неуказанных предельных отклонений;
- основная шероховатость поверхностей детали.

Для удобства дальнейшего применения и редактирования технологического процесса система должна автоматически сохранять следующую служебную информацию:

- уникальный код технологического процесса;
- дату проведения последнего изменения;
- номер последней редактируемой операции;
- признак блокировки информации о технологическом процессе (при работе с сетевой версией системы).

Анализ информационной модели показывает, что для автоматизации технологического процесса должны быть созданы базы данных фамилий разработчиков, нормировщиков и нормоконтролеров, а также базы материалов и видов заготовок. При этом вес заготовки и коэффициент использования материала (КИМ) автоматически определяются системой.



Рис. 1. Информационная модель формирования заголовочной части технологических документов
Fig. 1. Information model for the formation of the header part of technological documents

В качестве исходных заготовок в машиностроении принимают:

- 1) отливки, полученные литьем в песчаные либо металлические формы или другими способами;
- 2) прокат в виде горячекатаного проката обычной или повышенной точности, а также профильный, фасонный либо другой тип проката;
- 3) поковки, полученные обработкой давлением в подкладных кольцах и штампах;
- 4) штамповки (поковки), полученные путем обработки давлением – объемной горячей или холодной штамповкой;
- 5) сварные заготовки из листового материала.

Выбор вида заготовки представляет собой сложную многовариантную задачу. В общем случае вид заготовки должен быть выбран при конструировании детали, а оптимальный метод ее изготовления уточнен при технологическом проектировании. Способ выполнения этих работ определяется видом детали и основывается на одном из трех вариантов:

1. Если деталь обрабатывается со всех сторон и к ее физико-механическим свойствам и макроструктуре металла не предъявляются особые требования, то выбор заготовки производится только технологом.

2. Если деталь обрабатывается со всех сторон, но к ней предъявляются дополнительные требования, то конструктор, задавая эти требования (например, к расположению волокон металла), косвенно задает возможные методы получения заготовки. Окончательное решение о выборе заготовки принимает технолог.

3. Если деталь имеет необрабатываемые поверхности, то конструктор предварительно задает определенный тип заготовки, указывая штамповочные (литейные) радиусы, уклоны и другие параметры, а технолог определяет оптимальный метод получения заготовки. Окончательное решение о выборе заготовки в этом случае целесообразно принимать совместно (и конструктору, и технологу).

База данных материалов содержит информацию о группах материалов, марках, профилях материалов, стандартных обозначениях материала и его технических параметрах. Структура базы данных материалов показана на рис. 2. Технологические решения, связанные с формированием технологического маршрута при отсутствии в архиве технологического процесс-аналога, разработчик должен принимать директивно. При этом САПР ТП не позволит выбрать операцию, которую нельзя использовать из-за отсутствия соответствующего оборудования [6].

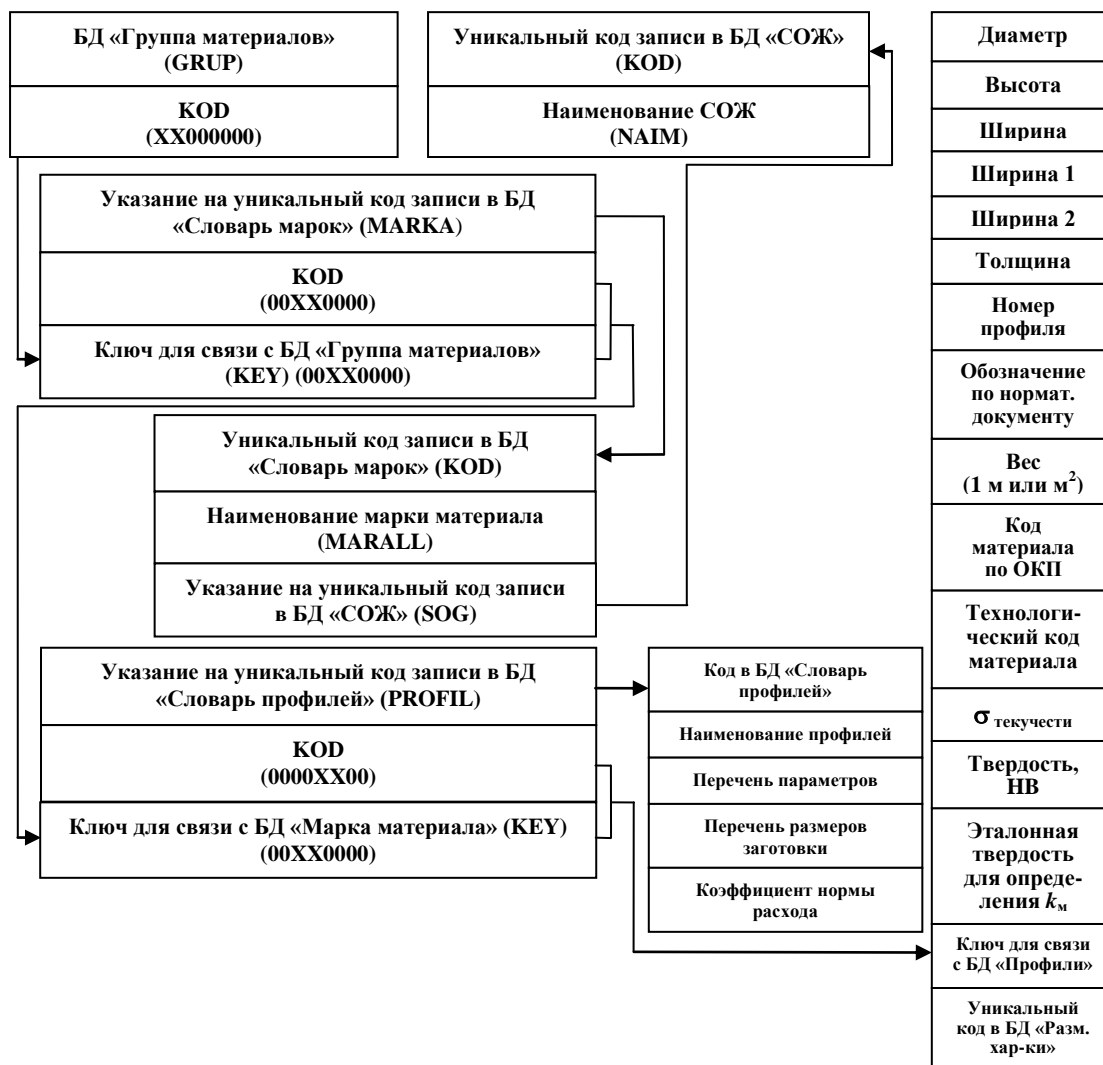


Рис. 2. Информационная модель формирования раздела «Материалы»

Fig. 2. Information model for the formation of the "Materials" section

На рис. 3 изображена информационная модель формирования технологического маршрута [7]. Анализ информационной модели показывает, что для автоматизации этого процесса должны быть созданы базы данных, содержащие информацию об оборудовании и исполнителях.

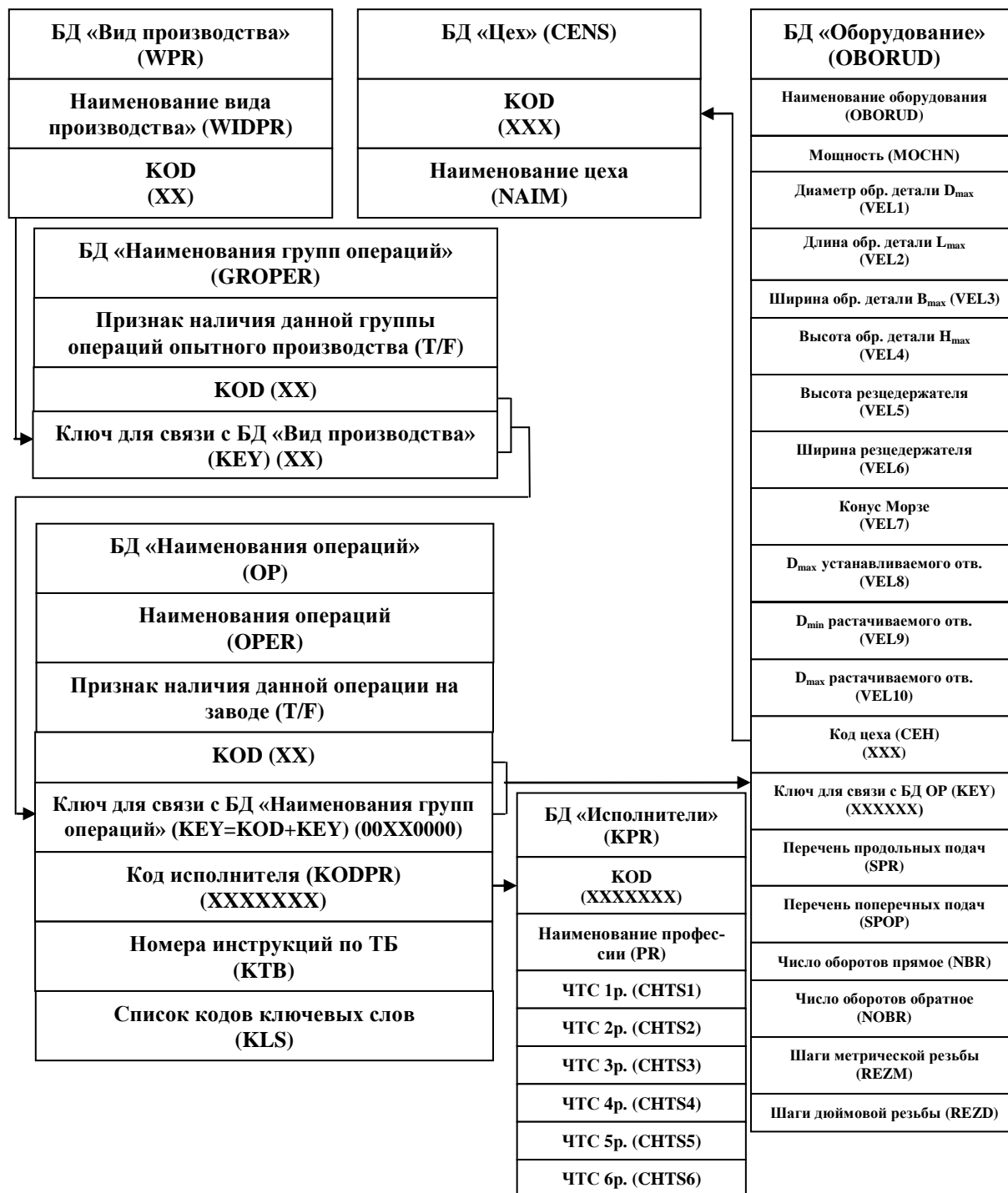


Рис. 3. Информационная модель формирования маршрута обработки
Fig. 3. Information model of processing route formation

Разработка универсальной схемы формирования операционного описания технологического процесса. В основу формирования операционного описания технологического процесса положена универсальная схема, показанная на рис. 4.

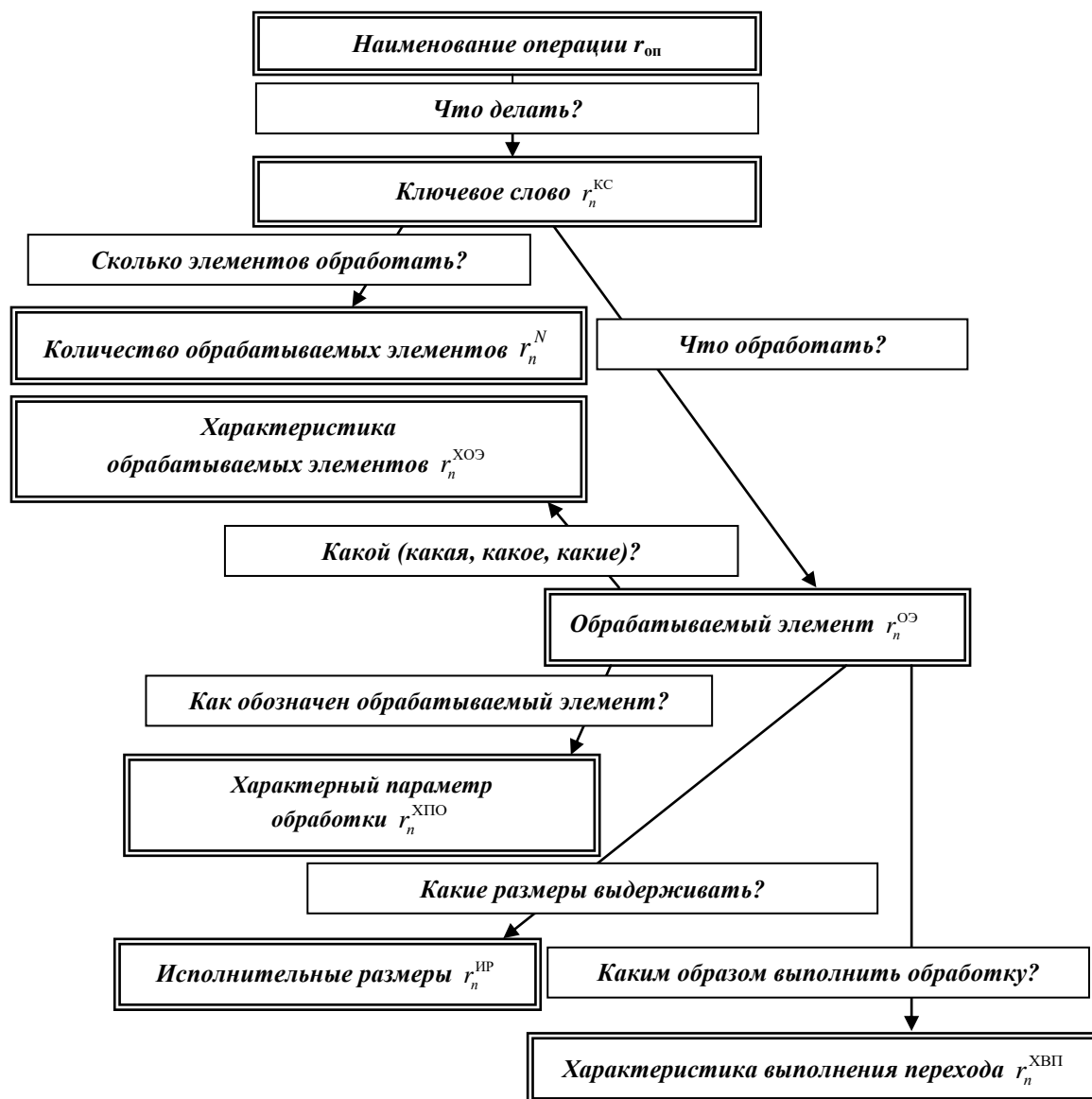


Рис. 4. Универсальная схема формирования технологических переходов

Fig. 4. Universal scheme for the formation of technological transitions

Для практической реализации универсальной схемы формирования технологических переходов в САПР ТП были созданы и взаимосвязаны следующие БД:

- «Установочные переходы» (FIRST);
- «Связь операций и ключевых слов» (SWOPKLS);
- «Ключевые слова» (KLS);
- «Предметы производства» (PREDPR);
- «Характеристики предмета производства» (DOP4_2);
- «Характеристики выполнения переходов» (DOP4_4);
- «Допуски» (POR500);
- «Типовые переходы» (PFIPKW);
- «Архив переходов» (ARCPER).

Информационная модель формирования операционного описания технологического процесса показана на рис. 5.

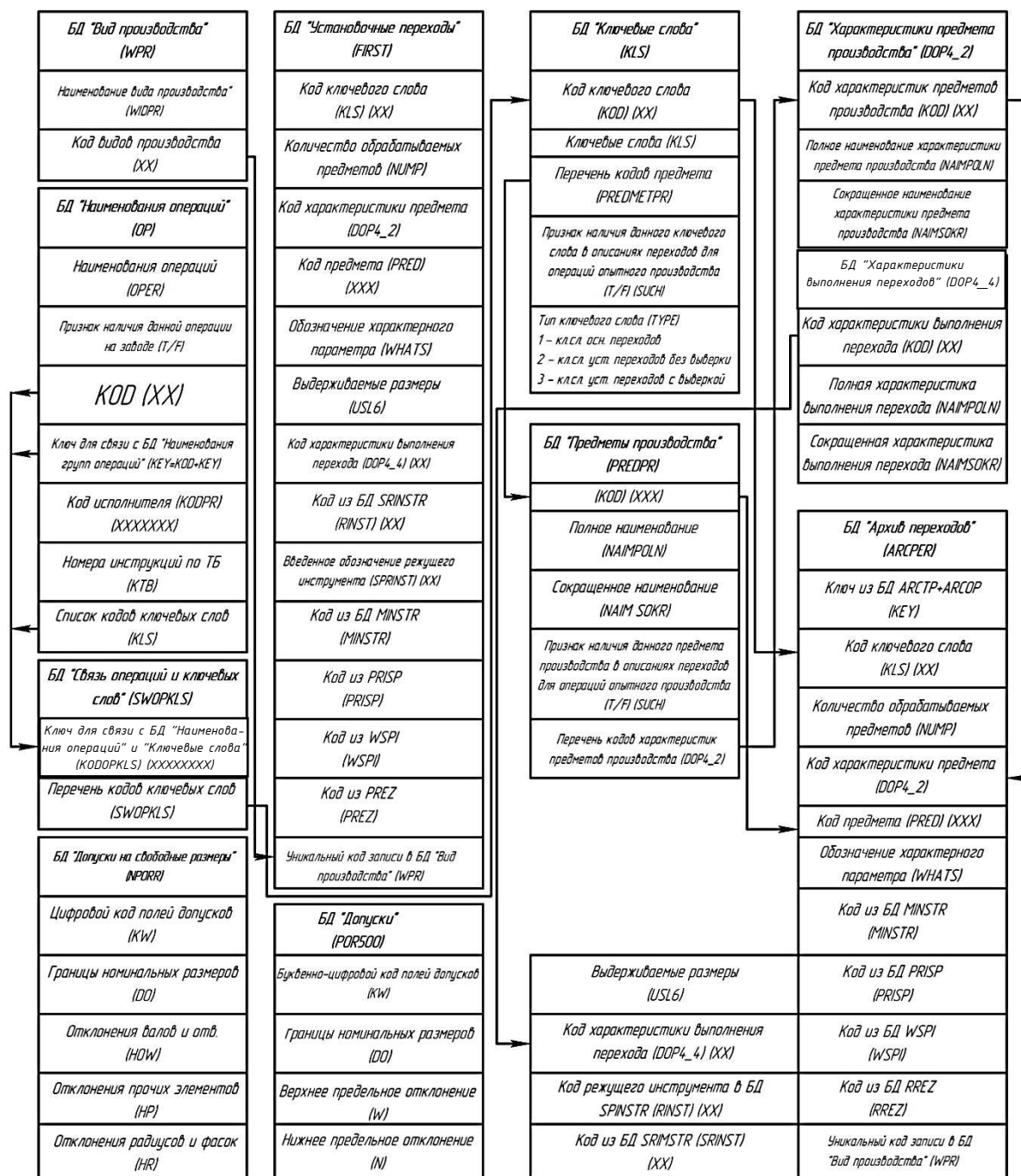


Рис. 5. Информационная модель формирования операционного описания технологического процесса

Fig. 5. Information model for the formation of the operational description of the technological process

Порядок использования исследуемого информационного обеспечения при формировании операционного описания технологического процесса. Формирование операционного описания технологического процесса начинается с установочного перехода. Выбор начального установочного перехода из БД «Установочные переходы» осуществляется по уникальному коду записи в БД «Вид производства». При этом автоматически определяется ряд параметров описания перехода, в частности код ключевого слова, код предмета и т. д.

При формировании основного перехода на базе универсальной схемы формирования технологических переходов [8] выполняется алгоритм последовательного выбора ряда параметров:

Шаг 1. Выбирается ключевое слово перехода r_n^{KC} .

Шаг 2. Вводится количество обрабатываемых при выполнении перехода элементов r_n^N .

Шаг 3. Выбирается обрабатываемый при выполнении перехода элемент r_n^{OO} .

Шаг 4. Выбирается характеристика обрабатываемого при выполнении перехода элемента r_n^{XOO} .

Шаг 5. Вводится характерный параметр обработки r_n^{XPO} .

Шаг 6. Вводятся исполнительные размеры r_n^{IP} . При вводе указанных параметров к ним автоматически добавляются верхние и нижние отклонения.

Шаг 7. Выбирается характеристика выполнения перехода $r_n^{XBП}$.

Итак, текст перехода сформирован окончательно, система переходит к выбору средств технологической оснастки.

Формирование описания средств технологической оснастки с использованием исследуемого информационного обеспечения. Вначале выбирается режущий инструмент. Для этого используется БД «Выбор режущих инструментов» (RINSTR), структура которой приведена в табл. 1.

Таблица 1
Структура полей БД «Выбор режущих инструментов» (RINSTR)

Table 1
The structure of the fields of the database "Choice of cutting tools" (RINSTR)

Наименование поля <i>Field name</i>	Вводимая информация <i>Input information</i>
KOD	Ключ из БД OP+ OBORUD+LS+REDPR+ DOP4_2
KODS	Список кодов режущих инструментов
NRR	Номер схемы для расчета режимов резания
NRRP	Номер подсхемы для расчета режимов резания

Выбор режущих инструментов выполняется с помощью указанной БД при соблюдении ряда условий, определенных в Справочнике режущих инструментов, структура записей которого приведена в табл. 2.

После выбора режущего инструмента система автоматически переходит в режим выбора вспомогательного инструмента. Автоматизированный выбор этого вида технологической оснастки обеспечивается взаимодействием с БД «Справочник вспомогательных инструментов» (SWSP0, табл. 3) и БД «Выбор вспомогательных инструментов» (WSP0, табл. 4).

Приспособления выбираются только для установочных переходов. За каждой моделью оборудования закреплен определенный комплект универсальных приспособлений. Кроме того, технолог имеет возможность при необходимости внести обозначения специального приспособления.

После выбора вспомогательных инструментов система переходит к выбору измерительных инструментов. Эта технологическая задача по аналогии с предыдущими решается при помощи БД «Измерительные инструменты» (SPRISP), структура которой приведена в табл. 5.

Таблица 2
Структура полей БД «Справочник режущих инструментов» (SRINSTR)

Table 2
The structure of the fields of the database "Directory of cutting tools" (SRINSTR)

Наименование поля <i>Field name</i>	Вводимая информация <i>Input information</i>
KOD	Код
VEL5	Ширина державки резца
VEL6	Высота державки резца
VEL7	Для инструментов с хвостовиком – номер конуса Морзе, для резьбовых резцов – угол в плане
VEL8	Диаметр инструмента
B	Ширина режущей кромки отрезного резца
UB	Условие выбора отрезного (канавочного) резца (ширина канавки B равна ширине режущей части резца)
L	Длина державки резца
UL	Условие выбора резца (длина резцедержателя должна быть больше длины резца)
M	Модуль нарезаемой инструментом зубчатой поверхности
UM	Условие выбора зубообрабатывающего инструмента (модуль зубчатой поверхности, указанный в тексте перехода как M, равен модулю нарезаемой инструментом зубчатой поверхности)
R	Радиус, обрабатываемой инструментом поверхности
UR	Условие выбора инструмента для обработки радиуса (радиус обрабатываемой поверхности, указанный в тексте перехода как R, равен радиусу, указанному в технической характеристике инструмента)
T	Шаг зубчатой звездочки, обрабатываемой инструментом
UT	Условие выбора инструмента для обработки звездочки (шаг звездочки, указанный в тексте перехода T, равен шагу, указанному в технической характеристике инструмента)
N	Количество шлицев поверхности, обрабатываемой инструментом
DM	Малый диаметр шлицев, обрабатываемых инструментом
UDM	Условие выбора инструмента для шлицевой поверхности (малый диаметр шлицев, указанный в тексте перехода как DM, равен DM, указанному в технической характеристике инструмента)
DB	Большой диаметр шлицев, обрабатываемых инструментом
UDB	Условие выбора инструмента для шлицевой поверхности (большой диаметр шлицев, указанный в тексте перехода как DB, равен DB, указанному в технической характеристике инструмента)
DOPUSK	Допуск на поверхность, обрабатываемую инструментом
UDOPUSK	Условие выбора инструмента по допуску на обрабатываемую поверхность (допуск на обрабатываемую поверхность, указанный в тексте перехода, равен допуску, указанному в технической характеристике инструмента)
USHLIC	Условие выбора инструментов (фрез и протяжек) для обработки шлицевых поверхностей (наружных и внутренних): параметры обрабатываемой поверхности, указанные в тексте перехода, равны параметрам, указанным в технической характеристике инструмента
UGOL	Угол, под которым расположены обрабатываемые поверхности
UUGOL	Условие выбора инструмента по углу, под которым находятся обрабатываемые поверхности: угол, под которым находятся обрабатываемые поверхности, указанный в тексте перехода, равен углу, указанному в технической характеристике инструмента
KONUS	Конусность обрабатываемых инструментом поверхностей
UKONUS	Условие выбора инструмента по конусности обрабатываемых поверхностей: конусность, указанная в тексте перехода, равна конусности, указанной в технической характеристике инструмента
KONMOR	Конус Морзе, обрабатываемый инструментом
UKONMOR	Условие выбора инструмента по конусу Морзе обрабатываемой поверхности: конус Морзе, указанный в тексте перехода, равен конусу Морзе, указанному в технической характеристике инструмента
REZ	Резьба, обрабатываемая инструментом
UREZ	Условие выбора инструмента по резьбе обрабатываемой поверхности: резьба, указанная в тексте перехода, равна резьбе, указанной в технической характеристике инструмента
DOTW	Диаметр посадочного отверстия инструмента (для выбора оправки)
LOTW	Длина посадочного отверстия инструмента (для выбора оправки)
H1	Высота державки резца для расточных резцов, вставляемых в борштангу
B1	Ширина державки резца для расточных резцов, вставляемых в борштангу
OBOZN	Полное обозначение инструмента по нормативному документу
MATERIAL	Материал режущей части инструмента
TYPE	Тип инструмента (для выбора вспомогательных инструментов)

Таблица 3
Структура полей БД «Справочник вспомогательных инструментов» (SWSP0)

Table 3
The structure of the fields of the database "Directory of auxiliary tools" (SWSP0)

Наименование поля <i>Field name</i>	Вводимая информация <i>Input information</i>
KOD	Код
OBOZN	Обозначение вспомогательных инструментов согласно нормативным документам

Таблица 4
Структура полей БД «Выбор вспомогательных инструментов» (WSPO)

Table 4
The structure of the fields of the database "Selection of auxiliary tools" (WSPO)

Наименование поля <i>Field name</i>	Вводимая информация <i>Input information</i>
TYPEI	Тип режущего инструмента
KMI	Конус Морзе режущего инструмента
KMS	Конус Морзе металлорежущего станка
DOTW	Диаметр оправки под посадку насадного режущего инструмента
LOTW	Длина посадочной части оправки для насадного режущего инструмента
KODS	Коды в БД «Справочник вспомогательных инструментов»

Таблица 5
Структура полей БД «Измерительные инструменты» (SPRISP)

Table 5
The structure of the fields of the database "Measuring tools" (SPRISP)

Наименование поля <i>Field name</i>	Вводимая информация <i>Input information</i>
KOD	Код мерительного инструмента
OBOZN	Обозначение мерительного инструмента
TYPE	Тип мерительного инструмента
MINR	Нижняя граница диапазона измерений
MAXR	Верхняя граница диапазона измерений
TCH	Точность измерения
R	Измеряемый радиус
REZ	Измеряемая резьба
MORZE	Измеряемый конус Морзе
D	Измеряемый диаметр
DOPUSK	Контролируемый допуск
SHLIC	Контролируемые шлицы
MMIN	Минимальный контролируемый шаг
MMAH	Максимальный контролируемый шаг

Разработка структуры архива технологических процессов. После окончания проектирования описание технологического процесса заносится в архив, структура которого показана на рис. 6. Фактически в архиве хранятся результаты принятых проектировщиком технологических решений. Таким образом, при изменении нормативной базы, например при изменении маркировки инструментов, автоматически происходит изменение технологических процессов, хранящихся в архиве, так как меняется только обозначение принятого технологом решения.

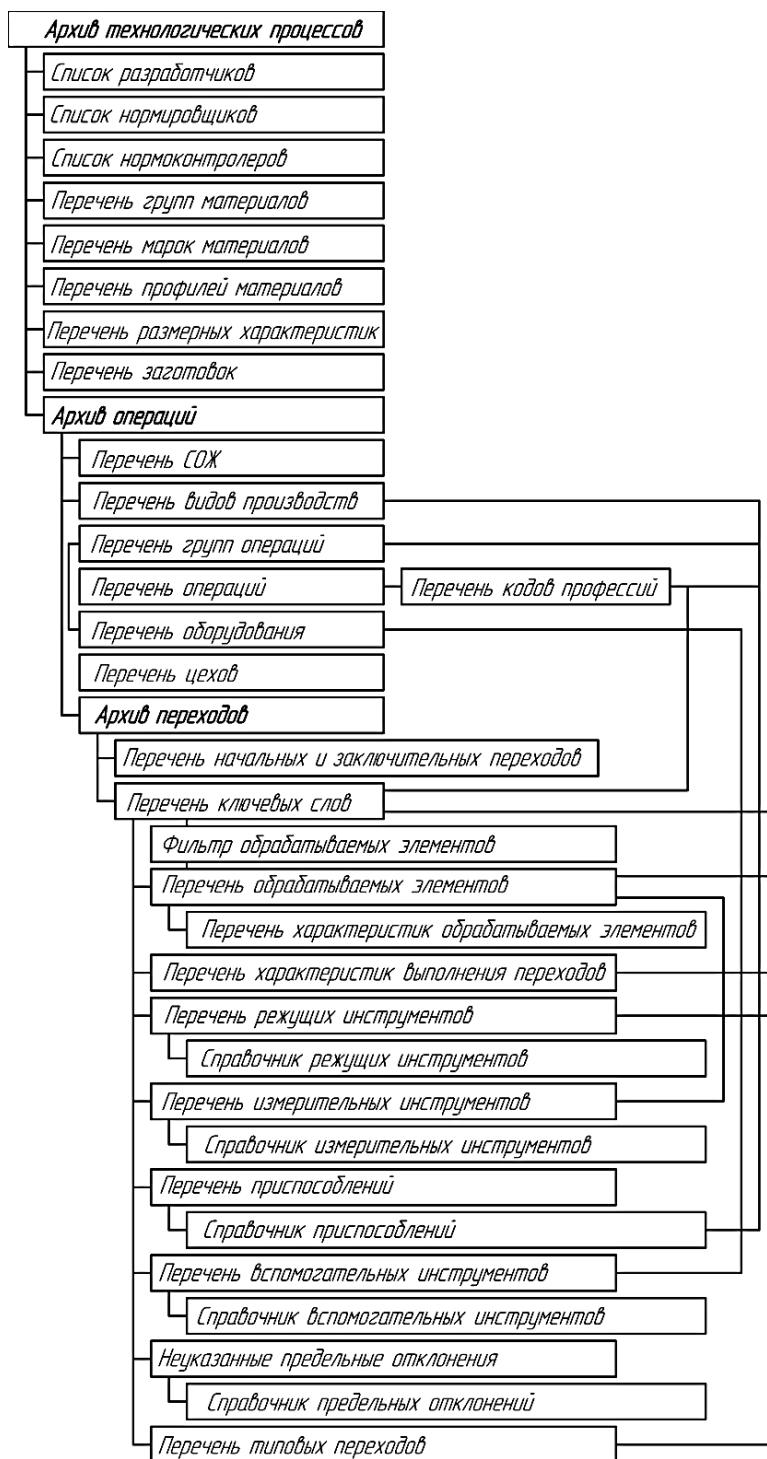


Рис. 6. Структура архива технологических процессов
Fig. 6. The structure of the archive of technological processes

Аналогичная ситуация возникает и при изменении парка оборудования. Введение нового оборудования просто расширяет возможности проектирования. Замена устаревшей модели на более современную решается корректировкой справочника. В случае вывода какого-либо оборудования из использования (и архива) можно получить справку о технологических процессах, в которых оно задействовано, и с помощью активации провести необходимые изменения.

Заключение. В результате исследования информационного обеспечения САПР ТП опытных образцов структурно-сложных изделий была решена задача определения структуры и последовательности использования информационного обеспечения на различных стадиях внедрения системы, разработаны универсальная схема формирования операционного описания технологического процесса и структура архива технологических процессов.

Внедрение САПР ТП опытных образцов кормоуборочной и зерноуборочной техники обеспечило повышение производительности проектирования в 7,8 раза.

Список использованных источников

1. Петухов, А. В. Автоматизация проектирования технологических процессов изготовления опытных образцов / А. В. Петухов // Тракторы и сельскохозяйственные машины. – 1993. – № 12. – С. 33–35.
2. Акулович, Л. М. Основы автоматизированного проектирования технологических процессов в машиностроении : учеб. пособие / Л. М. Акулович, В. К. Шелег. – Минск : Новое знание; М. : ИНФРА-М, 2012. – 488 с.
3. Автоматизация и управление в технологических комплексах / А. М. Русецкий [и др.] ; под общ. ред. А. М. Русецкого. – Минск : Беларус. навука, 2014. – 375 с.
4. Петухов, А. В. Системы автоматизированного проектирования технологических процессов : пособие по одноименному курсу для студентов специальности 36.01.01 «Технология машиностроения». – Гомель : ГГТУ им. П. О. Сухого, 2005. – 84 с.
5. Епифанова, О. В. Автоматизация технологической подготовки производства деталей, изготавливаемых на оборудовании с ЧПУ / О. В. Епифанова, Д. И. Троицкий // Изв. ТулГУ. Техн. науки. – Вып. 3. – Тула : Изд-во ТулГУ, 2011. – С. 53–61.
6. Епифанова, О. В. Методика оптимального выбора структуры САМ-системы / О. В. Епифанова, Д. И. Троицкий // Изв. ТулГУ. Техн. науки. – Вып. 3. – Тула : Изд-во ТулГУ, 2011. – С. 21–29.
7. Петухов, А. В. Системы автоматизированного проектирования технологических процессов : учеб. пособие / А. В. Петухов, Д. В. Мельников, В. М. Быстренков ; М-во образования Респ. Беларусь, Гомел. гос. техн. ун-т им. П. О. Сухого. – Гомель, 2011. – 143 с.
8. Петухов, А. В. Универсальная схема формирования переходов // Технологическая системотехника : сб. тр. Первой Междунар. электр. науч.-техн. конф. – Тула : Гриф и К, 2002. – С. 294–296.

References

1. Petukhov A. V. *Automation of the design of technological processes for the production of prototypes*. Traktory i sel'skhozajstvvennye mashiny [Tractors and Agricultural Machines], 1993, no. 12, pp. 33–35 (In Russ.).
2. Akulovich L. M., Sheleg V. K. *Osnovy avtomatizirovannogo proektirovaniya tekhnologicheskikh processov v mashinostroenii. Fundamentals of Computer-Aided Design of Technological Processes in Mechanical Engineering*. Minsk, Novoe znanie; Moscow, INFRA-M, 2012, 488 p. (In Russ.).
3. Ruseckij A. M., Vityaz' P. A., Hejfec M. L., Svirskij D. N., Averchenkov A. V., ..., Sheleg V. K. *Avtomatizaciya i upravlenie v tekhnologicheskikh kompleksah. Automation and Control in Technological Complexes*. In A. M. Ruseckij (ed.). Minsk, Belaruskaja navuka, 2014, 375 p. (In Russ.).
4. Petukhov A. V. *Sistemy avtomatizirovannogo proektirovaniya tekhnologicheskikh processov : posobie po odnoimennomu kursu dlja studentov special'nosti 36.01.01 "Tehnologija mashinostroenija". Systems of Computer-aided Design of Technological Processes : A Manual on the Course of the Same Name for Students of the Specialty 36.01.01 "Technology of Mechanical Engineering"*. Gomel, Sukhoi State Technical University of Gomel, 2005, 84 p. (In Russ.).
5. Epifanova O. V., Troickij D. I. *Automation of technological preparation for the production of parts manufactured on CNC equipment*. Izvestija Tul'skogo gosudarstvennogo universiteta. Tehnicheskie nauki [News

of the Tula State University. Technical Science], iss. 3, Tula, Izdatel'stvo Tul'skogo gosudarstvennogo universiteta, 2011, pp. 53–61 (In Russ.).

6. Epifanova O. V., Troickij D. I. Methodology for the optimal selection of the structure of the CAM system. *Izvestija Tul'skogo gosudarstvennogo universiteta. Tehnicheskie nauki [News of the Tula State University. Technical Science]*, iss. 3, Tula, Izdatel'stvo Tul'skogo gosudarstvennogo universiteta, 2011, pp. 21–29 (In Russ.).

7. Petukhov A. V., Melnikov D. V., Bystrenkov V. M. Sistemy avtomatizirovannogo proektirovanija tehnologicheskikh processov. *Systems of Computer-aided Design of Technological Processes*. Ministerstvo obrazovanija Respubliki Belarus', Sukhoi State Technical University of Gomel. Gomel', 2011, 143 p. (In Russ.).

8. Petukhov A. V. *Universal transition generation scheme*. Tehnologicheskaja sistemotehnika : sbornik trudov Pervoj Mezhdunarodnoj jelektronnoj nauchno-tehnicheskoi konferencii [*Technological System Engineering : Proceedings of the First International Electronic Scientific and Technical Conference*], Tula, Grif i K, 2002, pp. 294–296 (In Russ.).

Информация об авторе

Петухов Александр Владимирович, старший преподаватель кафедры «Технология машиностроения», Гомельский государственный технический университет имени П. О. Сухого.
E-mail: Petukhov_2000@gstu.by

Information about the author

Alexander V. Petukhov, Senior Lecturer of the Chair "Technology of Mechanical Engineering", Sukhoi State Technical University of Gomel.
E-mail: Petukhov_2000@gstu.by

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

INFORMATION TECHNOLOGIES



УДК 004.75
<https://doi.org/10.37661/1816-0301-2022-19-4-84-93>

Оригинальная статья
Original Paper

Модель балансировки нагрузки кластерной системы с учетом аппаратных характеристик серверного оборудования

А. Н. Марков

*Центр информатизации и инновационных разработок
Белорусского государственного университета
информатики и радиоэлектроники,
ул. Платонова, 39, Минск, 220013, Беларусь
E-mail: a.n.markov@bsuir.by*

Аннотация

Цели. Поставлены цели модернизировать и дополнить существующую модель балансировки нагрузки в многосерверных системах с учетом аппаратных характеристик серверного оборудования, а также его наиболее загруженных компонентов в кластере сервиса видео-конференц-связи при использовании в образовательных процессах и организации дистанционной формы образования.

Методы. Существующая математическая модель балансировки нагрузки рассмотрена как система массового обслуживания, в которой вводятся штрафы за простой оборудования, а штрафы за ожидание в очереди будут зависеть от загруженности аппаратных компонентов серверов в кластерной архитектуре сервиса видео-конференц-связи.

Результаты. Приведены формулы для вычисления суммарной производительности кластера из n серверов при максимальной и минимальной загрузках аппаратных компонентов серверов в кластере системы видео-конференц-связи.

Заключение. Разработан моделирующий комплекс для проверки математической модели на системе до $n < 10$ серверов в кластере системы видео-конференц-связи. По результатам вычислений моделирующего комплекса сделан вывод о необходимости доработки существующего алгоритма балансировки нагрузки на выбранный сервис видео-конференц-связи BigBlueButton.

Ключевые слова: видео-конференц-связь, сервис BigBlueButton, балансировка нагрузки в кластере, математическая модель балансировки, дистанционное образование

Для цитирования. Марков, А. Н. Модель балансировки нагрузки кластерной системы с учетом аппаратных характеристик серверного оборудования / А. Н. Марков // Информатика. – 2022. – Т. 19, № 4. – С. 84–93. <https://doi.org/10.37661/1816-0301-2022-19-4-84-93>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 13.07.2022

Подписана в печать | Accepted 26.07.2022

Опубликована | Published 29.09.2022

Cluster system load balancing model with consideration of hardware characteristics of server hardware

Aleksey N. Markov

*Center for Informatization and Innovation Development
of the Belarusian State University
of Informatics and Radioelectronics,
st. Platonova, 39, Minsk, 220013, Belarus
E-mail: a.n.markov@bsuir.by*

Abstract

Objectives. To upgrade and complement the existing load balancing model in multi-server systems, taking into account the hardware characteristics of the server equipment, as well as the most loaded components of the server equipment in the video conferencing service cluster in educational processes and distance education.

Methods. The existing mathematical model of load balancing as a mass exchange system is considered, when significant changes are introduced: penalties for equipment downtime and penalties for waiting in a queue will depend on the load on the server hardware components in the cluster architecture of video conferencing service.

Results. Formulas are given for calculating the total performance of a cluster of n servers with the maximum and minimum load of server hardware components in a videoconferencing system cluster.

Conclusion. A modeling complex has been developed to test the mathematical model on a system of up to $n < 10$ servers in a cluster of a videoconferencing system. Based on the results of calculations of the modeling complex, it was concluded that it is necessary to upgrade the existing algorithm for balancing the load on the selected BigBlueButton video conferencing service.

Keywords: video conferencing, service BigBlueButton, load balancing in a cluster, mathematical model of load balancing, distance education

For citation. Markov A. N. *Cluster system load balancing model with consideration of hardware characteristics of server hardware*. *Informatika [Informatics]*, 2022, vol. 19, no. 4, pp. 84–93 (In Russ.).
<https://doi.org/10.37661/1816-0301-2022-19-4-84-93>

Conflict of interest. The author declares of no conflict of interest.

Введение. В ходе реализации проекта цифровой трансформации образования на этапе перехода к смешанной модели обучения в рамках системы электронного обучения в БГУИР был внедрен сервис видео-конференц-связи BigBlueButton [1, 2]. В процессе интеграции видеосервиса в учебный процесс выявлена проблема повышенной нагрузки на единственный сервер, которая описана в работе [3]. На этапе тестовой эксплуатации эмпирическим путем было установлено, что увеличение аппаратных ресурсов сервера BigBlueButton относительно первоначально заданных параметров не приносит значительного прироста производительности сервиса в целом. Одной из главных причин нестабильности подключений и ненадлежащего качества видеопотока является превышение допустимого числа подключений к сервису. Для выхода из существующей проблемы был развернут сервис балансировки нагрузки системы видео-конференц-связи, решающий проблему распределения пользователей в зависимости от доступности каждого сервера и устранения задержек на этапах подключения к видеоконференции. Стандартный алгоритм балансировки нагрузки не учитывал параметры загрузки серверного оборудования. Встроенный метод балансировки нагрузки на основе алгоритма Round Robin DNS [4] учитывает только доступность ресурса без учета нагрузки на аппаратные компоненты серверов в кластере. Для устранения данного недостатка системы можно взять за основу математическую модель балансировки нагрузки в адаптивном алгоритме управления распределением нагрузки в многосерверных системах [5] и доработать ее с учетом загрузки аппаратных компонентов серверного оборудования в кластере. Другие математические модели для балансировки нагрузки (грид-вычисления) нельзя принять за основу, так как принципиальный под-

ход к балансировке нагрузки в них осуществляется на аппаратном уровне самого сервера (процессора и распределения потоков в операционной системе) без учета количества серверов в кластере.

Критерии качества работы многосерверной системы. Критериями качества многосерверных систем являются затраты, связанные с простоем серверов в кластере и временем ожидания подключения к конференции при наличии очереди на подключение. Оптимизация работы системы заключалась в использовании этих критериев для вычисления таких параметров работы серверов с учетом выбранного оборудования, при которых затраты будут минимальными. В процессе оптимизации подбираются доли распределения входного потока между серверами в кластере α_i и время доступа к ресурсу b_i (выраженное как время обработки одного запроса на вход в конференцию) с учетом аппаратных компонентов серверов при заданном количестве серверов n в кластере, где $i = \overline{1, n}$ – номер сервера в кластере.

Для определения минимальных значений критериев качества кластер выделен как система массового обслуживания типа $M | D | 1$. На основании системы $M | D | 1$ суммарный критерий качества работы системы (суммарные затраты на ожидание в очереди и суммарные затраты на простой сервера в кластере), выраженный через метод множителей Лагранжа, будет иметь следующий вид [5]:

$$\min \left\{ S(n, \lambda_1, \lambda_2 \dots \lambda_n, b_1, b_2 \dots b_n) = \sum_{i=1}^n \left\{ \frac{Pw_i \lambda_i^2 b_i^2}{2(1 - \lambda_i b_i)} + Pd_i \cdot (1 - \lambda_i b_i) \right\} \right\},$$

$$\sum_{i=1}^k \lambda_i = \Lambda, \lambda_i \geq 0,$$
(1)

где $i = \overline{1, n}$ – номер сервера в кластере;

n – количество серверов в вычислительном кластере;

λ_i – интенсивность потока запросов на сервер i ;

Λ – суммарная интенсивность потока запросов на кластер видео-конференц-связи, определяющая, на сколько пользователей какое количество запросов на вход в видеоконференцию распределено;

b_i – время обработки одного запроса на вход в конференцию, с;

Pw_i – штраф за ожидание в очереди на подключение к конференции;

Pd_i – штраф за простой сервера в кластере.

Сумма потоков заявок на все серверы равна общему потоку запросов на вход в конференцию, а значение потока запросов на каждый сервер – величина неотрицательная.

Отыскание точки минимума функции позволит определить параметры работы кластера с учетом штрафа за ожидание в очереди и штрафа за простой сервера исходя из выделенных аппаратных ресурсов серверов.

Зависимость величины штрафа за ожидание в очереди от времени доступа к серверу должна быть обратной: чем больше время обработки запроса, учитывающее загруженность сервера, тем «дешевле» должен быть сервер и меньше штраф за ожидание в очереди такого сервера.

Предположим, что при условии организации кластера видео-конференц-связи штраф за ожидание в очереди можно определить выражением

$$Pw_i = \frac{Pr_i}{b_i},$$
(2)

где Pr_i – суммарная загрузка сервера i , выраженная как произведение производительностей всех аппаратных компонентов:

$$Pr_i = (P_{cpu\max})_i \cdot (P_{ram\max})_i \cdot (P_{netw\max})_i \cdot (P_{hdd\max})_i, \quad i = \overline{1, n}. \quad (3)$$

Здесь $(P_{cpu\max})_i$ – производительность процессора в сервере i ;

$(P_{ram\max})_i$ – производительность памяти сервера i ;

$(P_{netw\max})_i$ – производительность локальной сети сервера в кластере;

$(P_{hdd\max})_i$ – производительность дискового хранилища сервера.

Учитывая специфику работы серверов и выделенные аппаратные ресурсы, параметрами, не оказывающими влияние на итоговую загруженность системы (при условии использования видео-конференц-связи на выбранном кластере), можно пренебречь согласно анализу графиков нагрузки на аппаратные компоненты [4]:

$$P_{ram\max} = 1, \quad P_{netw\max} = 1, \quad P_{hdd\max} = 1. \quad (4)$$

Между тем предположим, что кластер построен на различных выделенных серверах. Тогда его итоговая загруженность будет тем выше, чем больше загружены оперативная память, сеть и дисковое хранилище.

Для выделенного кластера с учетом аппаратных ресурсов общая загруженность серверов будет зависеть преимущественно от производительности центрального процессора:

$$Pr_i = (P_{cpu\max})_i, \quad i = \overline{1, n}. \quad (5)$$

Основываясь на методе расчета и анализа теоретической пиковой производительности процессоров [6–9] с изменениями ввиду специфики работы сервиса, предположим, что производительность процессора в серверах рассчитывается следующим образом:

$$(P_{cpu\max})_i = F_{cpu\max i} \cdot Fl_i \cdot v_i \cdot H_{осн i}, \quad (6)$$

где $F_{cpu\max i}$ – турбочастота процессора (для некоторых процессоров это максимальная частота работы). С учетом систем виртуализации выделенные аппаратные ресурсы будут напрямую зависеть от процессоров в сервере i ;

Fl_i – количество операций с плавающей запятой в секунду, выполняемых данной вычислительной системой, на одно ядро процессора, флопс;

v_i – количество ядер в выделенном процессоре сервера i ;

$H_{осн i}$ – усредненный коэффициент загрузки процессора сервера i на выполнение основных операций 0,95, показывающий долю от производительности процессора на выполнение задач распределения потока. Оставшаяся доля 0,05 – усредненный коэффициент, определяющий выполнение фоновых задач для сервера i на базе виртуализированной Unix-системы (коэффициенты 0,90 и 0,10 – для сервера на базе системы Windows).

Параметры, влияющие на общую производительность, а именно: величина кеш-памяти, объем физической адресуемой памяти, разрядность внешних шин данных и адреса, разрядность регистров, не учитываются, поскольку напрямую не влияют на показатели производительности процессора при обработке потока запросов в единицу времени.

Простой сервера в кластере (с учетом специфики работы видео-конференц-связи) является естественным процессом в отличие от простоя веб-серверов или серверов баз данных. Пользо-

вателей конференции может быть минимальное количество, и тогда все серверы в кластере не используются. Чем меньше время обслуживания запроса от клиента на доступ к конференции, тем большую итоговую производительность обеспечивает сервер, тем «дороже» время работы такого сервера и тем больше должен быть штраф за его простой. Предположим, что штраф за простой сервера будет рассчитываться исходя из минимальных значений производительности сервера:

$$Pd_i = \frac{Pr(\min)_i}{b_i}, \quad (7)$$

где $Pr(\min)_i$ – минимальная производительность сервера i в кластере, выраженная как произведение минимальной производительности всех аппаратных компонентов серверов:

$$Pr(\min)_i = (P_{cpu\ min})_i \cdot (P_{ram\ min})_i \cdot (P_{netw\ min})_i \cdot (P_{hdd\ min})_i, \quad i = \overline{1, n}. \quad (8)$$

Исходя из анализа графиков работы серверов в кластере и учитывая выделенные аппаратные ресурсы, можно сделать вывод, что параметрами, не оказывающими влияние на производительность системы, можно пренебречь [4]:

$$P_{ram\ min} = 1, \quad P_{netw\ min} = 1. \quad (9)$$

Между тем минимальная производительность системы хранения данных (или жесткого диска в выделенном сервере) будет оказывать влияние на работу сервера в целом. Поскольку $P_{hdd\ min} = 2$, простой дискового массива более критичен по сравнению с рабочим режимом массива в целом, что обусловлено временем доступа (обращения) к данным на диске в массиве.

Если кластер построен на различных выделенных серверах, итоговая минимальная величина учета его производительности будет тем ниже, чем менее производительными будут оперативная память, сеть и хранилище.

Для выделенного кластера с учетом ресурсов общая минимальная производительность сервера i может определяться выражением

$$Pd_i = (P_{cpu\ min})_i (P_{hdd\ min})_i, \quad i = \overline{1, n}. \quad (10)$$

Предположим, что аналогично расчету пиковой производительности минимальную производительность процессора для сервера i можно найти по формуле

$$Pr(\min)_i = (P_{cpu\ min})_i = F_{cpu\ mini} \cdot Fl_i \cdot v_i \cdot H_{фонов_i} \cdot m, \quad (11)$$

где $m = 3,5$ – поправочный коэффициент весомости простоя самого низкопроизводительного сервера. Физический смысл коэффициента заключается в том, что даже самый низкопроизводительный сервер в кластере должен быть минимально загружен для повышения общей производительности, нежели не быть использован вовсе. Коэффициент подбирается экспериментальным путем с учетом выделенных аппаратных ресурсов серверов в кластере. В случае построения кластера на различных процессорах в серверах коэффициент подбирается исходя из экспериментальных данных по загрузке и простоя каждого сервера в отдельности;

$F_{cpu\ mini}$ – базовая частота процессора сервера i (для некоторых процессоров это минимальная частота работы);

$H_{фонов_i}$ – усредненный коэффициент учета загрузки процессора сервера i на выполнение фоновых операций 0,2, показывающий долю от производительности одного процессора на выпол-

нение задач распределения потока. Оставшаяся доля 0,8 – усредненный коэффициент, учитывающий основные задачи для сервера на базе системы Unix (коэффициенты 0,3 и 0,7 – для сервера на базе системы Windows).

Задачи распределения входного потока между серверами и выбора производительности серверов. Оптимизация системы серверов, связанная одновременно с распределением входного потока и выбором производительности серверов, является задачей математического программирования с ограничениями, которую аналитически или численно можно решить, используя метод множителей Лагранжа. Для нахождения оптимальных значений α_i и b_i методом множителей Лагранжа необходимо решить систему уравнений, где частные производные по всем переменным равны нулю. С учетом подстановок и выражений функций получим систему уравнений

$$\left\{ \begin{array}{l} -\frac{Pw_1\lambda_1\Lambda b_1^2(\lambda_1 b_1 - 2)}{2(\lambda_1 b_1 - 1)^2} - Pd_1\Lambda b_1 + x = 0; \\ -\frac{Pw_2\lambda_2\Lambda b_2^2(\lambda_2 b_2 - 2)}{2(\lambda_2 b_2 - 1)^2} - Pd_2\Lambda b_2 + x = 0; \\ \dots\dots\dots \\ -\frac{Pw_n\lambda_n\Lambda b_n^2(\lambda_n b_n - 2)}{2(\lambda_n b_n - 1)^2} - Pd_n\Lambda b_n + x = 0; \\ \sum_{i=1}^k \alpha_i - 1 = 0; \\ -\frac{Pw_1\alpha_1^2\Lambda^2 b_1(\alpha_1\Lambda b_1 - 2)}{2(\alpha_1\Lambda b_1 - 1)^2} + Pd_1 \frac{(1 + \alpha_1\Lambda)}{b_1} = 0; \\ -\frac{Pw_2\alpha_2^2\Lambda^2 b_2(\alpha_2\Lambda b_2 - 2)}{2(\alpha_2\Lambda b_2 - 1)^2} + Pd_2 \frac{(1 + \alpha_2\Lambda)}{b_2} = 0; \\ \dots\dots \\ -\frac{Pw_n\alpha_n^2\Lambda^2 b_n(\alpha_n\Lambda b_n - 2)}{2(\alpha_n\Lambda b_n - 1)^2} + Pd_n \frac{(1 + \alpha_n\Lambda)}{b_n} = 0. \end{array} \right. \quad (12)$$

Решение системы уравнений (12) позволит вычислить оптимальное распределение входного потока запросов при оптимальных параметрах производительности серверов, а также сами эти параметры. Данная задача является многомерной. Решения многомерных задач всегда очень сложны, поэтому для простоты изложения предлагается упростить решение системы уравнений (12), разбив его на два этапа.

В качестве примера возьмем задачу оптимизации работы кластера из двух серверов, которая связана с поиском оптимального распределения входного потока между ними и заключается в нахождении минимума критерия качества по величине α_i . С учетом параметров загрузки оборудования задача распределения входного потока между серверами будет иметь следующий вид:

$$\min_{\alpha_i} \left\{ S(\alpha_1, \alpha_2 \dots \alpha_n) = \sum_{i=1}^n \left\{ \frac{Pw_i \alpha_i^2 \Lambda^2 b_i^2}{2(1 - \alpha_i \Lambda b_i)} + Pd_i(b_i) \cdot (1 - \alpha_i \Lambda b_i) \right\} \right\}, \quad \sum_{i=1}^n \alpha_i = 1, \quad \alpha_i \geq 0. \quad (13)$$

Сумма долей входного потока на кластер равна общему потоку заявок на вход в видеоконференции, а доля запросов на каждый сервер в кластере – величина неотрицательная.

Частный случай задачи может быть представлен через систему уравнений. Оптимизация работы кластера является задачей математического программирования. Для первичной оценки системы и решения данной задачи применим метод множителей Лагранжа. Система уравнений для решения задачи распределения потока для двух серверов в кластере с учетом преобразований будет иметь вид

$$\begin{cases} -\frac{Pw_1\alpha_1^2\Lambda^2b_1(\alpha_1\Lambda b_1-2)}{2(\alpha_1\Lambda b_1-1)^2} - Pd_1\Lambda + \frac{Pw_2(1-\alpha_1)^2\Lambda^2b_2((1-\alpha_1)\Lambda b_2-2)}{2((1-\alpha_1)\Lambda b_2-1)^2} + Pd_2\Lambda = 0; \\ x = \frac{Pw_2(1-\alpha_1)^2\Lambda^2b_2((1-\alpha_1)\Lambda b_2-2)}{2((1-\alpha_1)\Lambda b_2-1)^2} + Pd_2\Lambda; \\ \alpha_2 = 1 - \alpha_1. \end{cases} \quad (14)$$

Типовым подходом к решению системы уравнений n с $n - 1$ неизвестными будет использование численных методов. Для решения численными методами написана программа, с помощью которой был осуществлен перебор с заданным шагом (0,001) всех возможных вариантов α_1 и α_2 , удовлетворяющих условиям $\sum_{i=1}^n \alpha_i = 1$, $\alpha_i > 0$. Распределение потока запросов между двумя серверами показано на рис. 1.

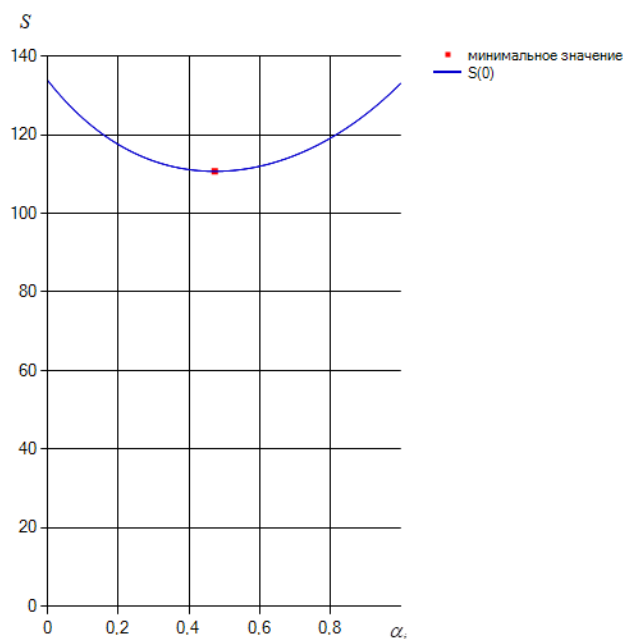


Рис. 1. График распределения потока запросов для кластера из двух серверов

Fig. 1. Diagram of the distribution of the flow of requests for a cluster of two servers

В результате пошаговых вычислений было рассчитано наиболее приемлемое распределение потока запросов в кластере из двух серверов: $\alpha_1 = 0,492$, $\alpha_2 = 0,508$ при $S(\alpha_1, \alpha_2) = 106,254$, что видно из рис. 1.

Оптимизация работы сервера в кластере по времени доступа к ресурсу также является задачей математического программирования. Для решения данной задачи применим метод множи-

телей Лагранжа. Выражение для решения задачи поиска времени доступа для серверов в кластере с учетом преобразований будет иметь вид

$$\min_{b_i} \left\{ S(b_i) = \sum_{i=1}^n \left\{ \frac{Pw_i \alpha_i^2 \Lambda^2 b_i^2}{2(1 - \alpha_i \Lambda b_i)} + Pd_i \cdot (1 - \alpha_i \Lambda b_i) \right\} \right\}. \quad (15)$$

Для кластера, состоящего из одного сервера, $\alpha = 1$. Тогда с учетом преобразований критерий качества по времени доступа к одному серверу будет определяться выражением

$$\min_b \left\{ S(b) = \frac{Pw\Lambda^2 b^2}{2(1 - \Lambda b)} + Pd(1 - \Lambda b) \right\}. \quad (16)$$

Величина времени доступа к серверу b_i , при которой рассчитанное значение критерия качества минимально, будет решением задачи оптимизации критерия качества.

Типовым подходом для решения задачи будет использование численных методов. Для решения численными методами написана программа, с помощью которой был осуществлен перебор с заданным шагом (0,001) всех возможных вариантов b . Для каждого значения времени доступа рассчитана величина критерия качества. Построим график зависимости величины критерия качества работы системы от величины времени доступа к серверу путем вычисления величины критерия качества $S(b)$ для значения времени доступа b , лежащего на отрезке $[0;2]$ (рис. 2).

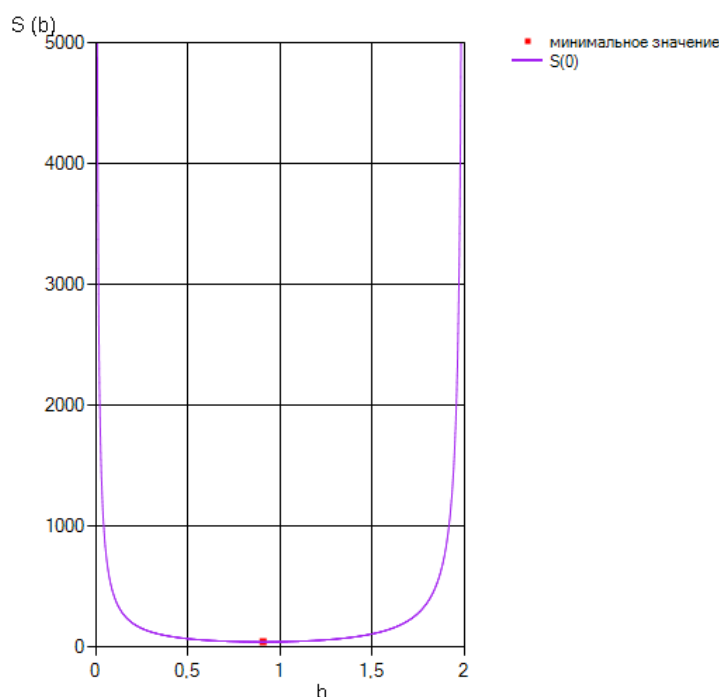


Рис. 2. График распределения производительности для одного сервера

Fig. 2. Diagram of performance distribution for one server

Пошаговые вычисления показали, что наиболее приемлемое значение критерия качества для различных распределений производительности сервера достигается при величине времени доступа $b = 0,908$, $S(b) = 43,768$ при $\Lambda = 0,5$, что соответствует графику на рис. 2.

Аналогично решению задачи для кластера из двух серверов задачу распределения можно решать в программном комплексе для n серверов при условии, что $n < 10$, так как с учетом по-

требностей в видео-конференц-связи на уровне крупных потребителей количество выделенных серверов редко превышает данные значения, а с учетом специфики учреждения образования и количества пользователей системы видео-конференц-связи – не превышает пяти. Ограничения в количестве серверов для кластера могут быть превышены, однако исходя из практики применения систем видео-конференц-связи, следует отметить, что у крупных компаний отсутствует «безграничный» подход к распределению вычислительных ресурсов. Так, зачастую оптимальнее иметь несколько высоконагруженных (более 60 % задействованных аппаратных ресурсов) серверов в кластере системы видео-конференц-связи, чем применять большое количество низконагруженных серверов.

Заключение. Для оптимального распределения больших потоков пользователей видеоконференции на базе кластерной архитектуры построения необходима доработка программного обеспечения балансировщика с учетом аппаратных составляющих серверов в кластере. Решение по распределению потока пользователей должно приниматься на основании загруженности всего кластера, что для системы видео-конференц-связи не всегда возможно ввиду специфики ее работы.

Доработанная математическая модель может быть адаптирована для веб-сервисов, вычислительных сервисов на базе грид-систем, а также для сервисов баз данных с учетом высоконагруженных аппаратных компонентов при работе данных сервисов и модификаций алгоритма распределения интенсивности входного потока запросов на системы.

Список использованных источников

1. Парамонов, А. И. Проблемы дистанционного образования и их прикладные решения в образовательных технологиях / А. И. Парамонов // Высшее техническое образование: проблемы и пути развития = Engineering education: challenges and developments : материалы X Междунар. науч.-метод. конф., Минск, 26 нояб. 2020 г. – Минск : БГУИР, 2020. – С. 182–187.
2. Марков, А. Н. Готовность учреждений высшего образования к цифровой трансформации процессов / А. Н. Марков, С. А. Мигалевич // Цифровая трансформация. – 2021. – № 2. – С. 64–68.
3. Марков, А. Н. Выбор сервиса видео-конференц-связи и его адаптация под учреждение образования / А. Н. Марков, Р. О. Игнатович, А. И. Парамонов // Информатика. – 2021. – Т. 18, № 4. – С. 17–25. <https://doi.org/10.37661/1816-0301-2021-18-4-17-25>
4. Lui, C. DNS and BIND / C. Lui, P. Albitz. – 5th ed. – O'Reilly Media, Inc., 2006. – 642 p.
5. Построение и исследование распределенной информационной системы веб-приложений и сервисов на базе кластерной архитектуры / В. А. Старых [и др.] // Качество. Инновации. Образование. – 2011. – № 12. – С. 91–99.
6. A survey of general-purpose computation on graphics hardware / J. D. Owens [et al.] // Computer Graphics Forum. – 2007. – Vol. 26, no. 1. – P. 80–113.
7. Dongarra, J. J. The LINPACK benchmark: Past, present and future / J. J. Dongarra, P. Luszczyk, A. Petitot // Concurrency and Computation: Practice and Experience. – 2003. – Vol. 15, no. 9. – P. 803–820.
8. High Performance Computing: Technology, Methods and Applications / J. Dongarra [et al.]. – Inc., 2015. – 444 p.
9. Dongarra, J. Adaptive linear solvers and eigensolvers / J. Dongarra // Argonne Training Program on Extreme-Scale Computing. – 2016. – Mode of access: https://extremecomputingtraining.anl.gov/files/2016/08/Dongara_830_AdaptiveLinear.pdf. – Date of access: 08.08.2022.

References

1. Paramonov A. I. *Distance education problems and their applied solutions in educational technologies*. Vysshee tehnikeskoe obrazovanie: problemy i puti razvitiya : materialy X Mezhdunarodnoj nauchno-metodicheskoy konferencii, Minsk, 26 nojabrja 2020 g. [*Engineering Education: Challenges and Developments : Materials of the X International Scientific and Methodological Conference, Minsk, 26 November 2020*]. Minsk, Belorusskij gosudarstvennyj universitet informatiki i radiojelektroniki, 2020, pp. 182–187 (In Russ.).
2. Markov A. N., Migalevich S. A. *Readiness of higher education institutions for digital transformation processes*. Cifrovaja transformacija [*Digital Transformation*], 2021, no. 2, pp. 64–68 (In Russ.).

3. Markov A. N., Ihnatovich R. O., Paramonov A. I. *Choosing a video conferencing service and its adaptation for educational institution*. Informatika [Informatics], 2021, vol. 18, no. 4, pp. 17–25 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-4-17-25>

4. Lui C., Albitz P. *DNS and BIND*, 5th edition. O'Reilly Media, Inc., 2006, 642 p.

5. Starykh V., Varyonov D., Ploskov S., Kuznetsov A. *Construction and research of a distributed information system of web applications and services based on cluster architecture*. Kachestvo. Innovacii. Obrazovanie [Quality. Innovation. Education], 2011, no. 12, pp. 91–99 (In Russ.).

6. Owens J. D., Luebke D., Govindaraju N., Harris M., Krüger J., ..., Purcell T. J. A survey of general-purpose computation on graphics hardware. *Computer Graphics Forum*, 2007, vol. 26, no. 1, pp. 80–113.

7. Dongarra J. J., Luszczek P., Petitet A. The LINPACK benchmark: Past, present and future. *Concurrency and Computation: Practice and Experience*, 2003, vol. 15, no. 9, pp. 803–820.

8. Dongarra J., Grandinetti L., Joubert G. R., Kowalik J. *High Performance Computing: Technology, Methods and Applications*. Inc., 2015, 444 p.

9. Dongarra J. Adaptive linear solvers and eigensolvers. *Argonne Training Program on Extreme-Scale Computing*, 2016. Available at: https://extremecomputingtraining.anl.gov/files/2016/08/Dongara_830_AdaptiveLinear.pdf (accessed 08.08.2022).

Информация об авторе

Марков Алексей Николаевич, аспирант, заместитель начальника Центра информатизации и инновационных разработок, Белорусский государственный университет информатики и радиоэлектроники.
ORCID <https://orcid.org/0000-0001-8508-1812>
E-mail: a.n.markov@bsuir.by

Information about the author

Aleksey N. Markov, Postgraduate Student, Deputy Head of the Center for Informatization and Innovative Development, Belarusian State University of Informatics and Radioelectronics.
ORCID <https://orcid.org/0000-0001-8508-1812>
E-mail: a.n.markov@bsuir.by



УДК 519.8
<https://doi.org/10.37661/1816-0301-2022-19-4-94-110>

Обзорная статья
Review Paper

Модели и методы машинного обучения для решения задач оптимизации и прогнозирования работы морских портов

М. Н. Лукашевич¹✉, М. Я. Ковалев²

¹*Белорусский государственный университет,
пр. Независимости, 4, Минск, 220050, Беларусь*
✉E-mail: mikhail.n.lukashevich@gmail.com

²*Объединенный институт проблем информатики
Национальной академии наук Беларуси,
ул. Сурганова, 6, Минск, 220012, Беларусь*

Аннотация

За последнее десятилетие существенно улучшились методы машинного обучения и расширилась сфера их применения, которая дополнилась рядом операционных задач, возникающих в грузовых портах. Это связано с накоплением и возможностью использования имеющихся в грузовых портах больших объемов данных. Статья посвящена обзору литературы по моделям и методам машинного обучения и их применению к оптимизации портовых операций. Основное внимание уделено планированию и развитию портов, их безопасности и охране, водным и сухопутным портовым операциям.

Ключевые слова: машинное обучение, прогнозирование работы портов, портовые операции, аналитический обзор, управление международными перевозками

Для цитирования. Лукашевич, М. Н. Модели и методы машинного обучения для решения задач оптимизации и прогнозирования работы морских портов / М. Н. Лукашевич, М. Я. Ковалев // Информатика. – 2022. – Т. 19, № 4. – С. 94–110. <https://doi.org/10.37661/1816-0301-2022-19-4-94-110>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 05.09.2022
Подписана в печать | Accepted 07.10.2022
Опубликована | Published 29.12.2022

Machine learning models and methods for solving optimization and forecasting problems of the work of seaports

Mikhail N. Lukashevich^{1✉}, Mikhail Y. Kovalyov²

¹Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220050, Belarus
✉E-mail: mikhail.n.lukashevich@gmail.com

²The United Institute of Informatics Problems
of the National Academy of Sciences of Belarus,
st. Surganova, 6, Minsk, 220012, Belarus

Abstract

Machine learning techniques have made significant advances and expanded application sphere over the past decade to include problems of port operations. This happened due to the growing amount of data available cargo ports. We review the literature on models and methods of machine learning and their application to optimization of port operations. A special attention is paid to the port planning and development a wide range of topics in port operations, including port planning and development, their safety and security, water and land port operations.

Keywords: machine learning, forecasting port forecasting, port operations, analytical review, international transportation management

For citation. Lukashevich M. N., Kovalyov M. Y. *Machine learning models and methods for solving optimization and forecasting problems of the work of seaports*. Informatika [Informatics], 2022, vol. 19, no. 4, pp. 94–110 (In Russ.). <https://doi.org/10.37661/1816-0301-2022-19-4-94-110>

Conflict of interest. The authors declare of no conflict of interest.

Введение. Морская торговля составляет почти девять десятых объема всей мировой торговли [1]. Морские торговые порты способствуют развитию мировой торговли, выступая в качестве перевалочных пунктов. Портовые операции разнообразны и могут быть разделены на следующие группы: планирование и развитие порта, безопасность и охрана, морские операции, операции складского двора и транспортной зоны [2]. Определение наиболее эффективного способа выполнения портовых операций может быть сложной задачей, включающей множество факторов, для которых методы машинного обучения (МО) являются приемлемыми инструментами решения. Однако применение МО для оптимизации портовых операций находится на стадии становления. На данный момент существует небольшое количество соответствующих публикаций, которые будут описаны ниже. Цель настоящей работы – короткий обзор основных методов МО в контексте задач оптимизации портовых операций. Начнем с описания самих задач оптимизации портовых операций и традиционных методов их решения.

Чен и др. [3] сосредоточились на составлении расписания для передачи контейнеров и кодировании проблемы с помощью двухкомпонентного хромосомного подхода, который затем реализуется с помощью модифицированного генетического алгоритма. Дахал и др. [4] объяснили, как моделировать, имитировать и оптимизировать портовые операции (хранение, обработку и отправку различных сыпучих грузов и т. д.) для достижения эффективного управления операциями. Де и др. [5] разработали модель смешанного целочисленного нелинейного программирования для решения проблемы устойчивой маршрутизации судов для ряда сценариев спроса и предложения в различных портах. Гафури и др. [6] предложили модель смешанного целочисленного линейного программирования для размещения гидролокаторов под водой с целью снижения риска терроризма, который может быть спровоцирован со стороны воды в портах и водных путях. Джон и др. [7] предложили новую методологию с использованием теории нечетких множеств, нечеткого аналитического иерархического аппарата и доказательного подхода для определения риска сбоя в работе морского порта с применением разнообразных и неточных данных, чтобы систематически оптимизировать эффективность работы системы. Ли и Ип [8] рассмотрели проблему планирования складских помещений и связанных

с ними домашних причалов в контейнерных терминалах и сформулировали ее в виде модели смешанного целочисленного линейного программирования. Линь и др. [9] проанализировали проблему планирования инвестиций для контейнерного терминала в порту Хумен (англ. Humen), используя моделирование с помощью программного обеспечения «Арена» (англ. Arena). Саид и Эль-Хорбати [10] предложили методику оптимизации обработки контейнеров с помощью генетических алгоритмов. Вентурини и др. [11] предложили модель целочисленного линейного программирования для задачи распределения причалов, интегрирующую оптимизацию скорости судов для нескольких портов на одной линии с учетом экологических соображений. Зейнебу и Абделлатиф [12] разработали математическую модель на основе генетического алгоритма, предлагающую распределение прибывающих судов на набережной таким образом, чтобы минимизировать общее расстояние, пройденное загруженными и разгруженными контейнерами в портовой зоне.

Основные методы машинного обучения. МО известно как область исследований, посвященная разработке вычислительных моделей и методов, интенсивно использующих большие объемы эмпирических данных. Оно берет свое начало от классических результатов в области статистики и искусственного интеллекта. За последние 10 лет МО превратилось в одну из самых популярных областей вычислительной науки благодаря усилиям таких крупных компаний, как Google, Microsoft, Facebook, Amazon и др. Благодаря их бизнес-процессам уже получены и продолжают накапливаться огромные объемы данных. Это дало шанс возродить статистические и вычислительные методы для автоматического создания эффективных моделей на основе данных. Методы МО можно разделить на три категории: обучение с учителем, обучение без учителя и обучение с подкреплением. Ниже описаны основные методы МО в этих трех категориях.

Обучение с учителем. Обучение с учителем использует размеченные обучающие данные и решает задачу построения взаимосвязей между входными и выходными данными. Эта задача называется классификацией, если выходные данные дискретны, и регрессией, если выходные данные непрерывны. Для получения выходных данных (прогнозов) на тестовых наборах, где заданы только входные данные, модели строятся на обучающих данных, состоящих из входов и выходов. Затем выходные данные модели сравниваются с сохраненными целевыми переменными, чтобы определить качество модели. В заключение описываются наиболее часто используемые алгоритмы обучения с учителем: линейная регрессия (англ. Linear Regression, LR), метод k -ближайших соседей (англ. k -nearest neighbors, KNN), искусственные нейронные сети (англ. Artificial Neural Networks, ANN), метод опорных векторов (англ. Support Vector Machine, SVM), случайный лес (англ. Random Forest, RF), градиентный бустинг (англ. Gradient Boosting, GB), байесовские сети и дерево решений (англ. Decision Tree, DT).

LR и KNN – одни из основных методов обучения с учителем. Для того чтобы получить (спрогнозировать) выходные данные, модели линейной регрессии делают предположения о структуре модели, например о значении ошибки и взаимосвязи между выходными и входными данными. KNN – непараметрический подход к обучению в отличие от линейной регрессии. Его идея заключается в том, что определяются k точек данных в обучающих (входных) данных, которые наиболее близки к новой точке входных данных. Для классификации и регрессии предсказание основывается на средней метке и метке большинства соответственно.

ANN – это метод обучения с учителем, который построен по принципу организации и функционирования биологических нейронных сетей (сетей нервных клеток живого организма). Его основная концепция заключается в извлечении линейных комбинаций признаков, а затем моделировании выходных данных как нелинейной функции признаков. Для обработки информации узлы нейронной сети связаны между собой таким образом, что позволяют им обучаться и адаптироваться с течением времени. ANN способны обучать очень сложные модели, но для обучения часто требуется большое количество данных.

SVM – еще один метод машинного обучения, в основном используемый для классификации. Расстояние между границами различных классов максимизируется, когда SVM создает линей-

ные границы в преобразованном виде пространства признаков. Метод может обрабатывать большое количество функций и решать проблему переобучения.

Ациклическая связная сеть, называемая деревом решений, используется для разделения набора данных на более управляемые однородные секции. Каждый узел в дереве представляет собой попарное сравнение признаков. Результаты сравнения показаны по каждой ветви. Прогноз можно сделать быстро и концептуально, просто используя дерево решений.

На основе описанных выше методов можно строить более сложные методы, комбинируя их друг с другом. Композитная модель может быть построена на основе данных нескольких отдельных моделей для повышения точности прогнозирования и устойчивости. Для этой цели также было разработано обучение ансамблей (англ. Ensemble Learning). В данном подходе есть два основных компонента: создание популяции базовых моделей на основе обучающих данных и объединение базовых моделей для получения композиционного предиктора. Методы ансамблевого обучения существуют в двух формах – бэггинга (англ. bagging) и бустинга (англ. boosting). В методе бэггинга несколько независимых моделей создаются из подмножеств обучающего набора данных параллельно, чтобы уменьшить дисперсию между моделями. При бустинге последовательные модели создаются упорядоченно, ориентируясь на ошибки предыдущих моделей, чтобы уменьшить ошибку предсказания. Наиболее часто используемыми моделями ансамблевого обучения являются случайные леса и градиентный бустинг. Кроме того, байесовские сети также можно рассматривать как форму ансамблевого обучения.

Обучение без учителя. Обучение без учителя использует немаркированные данные для поиска скрытых закономерностей в данных. Двумя наиболее широко используемыми методами являются метод главных компонент (англ. Principal Component Analysis, PCA) и кластерный анализ. PCA уменьшает размерность набора данных путем преобразования коррелированных переменных в меньшее количество некоррелированных главных компонент, каждая из которых является линейной проекцией данных. Метод используется как для исследовательского анализа данных, так и для прогнозирования, а также для извлечения и синтеза признаков. Концепция кластеризации относится к группированию данных в группы (кластеры) по сходству, но с поиском различий между группами. Анализируя данные в каждой группе, кластеры могут быть помечены. Кластеризацию также можно использовать для повышения производительности контролируемого машинного обучения путем добавления меток кластера в модель.

Обучение с подкреплением. Обучение с подкреплением (англ. Reinforcement Learning, RL) позволяет испытуемой системе (агенту) учиться методом проб и ошибок, используя отзывы о своих действиях и опыте. Целью обучения с подкреплением является максимизация заранее определенного кумулятивного вознаграждения. После того как действие завершено, агент получает вознаграждение, в котором подробно описывается немедленная реакция среды действия. Агент выбирает оптимальный план действий, исходя из вознаграждения и ситуации. В отличие от обучения с учителем, где обучающие данные уже содержат решение, в обучении с подкреплением эта информация отсутствует. Вместо этого агент выбирает, как выполнить конкретную миссию. Популярные методы обучения с подкреплением включают Q-обучение, состояние-действие, вознаграждение-состояние-действие (англ. State-Action, Reward-State-Action, SARSA) и глубокие Q-сети (англ. deep Q networks, DQN).

Сравнение классических методов машинного обучения и методов глубокого обучения. Естественные данные в необработанном формате и большом объеме представляют собой проблему для традиционных методов машинного обучения. Методы глубокого обучения (англ. Deep Learning, DL) позволяют бороться с этой проблемой. В отличие от традиционных подходов они могут различать сложные скрытые структуры в многомерных средах, что часто встречается в реальных условиях. Классические алгоритмы МО решают многие реальные проблемы, но они могут быть не в состоянии показать высокий результат при обработке многомерных или сложных наборов данных. Методы, основанные на искусственных нейронных сетях и глубоком обучении, предназначены для имитации принятия решений человеком посредством использования данных и в отличие от многих других алгоритмов МО могут обрабатывать многомерные данные. Глубокое обучение является представителем большого семейства методов МО и в за-

висимости от характера проблемы может быть классифицировано как обучение с учителем, обучение без учителя и обучение с подкреплением.

В следующих разделах рассмотрены области применения моделей и методов МО в портовых операциях.

Планирование и развитие портов. Планирование и развитие порта включает процесс стратегического планирования порта, сбор данных и прогнозирование, управление отношениями с заинтересованными сторонами и анализ эффективности проектов. Основные мероприятия, рассматриваемые в данном разделе (прогнозирование и прогнозное моделирование), не являются операциями по своей природе, но оказывают на них значительное влияние. Получение правильного прогноза позволяет минимизировать потенциальное недостаточное или избыточное использование мощностей порта, предотвращая финансовые потери и повышая его эффективность.

Авах и др. [13] использовали модели случайного леса и многослойного перцептрона (англ. Multi-Layer Perceptron, MLP) для прогнозирования оптимальной пропускной способности контейнеров, которую порт может физически обработать, учитывая уровень эффективности терминала. Было показано, что модель случайного леса обеспечивает более точный прогноз пропускной способности контейнеров в порту Дуала (англ. Douala), чем модель MLP или гибридная модель на основе оценок коэффициента детерминации, нормализованной среднеквадратичной ошибки, средней абсолютной ошибки и средней абсолютной процентной ошибки.

Чан и Сюй [14] разработали подход к прогнозированию пропускной способности контейнеров на основе метода опорных векторов регрессии (англ. Support Vector Regression, SVR), учитывая потенциальные факторы влияния, чтобы создать интеллектуальный логистический центр для портовых операций. В данной статье используются другие модели прогнозирования: метод опорных векторов и авторегрессионное интегрированное скользящее среднее (англ. Autoregressive Integrated Moving Average, ARIMA) для оценки эффективности предложенного подхода, и результаты показывают, что предложенный подход превосходит сравниваемые модели и может помочь в принятии решений заинтересованными сторонами. Чан и др. [15] сравнили несколько существующих методов прогнозирования временных рядов, а именно: скользящее среднее (англ. Moving Average, MA), многомерные адаптивные регрессионные сплайны (англ. Multivariate Adaptive Regression Splines, MARS), модель Грея (англ. Grey Model, GM), авторегрессионное интегрированное скользящее среднее, метод опорных векторов регрессии и искусственные нейронные сети, чтобы определить лучший метод для прогнозирования пропускной способности контейнеров. По результатам данного исследования метод опорных векторов регрессии представляется перспективным методом для прогнозирования пропускной способности контейнеров, тем более что он использует относительно небольшой размер выборки. Кроме того, исследование подтверждает, что, несмотря на различия в источниках данных, подходы МО все же более точны в прогнозировании, чем некоторые традиционные методы.

Куонг и др. [16] исследовали прогнозирование пропускной способности морских портов и политику восстановления после COVID-19, применяя эффективную стратегию принятия решений. Для прогнозирования пропускной способности портов Вьетнама они разработали ANN-модель, учитывающую динамику сотрудничества и конкуренции. Для обучения модели используются данные за 10 лет по объему контейнерных перевозок. В статье продемонстрирована эффективность предложенной стратегии с помощью системной динамики и теории управления. Согласно результатам моделирования устойчивое управление может поддерживать темпы роста пропускной способности контейнеров до 7,46 %.

Гамасса и Чен [17] предложили модель Грея, модель линейной регрессии, модель двойного экспоненциального сглаживания (англ. Double Exponential Smoothing, DES) и комбинированную модель прогнозирования для точного прогнозирования пропускной способности контейнеров Абиджанского порта (англ. Abidjan Port). Согласно оценке средней абсолютной процентной ошибки модель двойного экспоненциального сглаживания работает лучше, хотя комбинированная модель прогнозирования также хорошо себя зарекомендовала.

Генг и др. [18] представили робастную модель векторной регрессии (англ. Robust v-Support Vector Regression, RSVR) для прогнозирования пропускной способности порта. Авторы разработали алгоритм хаотического имитационного отжига (англ. Particle Swarm Optimization, PSO) с целью определения оптимальной комбинации параметров для модели RSVR. На основе данных о пропускной способности порта Шанхай с 1978 по 2012 г. этот подход определял инвестиции в основные фонды, импорт и экспорт, объем промышленного производства и ВВП, стоимость промышленной продукции, численность населения, общий объем розничных продаж потребительских товаров, объем грузоперевозок, объем грузоперевозок по шоссе и по железной дороге. Исследование показывает, что предложенная схема прогнозирования пропускной способности порта обеспечивает лучший результат прогнозирования по сравнению с шестью конкурирующими моделями с точки зрения ошибки прогнозирования.

Госасанг и др. [19] изучили проблему прогнозирования морского транспортного потока. Они сравнили многослойную персептронную нейронную сеть и линейную регрессию при прогнозировании количества контейнеров в порту Бангкока на основе исторических данных. Используемые данные охватывали период за 12 лет с 1999 по 2010 г., а выходными данными были количество контейнеров, объем импорта и объем экспорта. По результатам сравнения среднеквадратичной ошибки и средней абсолютной ошибки многослойный персептрон дает лучшие результаты, чем линейная регрессия, но процесс обучения проходит медленнее.

Для прогнозирования контейнерных перевозок в трех крупных морских портах Турции, включая Стамбул, Измир и Мерсин, Гёккуш и др. [20] провели всестороннее исследование. В данной работе были использованы четыре модели: искусственная нейронная сеть с искусственной пчелиной колонией (англ. Artificial Neural Network with Artificial Bee Colony, ANN-ABC), искусственная нейронная сеть, обученная алгоритмом Левенберга – Марквардта (англ. Artificial Neural Network with Levenberg – Marquardt, ANN-LM), множественная нелинейная регрессия с генетическим алгоритмом (англ. Multiple Nonlinear Regression with Genetic Algorithm, MNR-GA) и метод опорных векторов с наименьшим квадратом (англ. Least Square Vector Machine, LSSVM). Набор данных с 1989 по 2015 г. включал ВВП, численность населения, объем экспорта и пропускную способность контейнеров. Сравнение среднеквадратичной ошибки, средней абсолютной процентной ошибки и коэффициента детерминации для портов Измира и Стамбула показало, что LSSVM обеспечивает хорошие результаты, в то время как для порта Мерсин ANN-ABC и ANN-LM показали превосходные результаты.

Лам и др. [21] использовали искусственные нейронные сети для моделирования пропускной способности порта Гонконга. Входные переменные включали национальный ВВП, объемы импорта и экспорта, нераспределенного импорта, реэкспорта, внутреннего экспорта, численность населения, расходы на здания и сооружения, а также потребление электроэнергии в период с 1983 по 2001 г. На основании показателей коэффициента детерминации и средней абсолютной ошибки был сделан вывод, что все разработанные модели на основе искусственных нейронных сетей превосходят регрессионный анализ.

Миленкович и др. [22] предложили нечеткий нейросетевой подход к прогнозированию на основе метаэвристики для прогнозирования контейнерных потоков. Для разработки архитектуры нейронных сетей применялся как имитационный отжиг (англ. Simulated Annealing, SA), так и генетические алгоритмы. Нечеткие правила «если – то» (англ. the fuzzy if-then rules) были использованы в качестве эвристики для выбора между двумя методами. В данном исследовании непараметрические модели сравнивались с классическим параметрическим методом авторегрессионного интегрированного скользящего среднего на примере порта Барселоны.

Москосо-Лопес и др. [23] проанализировали и сравнили прогнозы (в килограммах) грузоперевозок импортных овощей в порту Альхесираса (англ. Algeciras). В своем исследовании ученые разработали и протестировали модели с помощью искусственных нейронных сетей и метода опорных векторов. Оба метода используют исторические данные и показали отличную способность предсказывать будущее на семь дней вперед.

Москосо-Лопес и др. [24] исследовали будущие значения потока скоропортящихся грузов Ro-Ro в порту залива Альхесирас с помощью системы прогнозирования на основе МО. В частности, две базы данных, соответствующие грузопотокам свежих фруктов и овощей в период

с 2010 по 2017 г., анализировались с использованием множественной линейной регрессии, метода опорных векторов и сетей с долговременной краткосрочной памятью (англ. Long Short-Term Memory, LSTM). Результаты анализа показали, что эта система, основанная на МО, достигла на 14,83 % лучшей производительности, чем базовая модель устойчивости, с точки зрения среднеквадратичной ошибки в наборе данных по свежим фруктам и на 11,3 % лучшей производительности в наборе данных по овощам.

Безопасность и охрана. Роль безопасности в работе порта становится все более существенной. Порт – это сложная организация, на которую влияют несколько социально-экономических и экологических факторов. Соответственно, любой инцидент в области безопасности может иметь серьезные экономические, экологические и человеческие последствия, способен потенциально повлиять на конкурентоспособность порта и, возможно, сделать его недоступным.

В портах могут возникнуть серьезные проблемы, если одновременно обслуживается большое количество судов. Следовательно, прогнозирование поступления судов может повысить безопасность порта. Основываясь на данных о погоде и зарегистрированных перемещениях судов с 2015 по 2020 г., Альварельос и др. [25] предложили искусственные нейронные сети и дерево принятия решений по повышению градиента (англ. Gradient Boosting Decision Tree, GBDT) для прогнозирования перемещений судов. Результаты показали, что искусственные нейронные сети достигли наилучших показателей с точки зрения среднеквадратичной ошибки и коэффициента детерминации.

Атак и Арсланогу [26] применили алгоритмы экстремального повышения градиента (англ. eXtreme Gradient Boosting, XGBoost), машину для повышения светового градиента (англ. Light Gradient Boosted Machine, Light GBM), метод опорных векторов, наивный Байес (англ. Naive Bayes, NB) и метод k -ближайших соседей для прогнозирования аварий на контейнерных терминалах в двух турецких портах, используя информацию о вместимости судов, времени работы, температуре, скорости ветра, влажности и частоте аварий. На основе исследования, представленного в данной работе, было показано, что алгоритмы экстремального повышения градиента, машина для повышения светового градиента и метод k -ближайших соседей дали наилучшие результаты с точностью более 98 %.

Ченг и др. [27] проанализировали проблемы, с которыми сталкиваются современные системы видеонаблюдения при обнаружении безопасности портовых операторов, и систематически исследуют методы извлечения информации и предупреждения поведения оператора. Проанализировав большое количество видеозаписей портовых операций и создав крупномасштабный набор данных сцен портовых операций под названием Harbor Dataset, они предложили алгоритм обнаружения объектов на основе глубокого обучения для определения безопасности оператора в порту. Эксперименты показали, что такой метод дает конкурентоспособный результат на базе данных Harbor Dataset при определении безопасности операторов и калибровке опасных зон работы порта.

Объединив методы МО и исследования операций, Хошино и др. [28] разработали эффективный и простой способ повышения эффективности работы морских портов Канадского агентства пограничных служб (англ. Canada Border Services Agency, CBSA). Авторы использовали концепцию, представленную Драммонд и Холте [29], для разработки кривой улучшения – новой метрики эффективности для прогнозных моделей. Кривые улучшения были использованы для измерения процессов оценки рисков в морских портах Канады. В статье [28] было показано, что даже разработка базовой прогностической модели могла бы повысить эффективность работы CBSA независимо от распределения по классам или затрат на ошибочную классификацию. Кроме того, теории, разработанные в данной статье, описывают простой способ оценки потенциального улучшения от внедрения прогностической модели в общий процесс оценки рисков для лиц, принимающих решения, без необходимости рассчитывать точное количество положительных и отрицательных контейнеров или знать точную стоимость вентиляционного или химического теста.

Ким и др. [30] исследовали аварии, происходящие в контейнерном порту, и предсказали их с помощью методов МО, включая нейросетевую модель, модель случайного леса и модель градиентного бустинга, используя временные ряды данных операций, аварий и погодных наблюдений в период с 2017 по 2020 г. для порта Пусан (англ. Busan), Республика Корея. Согласно полученным результатам модель глубокой нейронной сети и модель градиентного бустинга с интервалом в 6 ч обеспечивают наилучшую производительность с точки зрения метрик точности (англ. precision), полноты (англ. recall), среднего гармонического (англ. F1 score) и доли правильных ответов алгоритма (англ. accuracy).

Ли и др. [31] разработали стратегию МО для прогнозирования диапазона риска небезопасной скорости швартовки при подходе судна к порту. Чтобы определить, является ли скорость швартовки безопасной, авторы применили девять типов классификационных моделей МО: классификатор дерева решений (англ. Decision Tree Classifier, DTC), классификатор случайного леса (англ. Random Forest Classifier, RFC), бэггинг, классификатор дополнительных деревьев (англ. Extra Trees Classifier, ETC), классификатор повышения градиента (англ. Gradient Boosting Classifier, GBC), гауссовский наивный байесовский метод (англ. Gaussian Naive Bayes, GNB), метод k -ближайших соседей, метод опорных векторов и многослойный перцептрон. Для оценки эффективности моделей классификации МО использовались такие показатели, как доля правильных ответов алгоритма, полнота, точность и среднее гармоническое. В результате оценки классификатор дополнительных деревьев, классификатор случайного леса, бэггинг и классификатор повышения градиента показали наилучшие результаты.

Озгурк и др. [32] представили новую методику оценки риска столкновений на подходах к порту, основанную на МО и нечетких выводах, а также три новые характеристики (расстояние, площадь и скорость), которые могут улучшить модели риска навигационных столкновений, используемые в портовых бассейнах. Проведя эксперименты по имитации управления судном с участием 20 лоцманов-экспертов и проанализировав 140 маневров подхода лоцманов к порту на симуляторе, авторы исследования собрали большой массив данных. В результате применения модели случайного леса на основе полученных данных было установлено, что предложенный метод обеспечивает адекватное взвешивание серьезности маневров подхода к порту, а также относительной важности соответствующих параметров.

Сяо и др. [33] проанализировали детерминанты риска задержания судов, применив бинарную логистическую регрессию и дерево решений для оценки эффективности нового инспекционного режима (англ. New Inspection Regime, NIR). В исследовании использовался набор данных, содержащий 125 259 случаев инспектирования более чем из 20 стран в период с 2015 по 2017 г. Эмпирические результаты показали, что существенными являются следующие факторы: возраст судна, тип судна, показатели государства флага и количество недостатков, которые NIR считает существенными. Также было отмечено, что суда старше шести лет чаще оказываются некондиционными (англ. substandard), а если судно имеет пять и более недостатков, оно с большей вероятностью будет задержано.

Ян и др. [34] предложили модель искусственного интеллекта с учетом дисбаланса данных для отбора судов при контроле состояния порта на основе вероятности задержания. Используя 1600 записей инспекций в порту Гонконга за три года, они разработали модель классификации сбалансированного случайного леса (англ. Balanced Random Forest, BRF) для прогнозирования задержания судов. Численные эксперименты показали, что модель BRF может идентифицировать 81,25 % всех судов с задержанием в тестовом наборе, который содержит еще 400 записей инспекции. По сравнению с текущим методом отбора судов в порту Гонконга модель BRF намного эффективнее и может достичь среднего улучшения на 73,72 % при идентификации задержанных судов.

Территория складского двора. Территория складского двора считается самым сложным элементом контейнерных терминалов и представляет собой временную буферную зону, где контейнеры оставляют на время загрузки назначенного контейнеровоза или в ожидании отправки для распределения внутри порта. Чем больше контейнеровозов обрабатывает порт, тем большая требуется контейнерная складская площадка. Эффективность работы склада иногда

рассматривается как показатель конкурентоспособности терминала, поскольку она влияет на остальные показатели его работы.

Фотухи и др. [35] изучили задачу планирования работы крана на верфи, т. е. определения последовательности, в которой должны работать тягачи, чтобы свести к минимуму время их ожидания. Авторы смоделировали дворовые краны как агенты обучения с подкреплением и учли проблемы с помехами. Согласно экспериментальным результатам данная модель очень эффективно помогает крановщику выбрать следующий лучший ход. В связи с этим предлагаемая модель потенциально может быть интегрирована в существующие системы управления площадкой, чтобы автоматизировать процесс выбора грузовика и улучшить работу верфи.

Гао и др. [36, 37] исследовали прогнозирование объема контейнеров складского двора в порту и предложили модель рекуррентной нейронной сети (англ. Recurrent Neural Network, RNN) с сетями с долговременной краткосрочной памятью. Эти исследования предоставляют складскому двору возможность максимально использовать пространство и снизить скорость переворачивания контейнеров (англ. the rate of turning over containers), что может повысить эффективность работы.

Хотунг и др. [38] предлагают новый метод автоматизации трудоемкой и дорогостоящей разработки эвристики для задачи предварительной сортировки контейнеров (англ. Container Pre-Marshalling Problem, CPMP), который называется поиском эвристического дерева глубокого обучения (англ. Deep Learning Heuristic Tree Search, DLTS). Метод использует глубокие нейронные сети для изучения стратегий решения и нижних границ, настроенных для CPMP, исключительно путем анализа существующих почти оптимальных решений для экземпляров CPMP. Затем сети интегрируются в процедуру поиска по дереву, чтобы решить, какую ветвь выбрать следующей, и сократить дерево поиска. Модель обучена на основе 900 000 экземпляров, созданных Терней и др. [39]. На сегодняшний день DLTS производит эвристические решения высочайшего качества для CPMP с разрывами до оптимальности менее 2 % для экземпляров реального размера.

Канг и др. [40] исследовали проблему планирования мест хранения контейнеров с неопределенным весом и разработали подход на основе SA для ее решения. Новые стратегии, основанные на вычислительных экспериментах, фактически превосходят традиционную стратегию складирования групп одинакового веса (англ. same-weight-group stacking), поскольку значительно сокращают количество трудоемких операций по повторной обработке. Применение алгоритмов МО таких, как наивный Байес, дерево решений, алгоритм построения 1-правил (англ. 1R) и проективная адаптивная теория резонанса (англ. Projective Adaptive Resonance Theory, PART) к процессу создания хороших классификаторов продемонстрировало также возможность достижения более точной оценки веса, что приводит к еще лучшим решениям.

Куруниоти и др. [41] разработали методологическую основу, позволяющую операторам портов прогнозировать время простоя контейнеров и, следовательно, время прибытия грузовиков для вывоза импортных контейнеров. Для этого были использованы искусственные нейронные сети. Авторы проверили точность модели с помощью различных наборов независимых переменных и обнаружили, что ее можно повысить, если учесть больше информации, в то время как измеренная точность 65,17 % была не очень высокой.

Мойни и др. [42] рассмотрели три метода прогнозирования времени простоя в рабочих зонах контейнерного терминала: наивный Байес, дерево решений (C4.5) и гибридное байесовское дерево решений (англ. hybrid Bayesian decision tree). Основываясь на характеристиках C4.5, авторы использовали три сценария для изучения влияния изменений в определяющих факторах на время пребывания контейнера, пропускную способность двора и доход терминала от демеджа (англ. demurrage). Эти сценарии включали в себя изменение статуса контейнера с пустого на полный, закрытие ворот грузовика в условиях малого объема и изменение морского перевозчика. Несмотря на необходимость дальнейшего анализа, было продемонстрировано влияние изменений во времени простоя контейнеров и доходов от платы за демедж.

Сайкия и др. [43] предложили комбинацию эволюционных стратегий и методов обучения с подкреплением, чтобы найти приближенное оптимальное решение проблемы загрузки кон-

тейнера. На основе реальных данных модель была обучена с использованием синтетических данных, максимально приближенных к реальным данным. В данных было 23 слота для кораблей и блок из 49 контейнеров, сложенных в семь рядов по семь контейнеров в каждом штабеле. В этом исследовании модель достигла оптимального решения после приемлемого количества итераций.

Шен и др. [44] предложили метод на основе глубоких Q-сетей для решения проблем планирования загрузки судов. В этом методе используются восьмиуровневые глубокие Q-сети, которые обучаются на 19 слотах и 19 соответствующих контейнерах из четырехъярдовых отсеков в двух блоках в порту Нинбо (англ. Ningbo). Глубокие Q-сети в среднем тратят всего 0,131 с на планирование загрузки по сравнению с 237 с у портовых операторов. Таким образом, предложенный метод сопоставим с конкурентами-людьми и требует гораздо меньше времени.

Водные (морские) портовые операции. Стандартные задачи распределения причалов и кранов, прогнозирования времени прибытия судов, а также погрузки и выгрузки контейнеров на суда обычно решаются в рамках оптимизации морских операций на контейнерных терминалах. Сложность задачи возрастает по мере увеличения размера судна и роста ожиданий операторов морских перевозок в отношении качества и своевременности портовых услуг. Чтобы успешно конкурировать на глобальном уровне и достичь разумного времени вычислений для решения крупномасштабных задач, необходимо оптимизировать морские операции порта.

АбуАлхаол и др. [45] представили три показателя загруженности морских портов (англ. Port Congestion Indicators, PCI), а именно пространственную сложность, пространственную плотность и временную критичность, на основе статических и динамических сообщений автоматической идентификационной системы AIS. PCI были извлечены и рассчитаны на основе статических и динамических данных Big AIS. Используя три геопространственных алгоритма: площадь геохеша, площадь выпуклой части корпуса и среднюю близость судна, PCI для 2015 г. рассчитывались ежемесячно, а уровни перегрузки портов характеризовались с помощью кластеризации k -средних. По словам авторов, это первое исследование, в котором аналитика больших данных AIS применялась для оценки уровней загруженности морских портов.

Де Леон и др. [46] исследовали повышение производительности терминалов за счет управления такими ограниченными ресурсами, как причалы. Проблема распределения причалов (англ. Bulk Berth Allocation Problem, Bulk-BAP) была исследована с помощью системы на основе KNN, которая также позволяет неявно решать проблему выбора алгоритма (англ. Algorithm Selection Problem, ASP), когда доступно несколько алгоритмов. Оцениваются 12 подходов к решению, в основном жадный рандомизированный алгоритм, эвристический алгоритм «первым пришел – первым обслужен» и различные конфигурации метаэвристического поиска по большому соседству. Согласно исследованию качество решений улучшалось в случае, когда используемый алгоритм выбирался с учетом характеристик экземпляра, а не просто в случае применения наилучшего алгоритма.

Фанчелло и др. [47] исследовали прогнозирование времени прибытия и распределения человеческих ресурсов для контейнерных терминалов. Авторы статьи представили два алгоритма: алгоритм динамического обучения на основе нейронных сетей и алгоритм оптимизации распределения ресурсов. Использование этих алгоритмов позволяет судам прибывать в порт с меньшей неопределенностью и оптимизировать распределение операторов на весь день с учетом фактического спроса и работы терминала. Кроме того, поскольку эти алгоритмы основаны на общих переменных, их можно применять к любому перегрузочному терминалу.

Ким и др. [48] предложили объектно-ориентированные модели рассуждений на основе прецедентов (англ. Case-Based Reasoning framework, CBR) для раннего обнаружения задержек судов с использованием данных отслеживания S-AIS в режиме реального времени вместе с историческими данными (такими, как погрузочные данные). Этот подход позволяет обнаруживать задержки и прогнозировать схемы движения судов в режиме реального времени до их прибытия. Кроме того, авторы подчеркивают возможность улучшения прогнозов с помощью данных в реальном времени.

Сутрисновати и др. [49] обсудили использование байесовских сетей для анализа вероятности задержки при обработке контейнеров с учетом различных факторов. На основе журналов собы-

тий, созданных в системах управления портами, предлагаемая модель строится путем декомпозиции графа зависимостей. Предложенную модель можно использовать для определения достоверных вероятностей задержек операций и предоставления менеджерам портов предложений по улучшению существующих операций.

Верма и др. [50] описали структуру обучения с подкреплением для выбора и упорядочивания контейнеров для загрузки на корабли и сравнили ее эффективность с двумя метаэвристическими подходами, основанными на имитации отжига и генетических алгоритмах. Используя реальные оперативные данные из трех независимых наборов данных, авторы обучили предложенную модель на слоты в количестве 130–1391 и на контейнеры в количестве 25 000–7 000 000. Полученные результаты продемонстрировали, что предлагаемый подход решает вариант загрузки одного крана для целых судов с лучшими объективными значениями, чем подходы, которые используют стандартную метаэвристику.

Ван и др. [51] предложили модель, основанную на реляционном анализе Грея (англ. Gray Relational Analysis, GRA) и методе опорных векторов, для прогнозирования распределения контейнеров между контейнеровозами для рейса. Вначале в этой модели веса влияющих факторов определяются с помощью GRA. Взвешенные факторы служат входными данными модели метода опорных векторов, а параметры модели оптимизируются с помощью генетического алгоритма. Согласно полученным результатам предложенная модель может эффективно прогнозировать количество контейнеров для рейса контейнеровоза и демонстрирует высокую способность к обобщению и точность.

Ю и др. [52] изучили оценку времени прибытия судов и ее влияние на ежедневные операции контейнерного терминала. Авторы использовали три подхода к интеллектуальному анализу данных: нейронная сеть обратного распространения (англ. Back-Propagation Neural Network, BPNN), дерево классификации и регрессии (англ. Classification and Regression Tree, CART) и случайный лес. Результаты исследования показывают, что модель случайного леса превосходит модели BPNN и CART с точки зрения планирования работы терминала.

Сухопутные портовые операции (транспортная зона). Объем контейнерных перевозок значительно вырос во всем мире, что требует от крупных портов инвестиций и развития своих транспортных зон для повышения производительности.

Аль-Дик [53] представил два метода разработки моделей генерации грузовых рейсов в портах Майами и Джексонвилля, США: регрессионный анализ и нейронные сети с обратным распространением. Эти модели применяются для прогнозирования уровней входящего и исходящего потоков грузовых автомобилей в морских портах. Исследование продемонстрировало потенциал применения нейронных сетей для моделей генерации рейсов и модального разделения. Однако следует отметить, что набор данных был недостаточно большим для проведения исследования нейронных сетей.

Чу [54] проанализировал различные объекты порта с целью выявления факторов, оказывающих существенное влияние на формирование грузовых поездов в порту Гаосюн (англ. Kaohsiung), Тайвань. Было проведено сравнение точности прогнозирования модели множественной регрессии и нескольких моделей временных рядов (например, авторегрессионное интегрированное скользящее среднее, модели экспоненциального сглаживания и искусственной нейронной сети с обратным распространением). Результаты средних абсолютных процентных ошибок, средних абсолютных отклонений и средних квадратичных отклонений показали, что модель искусственной нейронной сети имеет наилучшую точность прогнозирования, за ней следуют модели регрессии и авторегрессионного интегрированного скользящего среднего, причем различия между ними относительно невелики.

Хилл и Бозе [55] описали систему поддержки принятия решений, которая предоставляет прогнозный анализ одновременно логистическим узлам и сотрудничающим автотранспортным компаниям. Авторы предложили гибкий облачный сервис, включающий искусственные нейронные сети для прогнозирования времени ожидания грузовых транспортных средств на складе пустых контейнеров. Исследователи обнаружили, что исключение ночных периодов и учет будних и праздничных дней при расчете входных переменных повышают точность модели.

Хьюн и Хатсон [56] изучили причины задержки движения грузовых транспортных средств на контейнерных терминалах, чтобы определить причины их аномально длительного времени оборота. В статье используются три модели дерева решений: автоматический детектор взаимодействия Хи-квадрат (англ. Chi-squared Automatic Interaction Detector, CHAID), дерево классификации и регрессии, и дерево решений (C4.5). Было продемонстрировано, что портовые власти и операторы терминалов могут эффективно использовать деревья решений для выявления причин аномально большого времени оборота грузовых транспортных средств за относительно короткий промежуток времени по сравнению с проведением исчерпывающего исследовательского анализа исторических данных.

Чон и др. [57] применили Q-обучение для маршрутизации автоматически управляемых транспортных средств (англ. Automated Guided Vehicles, AGV) на портовых контейнерных терминалах. В статье метод пытается найти маршруты с кратчайшим временем в пути для каждого заказа на доставку. Посредством моделирования производительность алгоритма обучения сравнивалась с эффективностью подхода, использующего маршруты с наименьшим расстоянием. Было обнаружено, что применение маршрутов на основе обучения вместо маршрутов кратчайшего расстояния может сократить время в пути на 17,3 %.

Фам и др. [58] исследовали оценку времени ожидания грузовиков в очереди у ворот морского терминала. Авторы статьи оценили и сравнили четыре метода регрессии для прогнозирования времени ожидания грузовиков в терминале: множественную линейную регрессию, нечеткую регрессию, нечеткую кластерную регрессию и метод опорных векторов. Результаты показателей производительности среднеквадратичной и средней абсолютной процентной ошибок показали, что модели нечеткой регрессии, метода опорных векторов и множественной линейной регрессии имеют сравнимую производительность, которая лучше, чем у модели регрессии с нечеткой кластеризацией.

Сарваредди и др. [59] изучили прогноз движения грузовых перевозок нефти, цитрусовых и пиломатериалов, которые не считаются контейнерными грузами. Кроме того, набор данных содержит сведения о количестве грузовиков и грузовых перевозках за год. Чтобы прогнозировать ежедневное количество грузовиков в порту Канаверал (англ. Canaveral), штат Флорида, были разработаны модели нейронной сети обратного распространения и полностью рекуррентной нейронной сети (англ. Fully Recurrent Neural Network, FRNN). Модель нейронной сети обратного распространения показала точность около 95 %, а модель полностью рекуррентной нейронной сети не имела достаточных данных для получения надежных результатов.

Ван Риссен и др. [60] изучили правила принятия решений в режиме реального времени для распределения контейнеров для внутренних перевозок. В статье был представлен общий четырехэтапный метод: сборка исторических данных, оптимизация исторических наборов, вывод дерева решений и применение дерева решений в режиме реального времени. Для решения исторических транспортных задач предлагаемый метод использует автономный подход к оптимальному планированию. Результаты переводятся в дерево решений с помощью метода логического вывода. Кроме того, математическая модель используется для получения нескольких автономных оптимальных значений с одинаковыми целевыми значениями в качестве входных данных для алгоритма обучения.

Цзэн и др. [61] предложили метод планирования, ориентированный на несколько кранов, для маршрутизации и диспетчеризации транспортных средств контейнерного терминала. Были разработаны две модели: одна для метода совместного использования частот между судами (англ. the inter-ship-based sharing method), другая – для метода совместного использования ресурсов судами (англ. the ship-based sharing method). Двухфазный алгоритм поиска Табу (англ. Tabu) и алгоритм Q-обучения предназначены для решения этих двух моделей соответственно. Численные результаты показали, что метод, основанный на межсудовых перевозках, может уменьшить порожний пробег трейлеров на верфи, оптимизировать работу трейлеров на верфи и тем самым повысить эффективность работы контейнерных терминалов. С помощью метода совместного использования судов можно уменьшить неравновесие различных рабочих линий и тем самым повысить эффективность погрузки или разгрузки.

Карта рассматриваемых в работе задач и методов МО

A map of the tasks and methods of MO considered in the work

Методы МО <i>MO methods</i>	Рассматриваемые в работе задачи <i>Tasks considered in the work</i>				
	БПиПА	МиУТС	ОМО	ОРСД	ППС
IR	-	-	-	+	-
ANN	+	+	+	+	+
ARIMA	-	+	-	-	+
Bagging	+	-	-	-	-
BPNN	-	+	+	-	-
BRF	+	-	-	-	-
CART	-	-	+	-	-
CBR	-	-	+	-	-
DES	-	-	-	-	+
DL	+	-	-	-	-
DLTS	-	-	-	+	-
DT	-	+	-	+	-
DTC	+	-	-	-	-
ETC	+	-	-	-	-
FRNN	-	+	-	-	-
GB	+	-	-	-	-
GBC	+	-	-	-	-
GBDT	+	-	-	-	-
GM	-	-	-	-	+
GNB	+	-	-	-	-
GRA	-	-	+	-	-
KNN	+	-	+	-	-
Light GBM	+	-	-	-	-
LR	+	+	-	-	+
LSSVM	-	-	-	-	+
LSTM	-	-	-	-	+
MA	-	-	-	-	+
MARS	-	-	-	-	+
MLP	+	-	-	-	+
MNR-GA	-	-	-	-	+
NB	+	-	+	+	-
PART	-	-	-	+	-
RF	+	-	+	-	+
RFC	+	-	-	-	-
RL	-	+	+	+	-
RNN	-	-	-	+	-
RSVR	-	-	-	-	+
SVM	+	+	-	-	+
SVR	-	-	-	-	+
XGBoost	+	-	-	-	-

Заключение. Все большее число практических задач решается с помощью методов МО. Не является исключением и портовая индустрия – отрасль, которая кажется консервативной и устойчивой к инновациям. Основные причины этого явления – прогресс в информационных технологиях и растущий спрос на портовые операции.

В настоящей статье представлен краткий обзор литературы по методам МО для улучшения портовых операций. Данная область быстро развивается, о чем свидетельствует ежегодный рост соответствующих публикаций. В таблице представлена карта рассмотренных в работе задач портовых операций и примененных к ним методов МО. В ней используются следующие сокращения: БПиПА – безопасность порта и прогнозирование аварий, МиУТС – маршрутизация и управление транспортными средствами, ОМО – оптимизация морских операций, ОРСД – оптимизация работ складского двора, ППС – прогнозирование пропускной способности. Знаком «+» отмечены задачи, для решения которых применялся указанный в таблице метод; знаком «-» – задачи, к которым данный метод не применялся.

References

1. Cheraghchi F., Abualhaol I., Falcon R., Abielmona R., Raahemi B., Petriu E. Big-data-enabled modelling and optimization of granular speed-based vessel schedule recovery problem. *2017 IEEE International Conference on Big Data (Big Data), Boston, USA, 11–14 December 2017*, pp. 1786–1794. <https://doi.org/10.1109/bigdata.2017.8258122>
2. Mekkaoui S. E., Benabbou L., Berrado A. A systematic literature review of machine learning applications for port's operation. *5th International Conference on Logistics Operations Management (GOL), Rabat, Morocco, 28–30 October 2020*, pp. 1–5. <https://doi.org/10.1109/GOL49479.2020.9314756>
3. Chen G., Govindan K., Yang Z. Managing truck arrivals with time windows to alleviate gate congestion at container terminals. *International Journal of Production Economics*, 2013, vol. 141, iss. 1, pp. 179–188. <https://doi.org/10.1016/J.IJPE.2012.03.033>
4. Dahal K., Galloway S., Hopkins I. Modelling, simulation and optimisation of port system management. *International Journal of Agile Systems and Management*, 2007, vol. 2, no. 1, pp. 92–108. <https://doi.org/10.1504/ijasm.2007.015683>
5. De A., Mamanduru V. K. R., Gunasekaran A., Subramanian N., Tiwari M. K. Composite particle algorithm for sustainable integrated dynamic ship routing and scheduling optimization. *Computers and Industrial Engineering*, 2016, vol. 96, pp. 201–215. <https://doi.org/10.1016/j.cie.2016.04.002>
6. Ghafoori A., Altiock T. A mixed integer programming framework for sonar placement to mitigate maritime security risk. *Journal of Transportation Security*, 2012, vol. 5, iss. 4, pp. 253–276. <https://doi.org/10.1007/s12198-012-0095-6>
7. John A., Paraskevadakis D., Bury A., Yang Z., Riahi R., Wang J. An integrated fuzzy risk assessment for seaport operations. *Safety Science*, 2014, vol. 68, pp. 180–194. <https://doi.org/10.1016/j.ssci.2014.04.001>
8. Li M. K., Yip T. L. Joint planning for yard storage space and home berths in container terminals. *International Journal of Production Research*, 2013, vol. 51, iss. 10, pp. 3143–3155. <https://doi.org/10.1080/00207543.2012.760852>
9. Lin J., Gao B., Zhang C. Simulation-based investment planning for Humen Port. *Simulation Modelling Practice and Theory*. 2014, vol. 40, pp. 161–175. <https://doi.org/10.1016/j.simpat.2013.09.009>
10. Said G. A. E.-N. A., El-Horbaty E.-S. M. An optimization methodology for container handling using genetic algorithm. *Procedia Computer Science*, 2015, vol. 65, pp. 662–671. <https://doi.org/10.1016/j.procs.2015.09.010>
11. Venturini G., Iris Ç., Kontovas C. A., Larsen A. The multi-port berth allocation problem with speed optimization and emission considerations. *Transportation Research Part D: Transport and Environment*, 2017, vol. 54, pp. 142–159. <https://doi.org/10.1016/j.trd.2017.05.002>
12. Zeinebou Z., Abdellatif B. Development of a model of decision support for optimization of physical flows in a container terminal. *2013 International Conference on Advanced Logistics and Transport, Tunisia, 29–31 May 2013*, pp. 421–426. <https://doi.org/10.1109/icadlt.2013.6568496>
13. Awah P. C., Nam H., Kim S. Short term forecast of container throughput: new variables application for the Port of Douala. *Journal of Marine Science and Engineering*, 2021, vol. 9, iss. 7, p. 720. <https://doi.org/10.3390/jmse9070720>
14. Chan H. K., Xu S. Intelligent port data management systems to improve capability. *2017 International Conference on Service Systems and Service Management, Tokyo, Japan, 26–30 November 2017*, pp. 1–3. <https://doi.org/10.1109/icsssm.2017.7996283>

15. Chan H. K., Xu S., Qi X. A comparison of time series methods for forecasting container throughput. *International Journal of Logistics Research and Applications*, 2019, vol. 22, iss. 3, pp. 294–303. <https://doi.org/10.1080/13675567.2018.1525342>
16. Cuong T. N., Kim H. S., You S. S., Nguyen D. A. Seaport throughput forecasting and post COVID-19 recovery policy by using effective decision-making strategy: A case study of Vietnam ports. *Computers and Industrial Engineering*, 2022, vol. 168, p. 108102. <https://doi.org/10.1016/j.cie.2022.108102>
17. Gamassa P. K. P., Chen Y. Application of several models for the forecasting of the container throughput of the Abidjan Port in Ivory Coast. *International Journal of Engineering Research in Africa*, 2017, vol. 28, pp. 157–168. <https://doi.org/10.4028/www.scientific.net/JERA.28.157>
18. Geng J., Li M. W., Dong Z. H., Liao Y. S. Port throughput forecasting by MARS-RSVR with chaotic simulated annealing particle swarm optimization algorithm. *Neurocomputing*, 2015, vol. 147, pp. 239–250. <https://doi.org/10.1016/j.neucom.2014.06.070>
19. Gosasang V., Chandraprakaikul W., Kiattisin S. A comparison of traditional and neural networks forecasting techniques for container throughput at Bangkok Port. *The Asian Journal of Shipping and Logistics*, 2011, vol. 27, no. 3, pp. 463–482. [https://doi.org/10.1016/s2092-5212\(11\)80022-2](https://doi.org/10.1016/s2092-5212(11)80022-2)
20. Gökkuş Ü., Yıldırım M. S., Aydın M. M. Estimation of container traffic at seaports by using several soft computing methods: a case of Turkish seaports. *Discrete Dynamics in Nature and Society*, 2017, vol. 2017, pp. 1–15. <https://doi.org/10.1155/2017/2984853>
21. Lam W. H. K., Ng P. L. P., Seabrooke W., Hui E. C. M. Forecasts and reliability analysis of port cargo throughput in Hong Kong. *Journal of Urban Planning and Development*, 2004, vol. 130, pp. 133–144. [https://doi.org/10.1061/\(asce\)0733-9488\(2004\)130:3\(133\)](https://doi.org/10.1061/(asce)0733-9488(2004)130:3(133))
22. Milenkovic M., Milosavljevic N., Bojovic N., Val S. Container flow forecasting through neural networks based on metaheuristics. *Operational Research*, 2021, vol. 21, iss. 2, pp. 965–997. <https://doi.org/10.1007/s12351-019-00477-1>
23. Moscoso-Lopez J. A., Turias I. T., Come M. J., Ruiz-Aguilar J. J., Cerban M. Short-term forecasting of intermodal freight using ANNs and SVR: case of the Port of Algeciras Bay. *Transportation Research Procedia*, 2016, vol. 18, pp. 108–114. <https://doi.org/10.1016/J.TRPRO.2016.12.015>
24. Moscoso-Lopez J. A., Urda D., Ruiz-Aguilar J. J., Gonzalez-Enrique J., Turias I. J. A machine learning-based forecasting system of perishable cargo flow in maritime transport. *Neurocomputing*, 2020, vol. 452, pp. 487–497. <https://doi.org/10.1016/j.neucom.2019.10.121>
25. Alvarellos A., Figuero A., Carro H., Costas R., Sande J., Guerra A., Rabuñal J. Machine learning based moored ship movement prediction. *Journal of Marine Science and Engineering*, 2021, vol. 9, iss. 8, p. 800. <https://doi.org/10.3390/jmse9080800>
26. Atak Ü., Arslanoğlu Y. Machine learning methods for predicting marine port accidents: a case study in container terminal. *Ships and Offshore Structures*, 2021, pp. 1–8. <https://doi.org/10.1080/17445302.2021.2003067>
27. Cheng G., Wang S., Guo T., Han X., Cai G., Gao F., Dong J. Abnormal behavior detection for harbour operator safety under complex video surveillance scenes. *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), Harbin Institute of Technology (Shenzhen), China, 15–17 December 2017*, pp. 324–328. <https://doi.org/10.1109/spac.2017.8304298>
28. Hoshino R., Coughtrey D., Sivaraja S., Volnyansky I., Auer S., Trichtchenko A. Applications and extensions of cost curves to marine container inspection. *Annals of Operations Research*, 2009, vol. 187, iss. 1, pp. 159–183. <https://doi.org/10.1007/s10479-009-0669-2>
29. Drummond C., Holte R. C. Cost curves: an improved method for visualizing classifier performance. *Machine Learning*, 2006, vol. 65, pp. 95–130. <https://doi.org/10.1007/s10994-006-8199-5>
30. Kim J. H., Kim J., Lee G., Park J. Machine learning-based models for accident prediction at a Korean container port. *Sustainability*, 2021, vol. 13, iss. 16, p. 9137. <https://doi.org/10.3390/su13169137>
31. Lee H. T., Lee J. S., Son W. J., Cho I. S. Development of machine learning strategy for predicting the risk range of ship's berthing velocity. *Journal of Marine Science and Engineering*, 2020, vol. 8, iss. 5, p. 376. <https://doi.org/10.3390/jmse8050376>
32. Ozturk U., Birbil S. I., Cicek K. Evaluating navigational risk of port approach manoeuvres with expert assessments and machine learning. *Ocean Engineering*, 2019, vol. 192, p. 106558. <https://doi.org/10.1016/j.oceaneng.2019.106558>
33. Xiao Y., Wang G., Lin K. C., Qi G., Li K. X. The effectiveness of the New Inspection Regime for Port State Control: Application of the Tokyo MoU. *Marine Policy*, 2020, vol. 115, p. 103857. <https://doi.org/10.1016/j.marpol.2020.103857>

34. Yan R., Wang S., Peng C. An artificial intelligence model considering data imbalance for ship selection in port state control based on detention probabilities. *Journal of Computational Science*, 2021, vol. 48, p. 101257. <https://doi.org/10.1016/j.jocs.2020.101257>
35. Fotuhi F., Huynh N., Vidal J. M., Xie Y. Modeling yard crane operators as reinforcement learning agents. *Research in Transportation Economics*, 2013, vol. 42, iss. 1, pp. 3–12. <https://doi.org/10.1016/j.retrec.2012.11.001>
36. Gao Y., Chang D., Chen C. H., Fang T. Deep learning with long short-term memory recurrent neural network for daily container volumes of storage yard predictions in port. *International Conference on Cyberworlds (CW), Singapore, 3–5 October 2018*, pp. 427–430. <https://doi.org/10.1109/cw.2018.00083>
37. Gao Y., Chang D., Fang T., Fan Y. The daily container volumes prediction of storage yard in port with long short-term memory recurrent neural network. *Journal of Advanced Transportation*, 2019, vol. 2019, pp. 1–11. <https://doi.org/10.1155/2019/5764602>
38. Hottung A., Tanaka S., Tierney K. Deep learning assisted heuristic tree search for the container pre-marshalling problem. *Computers and Operations Research*, 2020, vol. 113, p. 104781. <https://doi.org/10.1016/j.cor.2019.104781>
39. Tierney K., Pacino D., Voß S. Solving the pre-marshalling problem to optimality with A* and IDA*. *Flexible Services and Manufacturing Journal*, 2016, vol. 29, iss. 2, pp. 223–259. <https://doi.org/10.1007/s10696-016-9246-6>
40. Kang J., Ryu K. R., Kim K. H. Deriving stacking strategies for export containers with uncertain weight information. *Journal of Intelligent Manufacturing*, 2006, vol. 17, iss. 4, pp. 399–410. <https://doi.org/10.1007/s10845-005-0013-x>
41. Kourouniotti I., Polydoropoulou A., Tsiklidis C. Development of models predicting Dwell Time of import containers in port container terminals – an Artificial Neural Networks application. *Transportation Research Procedia*, 2016, vol. 14, pp. 243–252. <https://doi.org/10.1016/j.trpro.2016.05.061>
42. Moini N., Boile M., Theofanis S., Laventhal W. Estimating the determinant factors of container dwell times at seaports. *Maritime Economics and Logistics*, 2012, vol. 14, iss. 2, pp. 162–177. <https://doi.org/10.1057/mel.2012.3>
43. Saikia S., Verma R., Agarwal P., Shroff G., Vig L., Srinivasan A. Evolutionary RL for container loading. *ESANN 2018 Proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Belgium, 25–27 April 2018*. <https://doi.org/10.48550/arXiv.1805.06664>
44. Shen Y., Zhao N., Xia M., Du X. A Deep Q-Learning Network for ship stowage planning problem. *Polish Maritime Research*, 2017, vol. 24, no. s3, pp. 102–109. <https://doi.org/10.1515/pomr-2017-0111>
45. AbuAlhaol I., Falcon R., Abielmona R., Petriu E. Mining Port Congestion Indicators from Big AIS Data. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Brazil, 8–13 July 2018*, pp. 1–8. <https://doi.org/10.1109/ijcnn.2018.8489187>
46. De Leon A. D., Lalla-Ruiz E., Melian-Batista B., Marcos Moreno-Vega J. A Machine Learning-based system for berth scheduling at bulk terminals. *Expert Systems with Applications*, 2017, vol. 87, pp. 170–182. <https://doi.org/10.1016/j.eswa.2017.06.010>
47. Fancello G., Pani C., Pisano M., Serra M., Zuddas P., Fadda P. Prediction of arrival times and human resources allocation for container terminal. *Maritime Economics and Logistic*, 2011, vol. 13, iss. 2, pp. 142–173. <https://doi.org/10.1057/mel.2011.3>
48. Kim S., Kim H., Park Y. Early detection of vessel delays using combined historical and real-time information. *Journal of the Operational Research Society*, 2017, vol. 68, iss. 2, pp. 182–191. <https://doi.org/10.1057/s41274-016-0104-4>
49. Sutrisnowati R. A., Bae H., Song M. Bayesian network construction from event log for lateness analysis in port logistics. *Computers and Industrial Engineering*, 2015, vol. 89, pp. 53–66. <https://doi.org/10.1016/j.cie.2014.11.003>
50. Verma R., Saikia S., Khadilkar H., Agarwal P., Shroff G., Srinivasan A. A reinforcement learning framework for container selection and ship load sequencing in ports. *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 4), Montreal, 13–17 May 2019*, pp. 2250–2252.
51. Wang Y., Shi G., Sun X. A forecast model of the number of containers for containership voyage. *Algorithms*, 2018, vol. 11, no. 12, p. 193. <https://doi.org/10.3390/a11120193>
52. Yu J., Tang G., Song X., Yu X., Qi Y., Li D., Zhang Y. Ship arrival prediction and its value on daily container terminal operation. *Ocean Engineering*, 2018, vol. 157, pp. 73–86. <https://doi.org/10.1016/j.oceaneng.2018.03.038>
53. Al-Deek H. Which method is better for developing freight planning models at seaports-neural networks or multiple regression? *Transportation Research Record: Journal of the Transportation Research Board*, 2001, vol. 1763, iss. 1, pp. 90–97. <https://doi.org/10.3141/1763-14>

54. Chu C. H. Empirical method for predicting internal-external truck trips at a major port. *Journal of Transportation Engineering*, 2010, vol. 137, no. 7, pp. 496–508. [https://doi.org/10.1061/\(asce\)te.1943-5436.0000233](https://doi.org/10.1061/(asce)te.1943-5436.0000233)
55. Hill A., Bose J. W. A decision support system for improved resource planning and truck routing at logistic nodes. *Information Technology and Management*, 2017, vol. 18, iss. 3, pp. 241–251. <https://doi.org/10.1007/s10799-016-0267-3>
56. Huynh N., Hutson N. Mining the sources of delay for dray trucks at container terminals. *Transportation Research Record: Journal of the Transportation Research Board*, 2008, vol. 2066, iss. 1, pp. 41–49. <https://doi.org/10.3141/2066-05>
57. Jeon S. M., Kim K. H., Kopfer H. Routing automated guided vehicles in container terminals through the Q-learning technique. *Logistics Research*, 2010, vol. 3, iss. 1, pp. 19–27. <https://doi.org/10.1007/s12159-010-0042-5>
58. Pham Q., Huynh N., Xie Y. Estimating truck queuing time at marine terminal gates. *Transportation Research Record: Journal of the Transportation Research Board*, 2011, vol. 2222, iss. 1, pp. 45–53. <https://doi.org/10.3141/2222-06>
59. Sarvareddy P., Al-Deek H., Klodzinski J., Anagnostopoulos G. Evaluation of two modeling methods for generating heavy-truck trips at an intermodal facility by using vessel freight data. *Transportation Research Record: Journal of the Transportation Research Board*, 2005, vol. 1906, iss. 1, pp. 113–120. <https://doi.org/10.1177/0361198105190600114>
60. Van Riessen B., Negenborn R. R., Dekker R. Real-time container transport planning with decision trees based on offline obtained optimal solutions. *Decision Support Systems*, 2016, vol. 89, pp. 1–16. <https://doi.org/10.1016/j.dss.2016.06.004>
61. Zeng Q., Yang Z. Z., Lai L. Models and algorithms for multicrane oriented scheduling method in container terminals. *Transport Policy*, 2009, vol. 16, iss. 5, pp. 271–278. <https://doi.org/10.1016/J.TRANPOL.2009.08.006>

Информация об авторах

Лукашевич Михаил Николаевич, аспирант, факультет прикладной математики и информатики, Белорусский государственный университет.
E-mail: mikhail.n.lukashevich@gmail.com

Ковалев Михаил Яковлевич, член-корреспондент НАН Беларуси, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: kovalyov_my@newman.bas-net.by

Information about the authors

Mikhail N. Lukashevich, Postgraduate Student, the Faculty of Applied Mathematics and Computer Science, Belarusian State University.
E-mail: mikhail.n.lukashevich@gmail.com

Mikhail Y. Kovalyov, Corresponding Member of the National Academy of Sciences of Belarus, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.
E-mail: kovalyov_my@newman.bas-net.by

Правила для авторов

Редакция журнала «Информатика» просит авторов руководствоваться приведенными ниже правилами.

I. Статьи принимаются в редакцию через электронную систему подачи по адресу <http://inf.grid.by> в формате файлов текстовых редакторов Microsoft Word. Объем оригинальной статьи – от 8 до 16 стр., включая рисунки, таблицы и достаточное количество наиболее актуальных ссылок; объем обзорной статьи – от 16 до 32 стр., включая все основные ссылки. Текст набирается с переносами, шрифт Times New Roman 11 пт, интервал между строками – одинарный, абзацный отступ 0,5 см, поля по 2,5 см со всех сторон.

Материал статьи должен быть четко структурированным: Введение; основные разделы, в которых изложены цели и задачи, методы, результаты; Заключение (выводы).

II. Статьи о результатах работ, проведенных в научных учреждениях, должны иметь разрешение на публикацию (сопроводительное письмо за подписью руководителя или выписку из заседания ученого совета, отдела или кафедры, акт экспертизы).

III. Статьи в обязательном порядке должны включать аннотацию, ключевые слова, список литературы, информацию об авторах на русском и английском языках.

На титульной странице располагаются следующие метаданные:

1. Индекс по универсальной десятичной классификации (УДК); на русском и английском языках тип статьи (оригинальная или обзорная), название статьи, инициалы и фамилии всех авторов, полное наименование учреждений, где работают авторы, с указанием почтового адреса, при наличии указывается ученая степень и ORCID, e-mail ответственного лица.

2. Аннотация (Abstract) объемом 150–250 слов в оригинальной статье должна быть структурирована отдельными подразделами: Цели, Методы, Результаты, Заключение, а также максимально характеризовать содержательную часть рукописи. Сюда не следует включать впервые введенные термины, аббревиатуры (за исключением общеизвестных), ссылки на литературу.

3. Ключевые слова (Keywords) – наиболее значимые слова или словосочетания по теме работы, отражающие специфику темы, объекты и результаты исследования; перечень ключевых слов должен содержать 5–10 слов.

4. В разделе Благодарности (Acknowledgements) указываются все источники финансирования исследования, а также благодарности людям, которые участвовали в работе над статьей.

5. Автор обязан уведомить редакцию о реальном или потенциальном конфликте интересов, включив информацию в раздел Конфликт интересов (Conflict of interest).

6. Формулы, рисунки, таблицы в статье нумеруются в соответствии с порядком их упоминания в тексте. Ссылки на рисунки и таблицы в тексте обязательны. Рисунки должны быть выполнены с хорошим разрешением в масштабе, позволяющем четко различать надписи и обозначения. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала. Подрисуночные подписи с расшифровкой всех позиций, представленных на рисунке, и названия таблиц набираются шрифтом гарнитуры основного текста размером 9 пт. Перевод подрисуночной подписи и пояснений к рисунку, а также перевод названия таблицы, заголовки строк или столбцов располагаются курсивом после русскоязычной версии.

7. Набор формул выполняется в формульном редакторе Microsoft Equation или Math Type. Прямым шрифтом набираются: греческие и русские буквы; математические символы (\sin , \lg , ∞); символы химических элементов (C, Cl, CH₃); цифры (римские и арабские); индексы (верхние и нижние), являющиеся сокращениями слов. Курсивом набираются латинские буквы, символы физических величин (в том числе и в индексе).

8. Список использованной литературы оформляется в соответствии с требованиями Высшей аттестационной комиссии Республики Беларусь (ГОСТ 7.5–2008). Номер литературной ссылки в тексте дается порядковым номером в квадратных скобках. Ссылаться на неопубликованные работы не допускается.

10. Отдельно оформляется References со следующей структурой: авторы (транслитерация), транслитерированное название монографии, *Перевод названия монографии на английский язык*. Выходные данные с обозначениями на английском языке. От транслитераций названий статей можно отказаться.

Ссылки на учебно-методическую литературу, ГОСТы, авторефераты, статистические отчеты в список не включаются, а оформляются в виде сносок (с подробными рекомендациями можно ознакомиться на сайте журнала в разделе Правила для авторов).

11. В разделе Информация об авторах (Information about the authors) приводятся ФИО авторов полностью, ученая степень, звание, должность, название организации, ORCID (при наличии).

IV. Все поступающие в редакцию рукописи проходят предварительную проверку на соответствие Правилам для авторов. Статья может быть возвращена автору на доработку с просьбой устранить недостатки или дополнить информацию. После проверки на соответствие правилам статья направляется рецензенту с указанием сроков рецензирования.

V. При наличии замечаний рецензента автору предоставляется определенное время на доработку рукописи. Статьи, направляемые на доработку, должны быть возвращены в исправленном виде с ответами на все замечания. Окончательное решение о публикации или отклонении рукописи принимается редколлегией журнала. При положительном заключении рецензента статья передается редактору для подготовки к печати. Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи.

VI. Редакция журнала предоставляет возможность первоочередного опубликования статей, представленных лицами, которые осуществляют послевузовское обучение (аспирантура, докторантура, соискательство) в год завершения обучения.

VII. Авторы несут ответственность за направление в редакцию статей, уже опубликованных ранее или принятых к публикации другими изданиями.

ИНДЕКСЫ

00827

для индивидуальных
подписчиков

008272

для предприятий
и организаций

4 | 2022

ТОМ ■ VOL. 19

ИНФОРМАТИКА ■ INFORMATICS

ISSN 1816-0301 (Print)



9 771816 030000