

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

ИНФОРМАТИКА

INFORMATICS

TOM VOL. **18**

3 | **2021**

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

Том 18, № 3, июль-сентябрь 2021

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., член-корр. НАН Беларуси,
генеральный директор 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., член-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., ОАО «АГАТ – системы управления» – управляющая
компания холдинга «Геоинформационные системы управления» (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, ОАО «Гипросвязь»
(Минск, Беларусь)

Кундас Семен Петрович, д-р техн. наук, проф., БНТУ (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, ИПМ РАН (Москва, Россия)

Харин Юрий Семенович, д-р физ.-мат. наук, проф., член-корр. 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 Чандigarха (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 18, № 3, июль-сентябрь 2021

Ответственный за выпуск *Мойсейчик Светлана Сергеевна*
Редактор *Гончаренко Галина Борисовна*
Корректор *Михайлова Анна Антоновна*
Компьютерная верстка *Бутевич Ольга Борисовна*

Сдано в набор 23.08.2021. Подписано в печать 20.09.2021. Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс. Ризография. Усл. печ. л. 13,2. Уч.-изд. л. 12,9. Тираж 40 экз. Заказ 6.

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. 18, no. 3, Jule-September 2021

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, Dr. Sci. (Phys.-Math.), Prof., Corr. Member of NASB, General Director
of the UIIP NASB (Minsk, Belarus)

Deputy Editor-in-Chief

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

Editorial Board

Sergey V. Ablameyko, Dr. Sci. (Eng.), Prof., Academician of NASB, BSU (Minsk, Belarus)

Uladimir V. Anishchanka, Cand. Sci. (Eng.), Associate Professor, SoftClub Ltd. (Minsk, Belarus)

Petr N. Bibilo, Dr. Sci. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Mikhail N. Bobov, Dr. Sci. (Eng.), Prof., Open Joint-Stock Company "AGAT – Control Systems –
Managing Company of Geoinformation Control Systems Holding" (Minsk, Belarus)

Alexandre B. Dolgui, Dr. Sci. (Eng.), Prof., IMT Atlantique (Nantes, France)

Alexander N. Dudin, Dr. Sci. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

Alexey A. Karpov, Dr. Sci. (Eng.), Associate Professor, SPII RAS (Saint Petersburg, Russia)

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

Viktor V. Krasnoproshin, Dr. Sci. (Eng.), Prof., BSU (Minsk, Belarus)

Alexander M. Krot, Dr. Sci. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Sergey V. Kruglikov, Dr. Sci. (Milit.), Cand. Sci. (Eng.), Associate Professor, JSV "Giprosvjaz" (Minsk, Belarus)

Semen P. Kundas, Dr. Sci. (Eng.), Prof., BNTU (Minsk, Belarus)

Nikolai A. Likhoded, Dr. Sci. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

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

Valery A. Sklyarov, Dr. Sci. (Eng.), Professor, University of Aveiro (Aveiro, Portugal)

Yuri N. Sotskov, Dr. Sci. (Phys.-Math.), Prof., UIIP NASB (Minsk, Belarus)

Alexander L. Stempkovsky, Dr. Sci. (Eng.), Prof., Academician of RAS, IPPM RAS (Moscow, Russia)

Yuriy S. Kharin, Dr. Sci. (Phys.-Math.), Prof., Corr. Member of NASB, RI APMI BSU (Minsk, Belarus)

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

Vyacheslav N. Yarmolik, Dr. Sci. (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. 18, no. 3, July-September 2021

Issue Head *Sviatlana S. Maiseichyk*

Editor *Halina B. Hancharenka*

Corrector *Hanna A. Mikhailava*

Computer Imposition *Volha B. Butsevich*

Sent for press 23.08.2021. Output 20.09.2021. Format 60×84 1/8. Offset paper. Headset Times. Riesography. Printed sheets 13,2. Publisher's signatures 12,9. Circulation 40 copies. Order 6.

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

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ

- Золоторевич Л. А., Ильинков В. А.** Контроль надежности защиты интегральных схем от троянов: кодирование и декодирование комбинационных структур7
- Ярмолик В. Н., Мрозек И., Леванцевич В. А., Деменковец Д. В.** Неразрушающие тесты с четным повторением адресов для запоминающих устройств18

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Харин А. Ю.** Последовательное статистическое принятие решений в задачах анализа потоков данных 36
- Шушкевич Г. Ч.** Аналитическое решение задачи экранирования низкочастотного магнитного поля тонкостенным цилиндрическим экраном в присутствии цилиндра48
- Сеньков А. Г.** Автоматическое регулирование температуры воздуха в теплице с учетом данных прогноза погоды59

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

- Порхун М. И., Вашкевич М. И.** Метод моделирования эффекта ухудшения частотного разрешения слуха у больных нейросенсорной тугоухостью 68
- Старовойтов В. В., Голуб Ю. И.** Нормализация данных в машинном обучении83

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

- Соболь А. М., Козлова Е. И., Чернявский Ю. А.** Применение логики первого порядка для выявления организаторов и исполнителей противоправных действий в преступных группах97
- Золотой С. А., Страшко И. Б., Котов Д. С., Нестерович И. М., Рубо В. В., Костюк К. И.** Совершенствование инструментальных средств обработки и анализа космической информации 106

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

CONTENTS

COMPUTER-AIDED DESIGN

- Zolotorevich L. A., Ilyinkov V. A.** Monitoring the reliability of integrated circuits protection against Trojans: encoding and decoding of combinational structures.....7
- Yarmolik V. N., Mrozek I., Levantsevich V. A., Demenkovets D. V.** Transparent memory tests with even repeating addresses for storage devices18

MATHEMATICAL MODELING

- Kharin A. Y.** Sequential statistical decision making in problems of data flows analysis..... 36
- Shushkevich G. Ch.** Analytical solution of problem of shielding low-frequency magnetic field by thin-walled cylindrical screen in presence of cylinder 48
- Senkov A. G.** Automatic control of air temperature in the greenhouse considering the weather forecast data 59

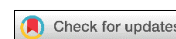
SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION

- Porhun M. I., Vashkevich M. I.** A method for simulation the effect of the reduced frequency resolution of the ear in patients with sensorineural hearing loss..... 68
- Starovoitov V. V., Golub Y. I.** Data normalization in machine learning..... 83

INFORMATION TECHNOLOGIES

- Sobol A. M., Kozlova E. I., Chernyavsky Y. A.** Application of first-order logic to identify organizers and perpetrators of illegal actions in teams of a limited circle of people 97
- Zolotoy S. A., Strashko I. B., Kotau D. S., Nestsiarovich I. M., Rouba V. V., Kostyuk K. I.** The improvement of the instrumental means for processing and analyzing of space information..... 106

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ COMPUTER-AIDED DESIGN



УДК 519.873:519.718.7
<https://doi.org/10.37661/1816-0301-2021-18-3-7-17>

Оригинальная статья
Original Paper

Контроль надежности защиты интегральных схем от троянов: кодирование и декодирование комбинационных структур

Л. А. Золоторевич[✉], В. А. Ильинков

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

Аннотация. Интегральные схемы и системы на кристалле являются ключевыми звеньями различных промышленных систем и систем обороноспособности государства. Появление контрафактных интегральных схем, проблемы пиратства, перепроизводства, несанкционированное вмешательство в проект микросхемы, аппаратные трояны требуют развития методов и средств их своевременного обнаружения. Трояны могут быть внесены в структуру интегральных схем при разработке и в процессе производства на этапах спецификации, проектирования, верификации и изготовления. Включение в структуру интегральных схем дополнительных элементов ставит под угрозу функциональную пригодность и надежность системы в целом. С целью аппаратной защиты проектов в настоящее время применяются методы аппаратного кодирования.

В работе рассматриваются особенности и надежность логического кодирования комбинационных схем. Предлагается алгоритм взлома кода комбинационных схем, основанный на описании закодированной структуры функцией разрешения и сведении задачи к КНФ-выполнимости. Исходными данными для декодирования структуры цифрового устройства являются структурная реализация закодированной схемы, полученная, например, методом обратного проектирования (проектирования по прототипу), а также активированный физический образец интегральной схемы, в защищенную от несанкционированного доступа память которой загружено правильное значение ключа. Этот образец может использоваться в виде модели черного ящика. Основная идея взлома ключа состоит в том, чтобы решить задачу, не прибегая к исследованиям на большом интервале значений входных и выходных переменных.

Ключевые слова: цифровое устройство, логическое кодирование, декодирование, функция разрешения, выполнимость КНФ-функции

Для цитирования. Золоторевич, Л. А. Контроль надежности защиты интегральных схем от троянов: кодирование и декодирование комбинационных структур / Л. А. Золоторевич, В. А. Ильинков // Информатика. – 2021. – Т. 18, № 3. – С. 7–17. <https://doi.org/10.37661/1816-0301-2021-18-3-7-17>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 21.04.2021
Подписана в печать | Accepted 24.05.2021
Опубликована | Published 29.09.2021

Monitoring the reliability of integrated circuits protection against Trojans: encoding and decoding of combinational structures

Lyudmila A. Zolotorevich[✉], Valery A. Ilyinkov

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

[✉]*E-mail: zolotorevichla.bsuir.by*

Abstract. Integrated circuits, systems on a chip are the key links in various industrial systems and state defense systems. The emergence of counterfeit integrated circuits, problems of piracy, overproduction, unauthorized interference in the design of microcircuit, hardware Trojans require the development of methods and means of their timely detection. Trojans can be introduced into the integrated circuits structure both on the development stage and during the production process, including the stages of specification, design, verification and manufacturing. The inclusion of additional elements in the integrated circuits structure jeopardizes the functional suitability and reliability of the system as a whole. For the purpose of hardware protection of projects, the methods of hardware coding are currently used.

The paper discusses the features and reliability of logical coding of combinational circuits. An algorithm for cracking the code of combinational circuits is proposed, based on the description of encoded structure by the resolution function and reducing the problem to SAT CNF. The initial data for decoding the structure of a digital device is the structural implementation of encoded circuit, obtained, for example, by reverse engineering (prototype design), as well as an activated physical sample of an integrated circuit, when into protected from unauthorized access memory the correct key value is loaded. This sample can be used as a black box model. The main idea of breaking a key is to solve a problem without research on a large interval of values of input and output variables.

Keywords: digital device, logical coding, decoding, resolution function, SAT CNF

For citation. Zolotorevich L. A., Ilyinkov V. A. Monitoring the reliability of integrated circuits protection against Trojans: encoding and decoding of combinational structures. *Informatiks*, 2021, vol. 18, no. 3, pp. 7–17 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-7-17>

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

Введение. Серьезными проблемами для электронной и оборонной промышленности в последние годы стали пиратство, перепроизводство и контрафакция, что привело к необходимости защиты проектов СБИС и систем на кристалле (СнК) от несанкционированного вмешательства в цикл проектирования и (или) производства интегральных схем (ИС) [1]. По оценкам службы обработки информации IHS (Technology Information Handling Services), финансовый риск из-за контрафактных и несанкционированных микросхем оценивается более чем в 169 млрд долл. в год, что примерно в 10 раз превышает ущерб от пиратства в области программного обеспечения [2]. Для оборонной промышленности важной задачей является возможность использования контрафактных ИС с модифицированными функциями, что в определенное время может деструктивно повлиять на функционирование структуры, ухудшить ее эксплуатационные характеристики, привести к раскрытию конфиденциальной информации и т. д. Кроме больших финансовых потерь, существует реальная проблема обеспечения национальной безопасности, так как 15 % ИС в системах оборонной промышленности являются контрафактными. В связи с этим стала очевидной необходимость защиты проектов на основе создания таксономии нарушений и отклонений, общего подхода к контролю СБИС и СнК, с моделями которых приходится работать при проектировании и организации контроля на всех этапах жизненного цикла цифровой системы с учетом злонамеренных внедрений в цикл проектирования и производства ИС. Как развитие теории контролепригодного проектирования (Design-for-Testability, DfT) в работе [3] предлагается подход к проектированию Design for-Trust (DfTr), который дополнительно включает средства для контроля и предотвращения аппаратных атак при проектировании и изготовлении СБИС.

В последние годы для защиты проектов ИС применяются методы и средства аппаратного кодирования комбинационных блоков. Для обеспечения надежности подобной защиты необходимы средства контроля эффективности применяемых методов кодирования и выявления внешних троянов на основе создания общего подхода к контролю проектов на всех этапах проектирования и производства.

В настоящей статье рассматриваются некоторые особенности метода логического кодирования структурных схем цифровых устройств комбинационного типа. Предлагается способ взлома кода при наличии информации о структуре закодированного объекта и возможности доступа к физической модели. Задача решается на основе описания закодированной структуры в виде КНФ-функции разрешения, решения задачи выполнимости (SAT) и физического моделирования объекта.

Логическое кодирование ИС как метод аппаратной защиты. В работе [2] проанализированы различные модели процесса злонамеренного искажения проекта, описывающие условия, при которых подобное искажение может внедриться в цифровую систему. В числе возможных источников искажений рассматриваются поставщики базовых функциональных блоков интеллектуальной собственности (IP's), которые приобретаются разработчиками СнК, собственно разработчики СнК, а также кремниевые фабрики – изготовители СнК. Методы несанкционированного доступа в проект могут быть различными. Например, они могут основываться на применении специальных средств САПР, способных исказить проект на RTL-уровне. В современных условиях наиболее уязвимым может оказаться этап производства.

Одним из методов борьбы с вышеупомянутыми угрозами является логическое кодирование, которое обеспечивает доступ к объекту только авторизованным пользователям [4]. Метод предполагает сокрытие функциональности проекта и использование ключа, который выводит систему в область правильного функционирования.

Основная идея кодирования состоит в том, чтобы изменить конструкцию ИС, добавив в нее дополнительные логические элементы и новые входы, называемые ключевыми, т. е. применить метод обфускации структуры объекта. В такой постановке если злоумышленник не владеет ключом, то ему недоступна внутренняя реализация объекта. Задача структурной обфускации и логического кодирования заключается в том, чтобы затруднить или сделать невозможным получение правильного ключа. Ключевые входы подсоединяются к защищенной от несанкционированного доступа памяти, а закодированная схема будет работать правильно, только если поданы верные значения на ее ключевые входы. Значения ключевых входов передаются после изготовления микросхем конечным пользователям (рис. 1). Таким образом, логическое кодирование основывается на предположении, что производитель не знает и не может вычислить правильные значения ключевых входов. В противном случае поиск правильного ключа должен быть для злоумышленника затруднителен.

Существуют различные методы кодирования комбинационной логики, в которых в качестве ключевых вентилей используются элементы XOR / XNOR [1, 5–7], AND / OR [8], мультиплексоры [9] или комбинации этих вентилей [10]. Выбор линии для включения вентиля и типа применяемого вентиля существенно влияет на эффективность кодирования. Воздействие неправильного ключа можно сравнить с влиянием неисправности константного типа на данной линии (рис. 1). В отличие от вентилей OR, NOR, AND, NAND при выборе в качестве ключевых вентилей XOR или XNOR применение неправильного ключа приводит к появлению неисправности константного типа в любом случае и на любом входном воздействии, что влияет в целом на эффективность кодирования. Кроме типа применяемого вентиля есть еще два основных способа увеличить влияние кодовых вентилей на значения выходов схемы. Один из них заключается в выборе линий, сигналы в которых влияют на максимально возможное количество выходов схемы, другой – в повышении чувствительности схемы в ответ на применение неправильного ключа.

Выбор линии для включения вентиля в большой степени влияет на эффективность кодирования. Один из подходов основан на случайном выборе линии схемы [11]. В работе [1] показана недостаточная эффективность этого подхода. Во-первых, вставка ключевого вентиля в случайно выбранную линию схемы не может гарантировать необходимое расстояние Хэмминга

между истинным выходным вектором и полученным в случае применения неправильного ключа. Оптимальное расстояние Хэмминга 50 % максимизирует двусмысленность злоумышленника относительно реакции схемы в случае применения неправильного ключа.

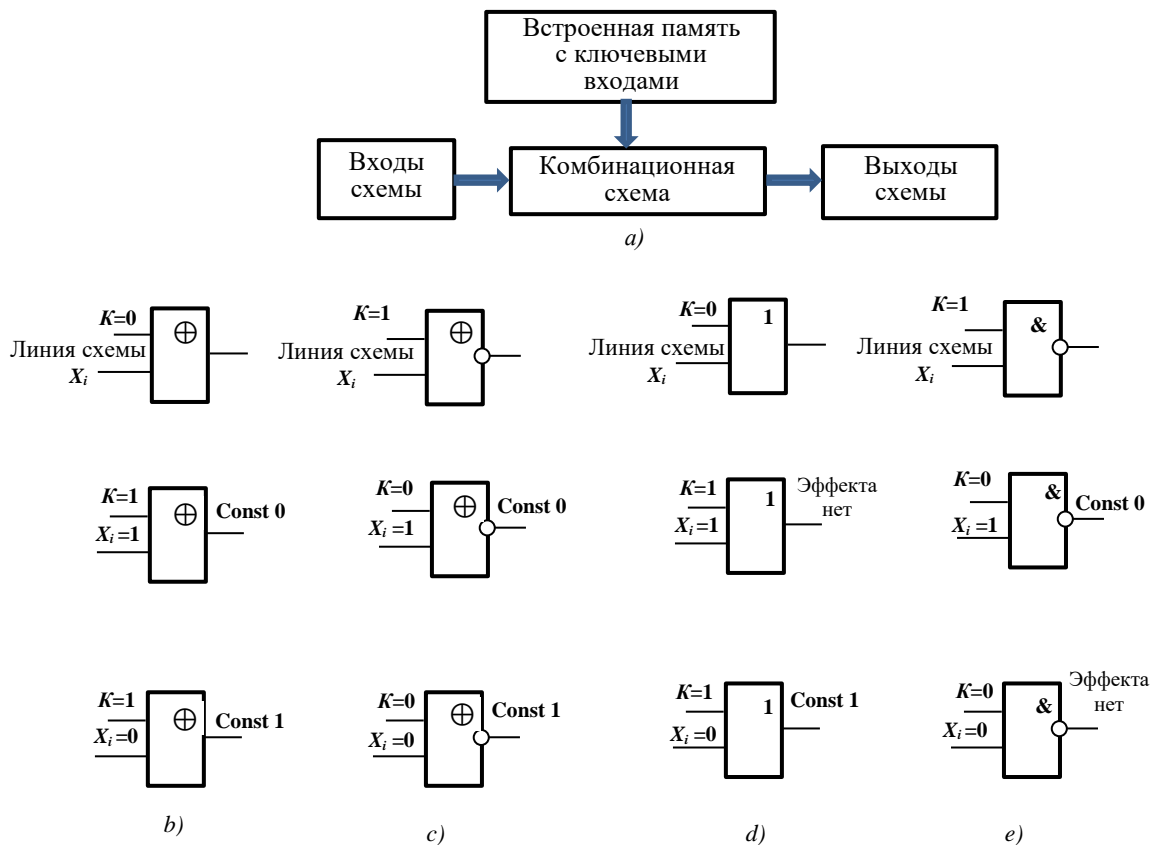


Рис. 1. Логическое кодирование цифровых устройств: а) общая идея кодирования; б) эффекты применения ключевого вентиля XOR-типа; в) NXOR; г) OR; д) NAND

Fig. 1. Logical coding of digital devices: a) the general idea of coding; b) the effects of using an XOR-type key gate; c) NXOR; d) OR; e) NAND

В работе [12] для характеристики эффективности выбора линии в схеме для введения ключевого вентиля предложено использовать метрику $M = N_0P_0 \cdot N_0O_0 + N_0P_1 \cdot N_0O_1$, где N_0P_0 (N_0P_1) – количество входных наборов, которые обнаруживают неисправность типа const 0 (const 1), а N_0O_0 (N_0O_1) – количество ошибочных бит выходного вектора в результате появления неисправности const 0 (const 1). Данная метрика может быть усовершенствована для получения возможности отслеживать в динамике параметры N_0O_0 (N_0O_1) для анализа неактивированных выходов при кодировании. Использование метрики M при кодировании можно сформулировать как нахождение множества неисправностей кодируемой схемы, которые вместе будут влиять на 50 % выходных линий при их активизации. Кодирование на основе использования метрики M требует моделирования схемы $Q = 2s \cdot 2^n$ раз, где s – общее количество линий схемы (переменных полного состояния схемы), n – количество входных переменных схемы. Для схемы на рис. 2 $M = 256 \cdot 34 = 8704$. Для реальных схем подобный подход практически неприемлем по причине высоких вычислительных затрат. В то же время с целью оптимизации вычислительных процедур предлагается эвристическое решение – сократить количество моделируемых входных наборов до 1000 [12].

Основная задача, которая должна быть решена при практической реализации данной общей идеи, заключается в том, чтобы определить оптимальное множество внутренних линий схемы

и количество ключевых элементов для создания максимальных трудностей для злоумышленника по поиску правильного ключа. При включении очередного вентиля в процессе кодирования логических устройств необходимо проводить анализ на появление эффекта маскирования неисправностей, который способен блокировать эффект кодирования [1, рис. 2]. При наличии избыточности отдельные линии схемы не могут быть активированы ни одним входным набором, поэтому вставка ключевого вентиля в данном случае может быть бесполезной [1, рис. 1].

В работе [1] на примере рассмотрена возможность повышения эффективности кодирования за счет включения в структуру схемы вентиля управления, которые позволяют активизировать влияние неправильного состояния каждого отдельного бита ключа на формирование выходного вектора закодированной схемы. Для того чтобы усилить влияние неправильного бита кодового слова на результат функционирования схемы, управляющие вентили объединяют биты кодового слова в группы, используя при этом их выходы в качестве входов ключевых вентилях. В этом случае реализуется групповое воздействие нескольких битов кодового слова на активацию ключевого вентиля. Если хотя бы один из ключевых входов, включенных в группу, принимает неправильное значение, ключевой вентиль окажется активированным.

Контроль надежности кодирования комбинационных схем. В работе [2] предлагается подход SAT-атаки для определения кода аппаратной защиты комбинационных схем цифровых устройств на структурном уровне. Подход основан на сведении задачи к определению выполнимости булевой функции.

Исходными данными для декодирования структуры цифрового устройства являются структурная реализация закодированной схемы, полученной, например, методом обратного проектирования (проектирования по прототипу), а также активированный физический образец ИС, в защищенную от несанкционированного доступа память которой заказчик загрузил правильное значение ключа. Этот образец может использоваться в виде модели черного ящика $Y = eval(X)$. Основная идея SAT-атаки взлома ключа состоит в том, чтобы определить правильный ключ, не прибегая к исследованиям на большом интервале входных (выходных) переменных [2].

Обозначим $\vec{Y} = f(\vec{X})$ функцию, реализуемую комбинационной схемой с первичными входами \vec{X} и выходами \vec{Y} , а $Cir_a(\vec{X}, \vec{Y})$ – КНФ функции разрешения исходной схемы. Сведем задачу получения ключа к описанию закодированной схемы в виде КНФ-представления булевой функции разрешения $Cir_b(\vec{X}, \vec{K}, \vec{Y})$, где \vec{X} – первичные входы схемы, $\vec{X} = (x_1, x_2, \dots, x_n)$; \vec{K} – ключевые входы схемы, $\vec{K} = (k_1, k_2, \dots, k_r)$; \vec{Y}_i – выходные линии схемы, $\vec{Y} = (y_1, y_2, \dots, y_m)$.

Если $F = f(\vec{X}, \vec{Y})$ – функция, реализуемая исходной схемой, то для любого \vec{X} $F = Cir_b(\vec{X}, \vec{K}, \vec{Y})$, если применить к закодированной схеме правильное значение ключа. Цель злоумышленника состоит в том, чтобы найти такой ключ $\vec{K} = (k_1, k_2, \dots, k_r)$, при котором $\forall \vec{X} \quad Cir_b(\vec{X}, \vec{K}, \vec{Y}) \wedge Cir_a(\vec{X}, \vec{Y})$. Однако злоумышленник не может получить формулу $Cir_a(\vec{X}, \vec{Y})$, так как ему недоступно структурное описание исходной схемы. Не получив доступ к структуре исходной схемы и не имея, таким образом, возможности построить отношение $Cir_a(\vec{X}, \vec{Y})$, злоумышленник может наблюдать реакцию схемы на требуемое входное воздействие по активированной ИС, выполнив функцию черного ящика $eval$:

$$\vec{X}_i = (x_1, x_2, \dots, x_n) \rightarrow \vec{Y}_i = (y_1, y_2, \dots, y_m).$$

Для заданного набора входных векторов $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_p$ и соответствующих выходных наблюдений $\vec{Y}_1, \vec{Y}_2, \dots, \vec{Y}_p$ определение ключевого значения, которое согласуется с p наблюде-

ниями, является достаточно простым, если свести задачу к решению выполнимости формулы $\bigwedge_{j=1}^p \text{Cir}_b(\vec{X}_j, \vec{K}, \vec{Y}_j)$. Однако если теперь выполнить новое наблюдение на физическом образце схемы $\text{eval}(\vec{X}_s) = \vec{Y}_s$, то нет гарантии, что удовлетворительное присваивание \vec{K} для формулы $\bigwedge_{j=1}^p \text{Cir}_b(\vec{X}_j, \vec{K}, \vec{Y}_j)$ также будет удовлетворительным присваиванием \vec{K} для формулы $\bigwedge_{j=p+1}^{2^n} \text{Cir}_b(\vec{X}_j, \vec{K}, \vec{Y}_j)$.

Для практической атаки при большом числе входных переменных функция eval может быть определена только на небольшом числе входных векторов $\text{Cir}_b(\vec{X}, \vec{K}, \vec{Y}) \Leftrightarrow \text{eval}(\vec{X}) = \vec{Y}$, в то время как $\exists \vec{K} : \forall \vec{X} \quad \text{Cir}_b(\vec{X}, \vec{K}, \vec{Y}) \wedge \text{Cir}_a(\vec{X}, \vec{Y})$.

Решение проблемы заключается в том, что вместо поиска правильного ключа выполняется определение ключа как члена класса эквивалентности ключей, который дает на выходах правильный результат для всех входных состояний.

Определение 1. Два ключа \vec{K}_1 и \vec{K}_2 являются эквивалентными ($\vec{K}_1 = \vec{K}_2$) тогда и только тогда, когда для входного значения \vec{X}_i закодированная схема выдает одинаковое выходное значение \vec{Y}_i для ключей \vec{K}_1 и \vec{K}_2 .

Для определения правильного ключа итеративно исключаются ключи из класса эквивалентности, которые выдают неправильные значения выходов по крайней мере для одного входного шаблона. Класс эквивалентных ключей определяется на некотором входном (выходном) векторе путем решения выполнимости функции $\text{Cir}_b(\vec{X}_j, \vec{K}, \vec{Y}_j)$ полным методом.

Определение 2. Входной вектор \vec{X}^d называется различающим, если реакция схемы при использовании ключа \vec{K}_1 равна \vec{Y}_1^d и отличается от реакции \vec{Y}_2^d при использовании ключа \vec{K}_2 .

При наличии различающего набора можно проверить реакцию активированной схемы для входа \vec{X}^d и использовать ее, чтобы исключить ключ \vec{K}_1 или \vec{K}_2 как не входящий в класс эквивалентности правильных ключей.

Алгоритм нахождения входного различающего набора:

1. $i := 1$.
2. $F_i = \text{Cir}_b(\vec{X}, \vec{K}_1, \vec{Y}_1) \wedge \text{Cir}_b(\vec{X}, \vec{K}_2, \vec{Y}_2)$.
3. Если условие $F_i \wedge \vec{Y}_1 \neq \vec{Y}_2$ не выполняется, переход к п. 8. Различающий набор не определен.
4. Решение $F_i = \text{Cir}_b(\vec{X}, \vec{K}_1, \vec{Y}_1) \wedge \text{Cir}_b(\vec{X}, \vec{K}_2, \vec{Y}_2) \wedge (\vec{Y}_1 \neq \vec{Y}_2)$, $\vec{X}_i^d := \vec{X}$. Входной набор \vec{X}_i^d является различающим.
5. $\vec{Y}_i^d := \text{eval}(\vec{X}_i^d)$.
6. $i = i + 1$.
7. $F_i = F_{i-1} \wedge \text{Cir}_b(\vec{X}_i^d, \vec{K}_1, \vec{Y}_i^d) \wedge \text{Cir}_b(\vec{X}_i^d, \vec{K}_2, \vec{Y}_i^d)$, переход к п. 3.
8. Выход.

Каждая итерация алгоритма исключает хотя бы один неверный член рассматриваемого класса эквивалентности ключей. Это связано с тем, что поиск различающего входного набора ведется с условием $\vec{Y}_1 \neq \vec{Y}_2$, т. е. при одинаковых входных данных выходные данные должны отличаться для разных ключей. Следовательно, хотя бы один ключ окажется неправильным. Алгоритм завершается, когда определен правильный ключ из класса эквивалентных ключей.

Покажем применение алгоритма на примере фрагмента схемы. Для полноты изложения рассмотрим получение функции разрешения F^f . Функция F^f , называемая функцией разрешения для логической функции f , зависит не только от аргументов функции f , но и от самой f и принимает значение логической единицы при всех допустимых состояниях входных и вы-

ходной переменных [14]. Функция F^f принимает значение 0 при всех недопустимых состояниях входных и выходной переменных. Представим функции разрешения F^f и запрета $\overline{F^f}$ в виде таблицы истинности для конъюнкции $f = a \cdot b$ (табл. 1).

Таблица 1. Функции разрешения и запрета для элемента AND
Table 1. Functions of permission and prohibition for AND element

a	b	f	F^f	$\overline{F^f}$
0	0	0	1	0
0	1	0	1	0
1	0	0	1	0
1	1	1	1	0
0	0	1	0	1
0	1	1	0	1
1	0	1	0	1
1	1	0	0	1

В табл. 2 приведены КНФ функций разрешения для некоторых типов вентилях элементов.

Таблица 2. Функции разрешения
Table 2. Functions of permission

Одноразрядные арифметические и логические уравнения	КНФ функций разрешения
$f = b \vee c$	$(\overline{b} \vee f)(\overline{c} \vee f)(b \vee c \vee \overline{f})$
$f = b \cdot c$	$(b \vee \overline{f})(c \vee \overline{f})(\overline{b} \vee \overline{c} \vee f)$
$f = a \oplus b$	$(a \vee b \vee \overline{f})(a \vee \overline{b} \vee f)(\overline{a} \vee b \vee \overline{f})(\overline{a} \vee \overline{b} \vee f)$
$f = a \sim b$	$(a \vee b \vee f)(a \vee \overline{b} \vee \overline{f})(\overline{a} \vee b \vee \overline{f})(\overline{a} \vee \overline{b} \vee f)$

Пример. На рис. 2, а изображена схема и вариант ее кодирования, которое выполнено путем включения дополнительных вентилях B_1 XOR и B_2 NXOR (рис. 2, б).

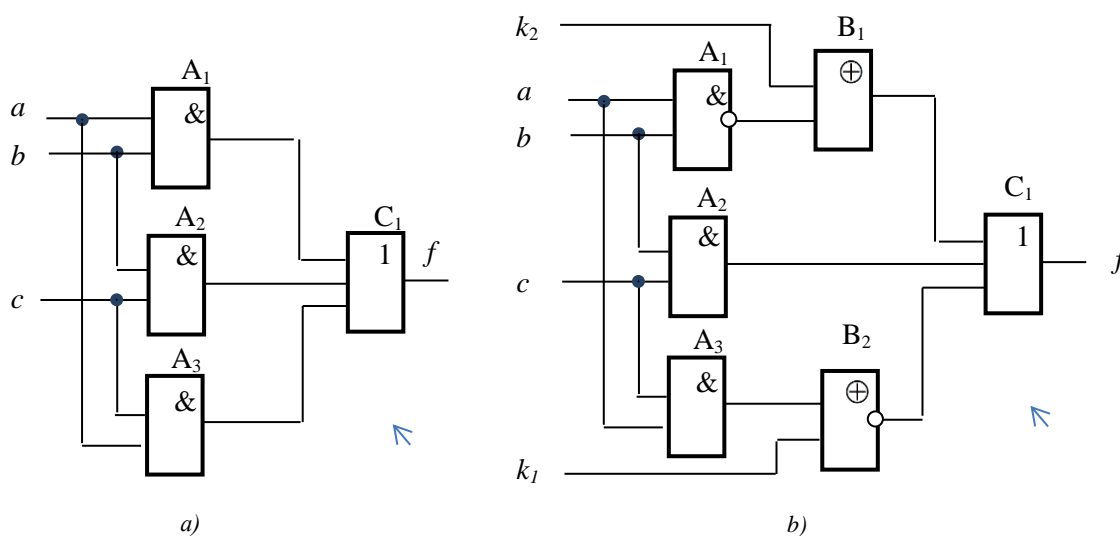


Рис. 2. Комбинационная схема для иллюстрации алгоритма взлома ключа:
а) исходная; б) закодированная

Fig. 2. Combinational scheme to illustrate the key cracking algorithm:
a) original; b) encoded

Приведем функцию разрешения закодированной схемы $Cir_b(\vec{X}, \vec{K}, \vec{Y})$. При формировании функции разрешения схемы a, b, c – входные переменные, а $a_1, a_2, a_3, b_1, b_2, c_1$ – выходы соответствующих элементов:

$$\begin{aligned} Cir_b &= (a \vee b \vee a_1)(a \vee \bar{b} \vee a_1)(\bar{a} \vee b \vee a_1)(\bar{a} \vee \bar{b} \vee a_1), \\ &(b \vee c \vee a_2)(b \vee \bar{c} \vee a_2)(\bar{b} \vee c \vee a_2)(\bar{b} \vee \bar{c} \vee a_2), \\ &(a \vee c \vee a_3)(a \vee \bar{c} \vee a_3)(\bar{a} \vee c \vee a_3)(\bar{a} \vee \bar{c} \vee a_3), \\ &(a_3 \vee k_1 \vee b_2)(a_3 \vee \bar{k}_1 \vee b_2)(\bar{a}_3 \vee k_1 \vee b_2)(\bar{a}_3 \vee \bar{k}_1 \vee b_2), \\ &(k_2 \vee a_1 \vee \bar{b}_1)(k_2 \vee \bar{a}_1 \vee b_1)(\bar{k}_2 \vee a_1 \vee b_1)(\bar{k}_2 \vee \bar{a}_1 \vee \bar{b}_1), \\ &(b_1 \vee a_2 \vee b_2 \vee \bar{c}_1)(b_1 \vee a_2 \vee \bar{b}_2 \vee c_1)(b_1 \vee \bar{a}_2 \vee b_2 \vee c_1)(b_1 \vee \bar{a}_2 \vee \bar{b}_2 \vee c_1), \\ &(\bar{b}_1 \vee a_2 \vee b_2 \vee c_1)(\bar{b}_1 \vee a_2 \vee \bar{b}_2 \vee c_1)(\bar{b}_1 \vee \bar{a}_2 \vee b_2 \vee c_1)(\bar{b}_1 \vee \bar{a}_2 \vee \bar{b}_2 \vee c_1). \end{aligned}$$

Для декодирования выполним следующие действия:

1. В качестве входного вектора для поиска ключей используем случайный вектор $\vec{X} = 110$, для которого определим \vec{Y} с помощью активированной схемы: $eval(\vec{X}) = 1$.

2. Найдем решение задачи SAT для функции $F = Cir_b ab\bar{c}c_1$ на основе полного алгоритма решения выполнимости:

$$F = \bar{a}_1 \bar{a}_2 \bar{a}_3 (k_1 \vee b_2)(\bar{k}_1 \vee \bar{b}_2)(k_2 \vee \bar{b}_1)(\bar{k}_2 \vee b_1) ab\bar{c}c_1.$$

Функция выполнима при следующих условиях:

$$F = k_1 k_2 \bar{a}_1 \bar{a}_2 \bar{a}_3 \bar{b}_1 \bar{b}_2 c_1, \quad \vec{K}_1 = 11;$$

$$F = \bar{k}_1 k_2 \bar{a}_1 \bar{a}_2 a_3 b_1 b_2 c_1, \quad \vec{K}_2 = 01;$$

$$F = \bar{k}_1 k_2 \bar{a}_1 \bar{a}_2 a_3 \bar{b}_1 b_2 c_1, \quad \vec{K}_3 = 00.$$

Таким образом, найдены три ключа: $\vec{K}_1 = 11$, $\vec{K}_2 = 01$, $\vec{K}_3 = 00$, которые составляют класс эквивалентных на данном этапе декодирования.

3. Найдем различающий входной набор для первых двух ключей $\vec{K}_1 = 11$ и $\vec{K}_2 = 01$ из найденного класса. Для этого необходимо вычислить булеву функцию

$$F_1 = Cir_b(\vec{X}, \vec{K}_1, \vec{Y}_1) \wedge Cir_b(\vec{X}, \vec{K}_2, \vec{Y}_2). \quad (1)$$

Для решения равенства (1) определим один из выполнимых входных (выходных) векторов для первого ключа $\vec{K}_1 = 11$. Задача решается на основе неполного алгоритма выполнимости функции

$$F = Cir_b k_1 k_2. \quad (2)$$

Получим $F = \bar{a}\bar{b}ck_1 k_2 a_1 \bar{a}_2 a_3 \bar{b}_1 b_2 c_1$. Таким образом определены новый входной $\vec{X} = 101$ и выходной $\vec{Y} = 1$ векторы. Проверим выполнимость функции $F = Cir_b \bar{a}\bar{b}ck_1 k_2 \bar{c}_1$ на входном

наборе $\vec{X} = 101$ при значении выходного вектора, отличного от полученного в равенстве (2) для второго ключа $\vec{K}_2 = 01$. В результате $F = \overline{a}b\overline{c}k_1k_2a_1a_2a_3\overline{b_1}b_2c_1$.

Следовательно, $\vec{X} = 101$ является различающим входным набором, так как разным ключам соответствуют разные выходы.

4. Определим вектор $\vec{Y} \Rightarrow \vec{X} = 101$, $\vec{Y} = eval(\vec{X}) = 1$ с помощью активированной схемы.

5. Вычислим функцию (1): для $\vec{K}_1 = 11$ $F_{K_1} = Cir_b\overline{a}b\overline{c}k_1k_2c_1$, для $\vec{K}_2 = 01$ $F_{K_2} = Cir_b\overline{a}b\overline{c}k_1k_2c_1$.

Функция $F_{K_2} = Cir_b\overline{a}b\overline{c}k_1k_2c_1$ не выполняется. Следовательно, ключ $\vec{K}_2 = 01$ исключается из класса эквивалентности.

6. Вычислим функцию (1) для $\vec{K}_3 = 00$: $F_{K_3} = Cir_b\overline{a}b\overline{c}k_1k_2c_1$. Функция F_{K_3} не выполняется, так как ключ $\vec{K}_3 = 00$ неправильный.

В связи с тем что ключи $\vec{K}_2 = 01$ и $\vec{K}_3 = 00$ оказались неверными, правильным является единственный ключ, оставшийся в классе эквивалентности правильных ключей, $\vec{K} = 11$.

Заключение. В работе рассмотрены некоторые особенности кодирования структурной реализации ИС на основе использования средств тестового диагностирования.

Для оценки надежности кодирования предлагается алгоритм декодирования, который проиллюстрирован на примере. Анализ надежности кодирования основан на решении SAT КНФ-функции разрешения, описывающей закодированную структуру.

Метод нахождения правильного ключа из класса эквивалентности предназначен для решения проблемы декодирования схем практических размеров без необходимости исследовать всю область возможных решений.

Вклад авторов. Л. А. Золоторевич – постановка задачи, анализ методов аппаратной защиты, разработка алгоритма декодирования; В. А. Ильинков – обоснование выбора вентиляного элемента для кодирования, анализ результатов на основе проведения компьютерного эксперимента, оформление статьи.

Список использованных источников

1. Золоторевич, Л. А. Аппаратная защита цифровых устройств / Л. А. Золоторевич // Вестник Томского гос. ун-та. Управление, вычислительная техника, информатика. – 2020. – № 50. – С. 69–78. <https://doi.org/10.17223/19988605/50/9>
2. Subramanyan, P. Evaluating the security of logic encryption algorithms / P. Subramanyan, S. Ray, S. Malik // 2015 IEEE Intern. Symp. on Hardware Oriented Security and Trust (HOST), Washington, DC, USA, 5–7 May 2015. – Washington, 2015. – P. 137–143.
3. Security analysis of integrated circuit camouflaging / J. Rajendran [et al.] // ACM SIGSAC Conf. on Computer & Communications Security, Berlin, Germany, 4–8 Nov. 2013. – Berlin, 2013. – P. 709–720.
4. Roy, J. A. EPIC: Ending piracy of integrated circuits / J. A. Roy, F. Koushanfar, I. L. Markov // IEEE Computer. – 2010. – Vol. 43, no. 10. – P. 30–38.
5. On improving the security of logic locking / M. Yasin [et al.] // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. – 2016. – Vol. 35, no. 9. – P. 1411–1424.
6. Logic encryption: a fault analysis perspective / J. Rajendran [et al.] // DATE '12 : Proc. of the Conf. on Design, Automation and Test in Europe, Dresden, Germany, March, 2012. – Dresden, 2012. – P. 953–958.
7. Fault analysis-based logic encryption / J. Rajendran [et al.] // IEEE Transactions on Computers. – 2015. – Vol. 64, no. 2. – P. 410–424.
8. A novel hardware logic encryption technique for thwarting illegal overproduction and hardware trojans / S. Dupuis [et al.] // 20th IEEE Intern. On-Line Testing Symp., Platja d'Aro, Catalunya, Spain, 7–9, July 2014. – Platja d'Aro, 2014. – P. 49–54.
9. Plaza, S. M. Solving the third-shift problem in IC piracy with test-aware logic locking / S. M. Plaza, I. L. Markov // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. – 2015. – Vol. 34, no. 6. – P. 961–971.

10. Lee, Y.-W. Improving logic obfuscation via logic cone analysis / Y.-W. Lee, N. Touba // Proc. Latin-American Test Symp., Puerto Vallarta, Mexico, 25–27 March 2015. – Puerto Vallarta, 2015. – P. 1–6.
11. Roy, J. A. Ending piracy of integrated circuits / J. A. Roy, F. Koushanfar, I. L. Markov // IEEE Computer. – 2010. – Vol. 43, no. 10. – P. 30–38.
12. Weighted logic locking: a new approach for IC piracy protection / N. Karousos [et al.] // IEEE 23rd Intern. Symp. on On-Line Testing and Robust System Design (IOLTS), Thessaloniki, Greece, 3–5 July 2017. – Thessaloniki, 2017. – P. 221–226.
13. Золоторевич, Л. А. Исследование методов и средств верификации проектов и генерации тестов МЭС / Л. А. Золоторевич // Сб. науч. тр. Всерос. науч.-техн. конф. «Проблемы разработки перспективных микроэлектронных систем» (МЭС–2006) / под общ. ред. А. Л. Стемповского. – М. : ИППМ РАН, 2006. – С. 163–168.
14. Zolotorevich, L. A. Project verification and construction of superchip tests at the RTL level / L. A. Zolotorevich // Automation and Remote Control. – N. Y. : Plenum Press, 2013. – Vol. 74, iss. 1. – P. 113–122.

References

1. Zolotorevich L. A. *Hardware protection of digital devices*. Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika, informatika [Bulletin of Tomsk State University. Management, Computer Technology, Informatics], 2020, no. 50, pp. 69–78. <https://doi.org/10.17223/19988605/50/9> (In Russ.).
2. Subramanyan P., Ray S., Malik S. Evaluating the security of logic encryption algorithms. *2015 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), Washington, DC, USA, 5–7 May 2015*. Washington, 2015, pp. 137–143.
3. Rajendran J., Sam M., Sinanoglu O., Karri R. Security analysis of integrated circuit camouflaging. *ACM SIGSAC Conference on Computer & Communications Security, Berlin, Germany, 4–8 November 2013*. Berlin, 2013, pp. 709–720.
4. Roy J. A., Koushanfar F., Markov I. L. EPIC: Ending piracy of integrated circuits. *IEEE Computer*, 2010, vol. 43, no. 10, pp. 30–38.
5. Yasin M., Rajendran J., Sinanoglu O., Karri R. On improving the security of logic locking. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2016, vol. 35, no. 9, pp. 1411–1424.
6. Rajendran J., Pino Y., Sinanoglu O., Karri R. Logic encryption: a fault analysis perspective. *DATE '12: Proceedings of the Conference on Design, Automation and Test in Europe, Dresden, Germany, March, 2012*. Dresden, 2012, pp. 953–958.
7. Rajendran J., Zhang H., Zhang C., Rose G. S., Pino Y., ..., Karri R. Fault analysis-based logic encryption. *IEEE Transactions on Computers*, 2015, vol. 64, no. 2, pp. 410–424.
8. Dupuis S., Ba P., Natale G. D., Flottes M., Rouzeyre B. A novel hardware logic encryption technique for thwarting illegal overproduction and hardware trojans. *20th IEEE International On-Line Testing Symposium, Platja d'Aro, Catalunya, Spain, 7–9 July 2014*. Platja d'Aro, 2014, pp. 49–54.
9. Plaza S. M., Markov I. L. Solving the third-shift problem in IC piracy with test-aware logic locking. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2015, vol. 34, no. 6, pp. 961–971.
10. Lee Y.-W., Touba N. Improving logic obfuscation via logic cone analysis. *Proceedings Latin-American Test Symposium, Puerto Vallarta, Mexico, 25–27 March 2015*. Puerto Vallarta, 2015, pp. 1–6.
11. Roy J. A., Koushanfar F., Markov I. L. Ending piracy of integrated circuits. *IEEE Computer*, 2010, vol. 43, no. 10, pp. 30–38.
12. Karousos N., Pexaras K., Karybali I. G., Kalligeros E. Weighted logic locking: a new approach for IC piracy protection. *IEEE 23rd International Symposium on On-Line Testing and Robust System Design (IOLTS), Thessaloniki, Greece, 3–5 July 2017*. Thessaloniki, 2017, pp. 221–226.
13. Zolotorevich L. A. Research of methods and means of project verification and test generation of MES. Sbornik nauchnyh trudov Vserossijskoj nauchno-tehnicheskoy konferencii "Problemy razrabotki perspektivnyh mikroelektronnyh sistem" (MJeS–2006) [Collection of Scientific Papers of the All-Russian Scientific and Technical Conference "Problems of Development of Promising Microelectronic Systems" (MES–2006)]. In Stempkovskij A. L. (ed.). Moscow, Institut problem proektirovaniya v mikroelektronike Rossijskoj akademii nauk, 2006, pp. 163–168 (In Russ.).
14. Zolotorevich L. A. Project verification and construction of superchip tests at the RTL level. *Automation and Remote Control*. New York, Plenum Press, 2013, vol. 74, iss. 1, pp. 113–122.

Информация об авторах

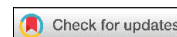
Золоторевич Людмила Андреевна, кандидат технических наук, доцент, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: zolotorevichla@bsuir.by

Ильинков Валерий Андреевич, кандидат технических наук, доцент, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: v.ilyinkov@gmail.com

Information about the authors

Lyudmila A. Zolotorevich, Cand. Sci. (Eng.), Assoc. Prof., Belarusian State University of Informatics and Radioelectronics.
E-mail: zolotorevichla@bsuir.by

Valery A. Ilyinkov, Cand. Sci. (Eng.), Assoc. Prof., Belarusian State University of Informatics and Radioelectronics.
E-mail: v.ilyinkov@gmail.com



УДК 004.33.054

<https://doi.org/10.37661/1816-0301-2021-18-3-18-35>

Оригинальная статья

Original Paper

Неразрушающие тесты с четным повторением адресов для запоминающих устройств

В. Н. Ярмолик¹✉, И. Мрозек², В. А. Леванцевич¹, Д. В. Деменковец¹¹Белорусский государственный университет

информатики и радиоэлектроники,

ул. П. Бровки, 6, 220013, Минск, Беларусь

✉E-mail: yarmolik10ru@yahoo.com

²Белостокский технический университет,

ул. Вейска, 45А, 15-351, Белосток, Польша

Аннотация. Показывается актуальность задачи тестирования запоминающих устройств современных вычислительных систем. Исследуются математические модели неисправностей этих устройств и используемые методы тестирования наиболее сложных из них на базе классических неразрушающих маршевых тестов. Вводится понятие адресных последовательностей (pA) с четным повторением адресов, которые являются основой базового элемента, входящего в структуру новых неразрушающих маршевых тестов $March_pA_1$ и $March_pA_2$. Приводятся алгоритмы формирования подобных последовательностей и примеры их реализации. Показывается максимальная диагностическая способность новых тестов для случая простейших неисправностей, таких как константные (SAF) и переходные (TF), а также сложных кодочувствительных неисправностей ($PNPSFk$). Отмечается существенно меньшая временная сложность тестов $March_pA_1$ и $March_pA_2$ по сравнению с классическими неразрушающими тестами, которая достигается за счет меньших временных затрат на получение эталонной сигнатуры. Вводятся новые метрики расстояния для количественного сравнения эффективности применяемых pA при однократной реализации тестов $March_pA_1$ и $March_pA_2$. В основе новых метрик лежит расстояние $D(A(j), pA)$, определяемое разностью индексов повторяющихся адресов $A(j)$ в последовательности pA . Исследуются свойства новых характеристик последовательностей pA и оценивается их применимость для выбора оптимальных тестовых последовательностей pA , обеспечивающих высокую эффективность новых неразрушающих тестов. Приводятся примеры вычисления метрик расстояний и показывается зависимость эффективности новых тестов от численных значений метрик расстояния. Как и в случае классических неразрушающих тестов, рассматривается многократное применение тестов $March_pA_1$ и $March_pA_2$. Вводится характеристика $V(pA)$, которая численно равняется количеству отличающихся значений расстояния $D(A(j), pA)$ адресов $A(j)$ последовательности pA . Экспериментально показывается справедливость аналитических оценок и подтверждается высокая эффективность обнаружения неисправностей однократными и многократными тестами типа $March_pA_1$ и $March_pA_2$ на примере неисправностей взаимного влияния для $p = 2$.

Ключевые слова: тестирование вычислительных систем, запоминающие устройства, неразрушающие маршевые тесты, адресные последовательности с четным повторением адресов, многократное неразрушающее тестирование

Для цитирования. Неразрушающие тесты с четным повторением адресов для запоминающих устройств / В. Н. Ярмолик [и др.] // Информатика. – 2021. – Т. 18, № 3. – С. 18–35. <https://doi.org/10.37661/1816-0301-2021-18-3-18-35>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 07.04.2021

Подписана в печать | Accepted 14.05.2021

Опубликована | Published 29.09.2021

Transparent memory tests with even repeating addresses for storage devices

Vyacheslav N. Yarmolik¹, Ireneusz Mrozek², Vladimir A. Levantsevich¹, Denis V. Demenkovets¹

¹Belarusian State University of Informatics and Radioelectronics,
st. P. Brovki, 6, 220013, Minsk, Belarus
✉E-mail: yarmolik10ru@yahoo.com

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

Abstract. The urgency of the problem of memory testing of modern computing systems is shown. Mathematical models describing the faulty states of storage devices and the methods used for their detection are investigated. The concept of address sequences (pA) with an even repetition of addresses is introduced, which are the basis of the basic element included in the structure of the new transparent march tests $March_{pA_1}$ and $March_{pA_2}$. Algorithms for the formation of such sequences and examples of their implementations are given. The maximum diagnostic ability of new tests is shown for the case of the simplest faults, such as constant (SAF) and transition faults (TF), as well as for complex pattern sensitive faults ($PNPSFk$). There is a significantly lower time complexity of the $March_{pA_1}$ and $March_{pA_2}$ tests compared to classical transparent tests, which is achieved at the expense of less time spent on obtaining a reference signature. New distance metrics are introduced to quantitatively compare the effectiveness of the applied pA address sequences in a single implementation of the $March_{pA_1}$ and $March_{pA_2}$ tests. The basis of new metrics is the distance $D(A(j), pA)$ determined by the difference between the indices of repeated addresses $A(j)$ in the sequence pA . The properties of new characteristics of the pA sequences are investigated and their applicability is evaluated for choosing the optimal test pA sequences that ensure the high efficiency of new transparent tests. Examples of calculating distance metrics are given and the dependence of the effectiveness of new tests on the numerical values of the distance metrics is shown. As well as in the case of classical transparent tests, multiple applications of new $March_{pA_1}$ and $March_{pA_2}$ tests are considered. The characteristic $V(pA)$ is introduced, which is numerically equal to the number of different values of the distance $D(A(j), pA)$ of addresses $A(j)$ of the sequence pA . The validity of analytical estimates is experimentally shown and high efficiency of fault detection by the tests $March_{pA_1}$ and $March_{pA_2}$ is confirmed by the example of coupling faults for $p = 2$.

Keywords: testing of computer systems, memory, transparent march tests, address sequences with even repeating addresses, multiple transparent testing

For citation. Yarmolik V. N., Mrozek I., Levantsevich V. A., Demenkovets D. V. Transparent memory tests with even repeating addresses for storage devices. *Informatics*, 2021, vol. 18, no. 3, pp.18–35 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-18-35>

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

Введение. Быстро прогрессирующие технологии производства изделий микроэлектроники позволяют создавать схемы запоминающих устройств (ЗУ) с высокой степенью интеграции, которые в современных вычислительных системах занимают до 94 % площади кристалла [1, 2]. Следствием использования новых технологий является увеличение вероятности возникновения различных неисправностей в ЗУ, которые могут составлять до 70 % от общего числа отказов системы в целом [3, 4]. Помимо традиционных методологий тестирования ЗУ в настоящее время широко применяется самотестирование памяти (*memory built-in self-test*) [5–7]. Такой подход позволяет периодически выполнять тестирование на рабочей частоте без применения внешнего оборудования. Из-за больших объемов современных ЗУ возможно применение только тестов сложности $O(N)$, где N – емкость в битах ЗУ, которые называются маршевыми тестами [3, 5, 6].

При реализации маршевого теста со всеми N ячейками ЗУ в заданном порядке (\uparrow – от младших адресов к старшим, \downarrow – в обратной последовательности, $\uparrow\downarrow$ – в любом направлении) выполняются операции, определяемые несколькими элементами, которые представляют собой

маршевый тест. Для этого используются: $w0, w1$ – запись в запоминающую ячейку памяти значения 0 или 1; $r0, r1$ – чтение текущего значения из элемента памяти и сравнение его со значением 0 или 1. Достоинствами маршевых тестов являются относительно высокая скорость выполнения и их обнаруживающая способность, а также простота реализации как средства тестирования. Наиболее значимый недостаток классических маршевых тестов заключается в потере хранящейся в памяти информации после выполнения теста.

Для надежной работы ЗУ современных вычислительных систем помимо методов, использующих избыточные данные (коды Хэмминга, контроль на четность и др.), применяются методы неразрушающего тестирования (*transparent testing*), основанные на классических маршевых тестах [8–10].

При анализе эффективности маршевых тестов рассматриваются математические модели, описывающие проявление физических дефектов ЗУ в процессе их функционирования [3, 4, 9, 10]. Наиболее распространенные модели неисправностей ЗУ приведены ниже [4].

Константные неисправности (*stuck-at fault, SAF*) характеризуются тем, что при их возникновении логическое значение ячейки памяти всегда равно 0 (*SA0*) или 1 (*SA1*) независимо от операции, производимой с этой или другими ячейками ЗУ.

Переходные неисправности (*transition fault, TF*) отличаются тем, что ячейка неспособна осуществлять переход из состояния логического 0 в состояние логической 1 ($TF\uparrow$) либо наоборот ($TF\downarrow$).

При возникновении неисправности взаимного влияния (*coupling fault, CF*) изменение логического значения одной влияющей ячейки (агрессора) отражается на значении второй зависимой ячейки (жертвы). Ячейка с меньшим адресом может влиять на ячейку со старшим адресом (\wedge), и наоборот (\vee). Существует три типа неисправностей *CF*. Во-первых, инверсная неисправность *CF* (*inversion CF, CF_{in}*), для которой изменение значения влияющей ячейки вызывает инвертирование значения зависимой. Возможны следующие виды неисправностей *CF_{in}*: $\wedge(\uparrow, \bar{b}), \wedge(\downarrow, \bar{b}), \vee(\uparrow, \bar{b}), \vee(\downarrow, \bar{b})$, где $b \in \{0, 1\}$ – текущее состояние ячейки жертвы. Во-вторых, неисправность *CF* прямого действия (*idempotent CF, CF_{id}*). В случае *CF_{id}* изменение значения влияющей ячейки переводит зависимую ячейку в определенное состояние. Возможны восемь видов неисправностей *CF_{id}*: $\wedge(\uparrow, 0), \wedge(\uparrow, 1), \wedge(\downarrow, 0), \wedge(\downarrow, 1), \vee(\uparrow, 0), \vee(\uparrow, 1), \vee(\downarrow, 0), \vee(\downarrow, 1)$. В-третьих, константная неисправность *CF_{st}* (*state CF, CF_{st}*), состоящая в том, что переход зависимой ячейки в какое-либо состояние возможен только при определенном значении влияющей ячейки. Различают следующие неисправности *CF_{st}*: $\wedge(0, 0), \wedge(0, 1), \wedge(1, 0), \wedge(1, 1), \vee(0, 0), \vee(0, 1), \vee(1, 0), \vee(1, 1)$.

Кодочувствительные неисправности (*pattern sensitive faults, PSF*) описывают поведение нескольких ячеек памяти вплоть до N ячеек, где N – емкость памяти в битах [4, 5]. Для подобных неисправностей логическое состояние одной ячейки памяти, называемой базовой (*base cell*), может зависеть от содержимого (0 или 1) или от логических переходов из 1 в 0 или из 0 в 1 в соседних ячейках (*neighborhood cells*) ЗУ. Среди множества разновидностей кодочувствительных неисправностей выделяют ограниченные (*restricted*) и граничные (*neighborhood, NPSF*) неисправности [4, 5]. При тестировании современных ЗУ обычно придерживаются модели кодочувствительных неисправностей, для которой рассматривается небольшое число $k \leq 9$ ячеек памяти, входящих в неисправность *NPSF_k*, а их местоположение может быть произвольным [4, 5, 9]. В качестве объекта исследования чаще всего рассматриваются пассивные кодочувствительные неисправности (*PNPSF_k*), где k обозначает количество произвольных ячеек памяти емкостью N бит, участвующих в конкретной неисправности. Для *PNPSF_k* содержимое базовой ячейки не может быть изменено из-за определенного набора данных в соседних $k - 1$ ячейках [4, 5]. Отметим, что результаты, полученные для *PNPSF_k*, легко обобщаются и для других классов кодочувствительных неисправностей в силу того, что *PNPSF_k* является моделью наиболее трудно обнаруживаемых неисправностей памяти, покрывающей другие виды неисправностей [4, 5].

В статье предлагается подход к построению неразрушающих тестов, основанный на применении адресных последовательностей с четным повторением адресов, позволяющий уменьшить временную сложность тестов и увеличить их диагностическую способность при сохранении обнаруживающей способности. Показывается зависимость эффективности новых неразрушающих тестов от применяемых адресных последовательностей и предоставляется оптимальный выбор для их многократной реализации.

1. Неразрушающие маршевые тесты с четным повторением адресов. Основным требованием, предъявляемым к неразрушающим тестам, является необходимость восстановления исходного состояния объекта после выполнения процедуры его тестирования. Поэтому все тестовые воздействия, направленные на активизацию неисправностей и проявление их в виде ошибок на выходе ЗУ, должны носить обратимый характер [10].

В настоящей работе при построении нового класса эффективных неразрушающих тестов ЗУ предлагается использовать модифицированные адресные последовательности. Первоначально рассмотрим общие свойства адресных последовательностей и их модификацию для реализации новых неразрушающих тестов.

Определение 1. Адресной последовательностью называется последовательность $A(j) = a_{m-1}a_{m-2}\dots a_2a_1a_0$, $j \in \{0, 1, 2, \dots, 2^m - 1\}$, где $a_i \in \{0, 1\}$ для $i \in \{0, 1, 2, \dots, m - 1\}$, состоящая из всех возможных 2^m m -разрядных двоичных комбинаций $a_{m-1}a_{m-2}\dots a_2a_1a_0$, генерируемых в произвольном порядке, причем каждый адрес формируется только один раз [5].

Для любой последовательности бит a_i , $i \in \{0, 1, 2, \dots, m - 1\}$, адресной последовательности A существует 2^{m-1} различных двоичных комбинаций $a_{m-1}a_{m-2}\dots a_{i+1}a_{i-1}\dots a_2a_1a_0$ для $a_i = 0$ и такое же количество комбинаций $a_{m-1}a_{m-2}\dots a_{i+1}a_{i-1}\dots a_2a_1a_0$ для $a_i = 1$. Для общего случая это свойство можно сформулировать следующим образом: для любого числа $1 \leq r < m$ последовательностей бит $a_\alpha, a_\beta, \dots, a_\chi$ адресов $A = a_{m-1}a_{m-2}\dots a_2a_1a_0$ для каждой из 2^r двоичных комбинаций $00\dots 0, 00\dots 1, \dots$ и $11\dots 1$ значений $a_\alpha, a_\beta, \dots, a_\chi$ существует ровно 2^{m-r} различных двоичных комбинаций в остальных $m - r$ последовательностях бит. Основываясь на данном свойстве, можно сформулировать формальный метод построения генераторов адресных последовательностей, в которых каждый адрес имеет кратность повторения адресов, равную $q = 2^r$. Приведем указанный метод в виде утверждения.

Утверждение. Удаление произвольных $1 \leq r < m$ бит $a_\alpha, a_\beta, \dots, a_\chi$ из исходной адресной последовательности $A = a_{m-1}a_{m-2}\dots a_2a_1a_0$, удовлетворяющей определению 1, приводит к тому, что оставшиеся биты последовательности будут формировать последовательность адресов, каждый из которых повторяется ровно $q = 2^r$ раз.

Подобные последовательности далее будем обозначать qA . При этом отметим, что алгоритм формирования qA такой же, как и для A . Отличием является использование в qA не всех бит исходной адресной последовательности A . В случае формирования каждого адреса дважды ($q = 2$) утверждение для получения последовательности $2A$ формулируется следующим образом: произвольная совокупность любых $m - 1$ разрядов $a_{m-1}a_{m-2}\dots a_{i+1}a_{i-1}\dots a_2a_1a_0$ из m разрядов $a_{m-1}a_{m-2}\dots a_2a_1a_0$ исходной адресной последовательности A формирует адресную последовательность $2A$, в которой каждый $m - 1$ разрядный адрес генерируется дважды.

Обобщая предыдущие рассуждения, сформулируем определение адресной последовательности pA для произвольного четного p .

Определение 2. Адресной последовательностью pA , состоящей из $p2^m$ адресов, называется упорядоченная последовательность адресов $A(j) = a_{m-1}a_{m-2}\dots a_2a_1a_0$, $j \in \{0, 1, 2, \dots, 2^m - 1\}$, где $a_i \in \{0, 1\}$ для $i \in \{0, 1, 2, \dots, m - 1\}$, состоящая из всех возможных 2^m m -разрядных двоичных комбинаций $a_{m-1}a_{m-2}\dots a_2a_1a_0$, каждая из которых формируется ровно p раз.

В качестве примера рассмотрим счетчиковую (пересчетную) адресную последовательность $A_C = c_{m-1}c_{m-2}\dots c_2c_1c_0$, где $c_i \in \{0, 1\}$, $i \in \{0, 1, 2, \dots, m - 1\}$, формируемую в соответствии с алгоритмом двоичного m -разрядного суммирующего счетчика [5]. Для исходной счетчиковой последовательности $A_C = c_3c_2c_1c_0$ в табл. 1 даны примеры последовательностей $2A_C$ и $4A_C$, полученных согласно утверждению.

Таблица 1. Адресная последовательность A_C и ее модификации $2A_C$ и $4A_C$ Table 1. Address sequences A_C and its $2A_C$ and $4A_C$ modifications

$A_C = c_3c_2c_1c_0$	$2A_C = c_3c_2c_1$	$2A_C = c_3c_2c_0$	$2A_C = c_3c_1c_0$	$2A_C = c_1c_0c_3$	$4A_C = c_3c_2$	$4A_C = c_3c_1$	$4A_C = c_3c_0$	$4A_C = c_0c_3$
0000 (0)	000 (0)	000 (0)	000 (0)	000 (0)	00 (0)	00 (0)	00 (0)	00 (0)
0001 (1)	000 (0)	001 (1)	001 (1)	010 (2)	00 (0)	00 (0)	01 (1)	10 (2)
0010 (2)	001 (1)	000 (0)	010 (2)	100 (4)	00 (0)	01 (1)	00 (0)	00 (0)
0011 (3)	001 (1)	001 (1)	011 (3)	110 (6)	00 (0)	01 (1)	01 (1)	10 (2)
0100 (4)	010 (2)	010 (2)	000 (0)	000 (0)	01 (1)	00 (0)	00 (0)	00 (0)
0101 (5)	010 (2)	011 (3)	001 (1)	010 (2)	01 (1)	00 (0)	01 (1)	10 (2)
0110 (6)	011 (3)	010 (2)	010 (2)	100 (4)	01 (1)	01 (1)	00 (0)	00 (0)
0111 (7)	011 (3)	011 (3)	011 (3)	110 (6)	01 (1)	01 (1)	01 (1)	10 (2)
1000 (8)	100 (4)	100 (4)	100 (4)	001 (1)	10 (2)	10 (2)	10 (2)	01 (1)
1001 (9)	100 (4)	101 (5)	101 (5)	011 (3)	10 (2)	10 (2)	11 (3)	11 (3)
1010 (10)	101 (5)	100 (4)	110 (6)	101 (5)	10 (2)	11 (3)	10 (2)	01 (1)
1011 (11)	101 (5)	101 (5)	111 (7)	111 (7)	10 (2)	11 (3)	11 (3)	11 (3)
1100 (12)	110 (6)	110 (6)	100 (4)	001 (1)	11 (3)	10 (2)	10 (2)	01 (1)
1101 (13)	110 (6)	111 (7)	101 (5)	011 (3)	11 (3)	10 (2)	11 (3)	11 (3)
1110 (14)	111 (7)	110 (6)	110 (6)	101 (5)	11 (3)	11 (3)	10 (2)	01 (1)
1111 (15)	111 (7)	111 (7)	111 (7)	111 (7)	11 (3)	11 (3)	11 (3)	11 (3)

Для каждой адресной последовательности в табл. 1 приведены двоичные значения адресов и в скобках – их десятичные эквиваленты. В дальнейшем модифицированную последовательность qA в соответствии с утверждением будем обозначать как qA_{M,a_i,a_j,\dots,a_q} , где индекс M указывает на метод формирования исходной адресной последовательности A_M , а остальные значения a_i, a_j, \dots, a_q – на биты, удаленные из A_M . Например, $2A_C = c_3c_2c_1$ будем обозначать как $2A_{C,c_0}$. Из табл. 1 видно, что двойные $2A_C$ и учетверенные $4A_C$ адресные последовательности зависят от выбранных разрядов исходной адресной последовательности A_C : $2A_{C,c_0} = c_3c_2c_1$, $2A_{C,c_1} = c_3c_2c_0$, $2A_{C,c_2} = c_3c_1c_0$, $4A_{C,c_1c_0} = c_3c_2$, $4A_{C,c_2c_0} = c_3c_1$, $4A_{C,c_2c_1} = c_3c_0$ и от перестановок разрядов в qA_C : $2A_C = c_1c_0c_3$, $4A_C = c_0c_3$. Соответственно, общее количество последовательностей qA ($q = 2^r$, $1 \leq r < m$), полученных из исходной последовательности A , определяется выражением

$$\binom{m}{m - \log_2 q} \cdot (m - \log_2 q)! = \binom{m}{m - r} \cdot (m - r)!, \quad 1 \leq r < m. \quad (1)$$

В случае двойной адресной последовательности $2A_C$ ($r = 1$) количество последовательностей, полученных только по одному из методов и только для одной исходной адресной последовательности A_C , согласно (1) равняется $m!$. Для $m = 4$ количество подобных последовательностей равняется 24, а для реальных значений m их число чрезвычайно велико. Следует подчеркнуть, что в качестве исходной последовательности A может быть использована любая адресная последовательность. Хорошо апробированы и применимы на практике такие последовательности, как пересчетные (счетчиковые), последовательности Грея, анти-Грея, последовательности с максимальной переключающей активностью, последовательности с заданным расстоянием Хэмминга, ЛП_r-последовательности, М-последовательности и ряд других [11–15]. Очевидно, что число подобных последовательностей, а также разнообразие алгоритмов для их формирования и свойств получаемых последовательностей велико.

Основная идея неразрушающих маршевых тестов заключается в применении адресных последовательностей pA , где p , в том числе и $p = q = 2^r$, – четное целое число, согласно которому при четном инвертировании содержимого ячейки ЗУ его значение останется прежним. В соответствии с этим простейшим свойством операции инвертирования построим базовый элемент неразрушающего маршевого теста с использованием четной адресной последовательности pA . Как и в классических неразрушающих тестах, маршевый элемент должен начинаться с операции чтения rb содержимого $b \in \{0, 1\}$ текущей ячейки ЗУ. Это необходимо для однозначных последующих действий с текущей ячейкой ЗУ, которые основываются на прочитанном значении ее содержимого. Следующей должна быть операция записи инверсного значения \bar{b} по отношению к только что полученному значению b , так как подобная операция является необхо-

димым условием активизации подавляющего большинства неисправностей ЗУ [4, 5]. За операцией записи следует операция чтения этой же текущей ячейки ЗУ для проверки правильности выполнения операции инвертирования ее содержимого. Использование четных адресных последовательностей pA обеспечивает повторное инвертирование каждой ячейки ЗУ, в итоге сохраняя его исходное состояние. Далее переходим к следующему запоминающему элементу ЗУ в соответствии с адресной последовательностью pA . Возрастающую четную (p) адресную последовательность pA обозначим как $p\uparrow$, а убывающую – как $p\downarrow$. Таким образом, базовый элемент будет иметь вид

$$p\uparrow(rb, w\bar{b}, rb). \quad (2)$$

Применение в базовом элементе четной адресной последовательности ($p\uparrow$) приведет к тому, что каждая ячейка ЗУ последовательно выполнит переход из b в \bar{b} и, наоборот, из \bar{b} в b , сохранив таким образом свое начальное значение. Количество таких пар переходов равняется $p/2$. Правильность выполнения обоих переходов (\uparrow и \downarrow), а также операций чтения нулевых и единичных значений обеспечивает вторая операция чтения rb базового элемента (2). Для иллюстрации реализации базового элемента (2) рассмотрим процедуру тестирования ЗУ, содержащего $N = 8$ запоминающих элементов с исходным содержимым 0 1 0 1 0 1 1 0. В качестве адресной последовательности используем последовательности двойных адресов отраженного кода Грея $2A_{Gg3} = g_2g_1g_0$ [16] и последовательности Соболя $2A_{Ss0} = s_3s_2s_1$ [17, 18] для исходного значения $m = 4$. Пошаговое изменение содержимого ЗУ при реализации базового элемента (2) для обоих случаев двойной адресации представлено в табл. 2.

Таблица 2. Процедура реализации базового элемента (2) неразрушающего теста для двух видов адресации
Table 2. The procedure for the implementation of the basic element (2) of transparent test for two types of addressing

Адреса ЗУ RAM addresses		7	6	5	4	3	2	1	0	Адреса ЗУ RAM addresses		7	6	5	4	3	2	1	0
Содержимое Contents		0	1	1	0	1	0	1	0	Содержимое Contents		0	1	1	0	1	0	1	0
$2A_{Gg3}$	000 (0)	0	1	1	0	1	0	1	\uparrow	$2A_{Ss0}$	100 (4)	0	1	1	\uparrow	1	0	0	0
	001 (1)	0	1	1	0	1	0	\downarrow	1		010 (2)	0	1	1	1	1	\uparrow	0	0
	011 (3)	0	1	1	0	\downarrow	0	0	1		110 (6)	0	\downarrow	1	1	1	1	0	0
	010 (2)	0	1	1	0	0	\uparrow	0	1		011 (3)	0	0	1	1	\downarrow	1	0	0
	110 (6)	0	\downarrow	1	0	0	1	0	1		111 (7)	\uparrow	0	1	1	0	1	0	0
	111 (7)	\uparrow	0	1	0	0	1	0	1		001 (1)	1	0	1	1	0	1	\uparrow	0
	101 (5)	1	0	\downarrow	0	0	1	0	1		101 (5)	1	0	\downarrow	1	0	1	1	0
	100 (4)	1	0	0	\uparrow	0	1	0	1		000 (0)	1	0	0	1	0	1	1	\uparrow
	100 (4)	1	0	0	\downarrow	0	1	0	1		100 (4)	1	0	0	\downarrow	0	1	1	1
	101 (5)	1	0	\uparrow	0	0	1	0	1		010 (2)	1	0	0	0	0	\downarrow	1	1
	111 (7)	\downarrow	0	1	0	0	1	0	1		110 (6)	1	\uparrow	0	0	0	0	1	1
	110 (6)	0	\uparrow	1	0	0	1	0	1		011 (3)	1	1	0	0	\uparrow	0	1	1
	010 (2)	0	1	1	0	0	\downarrow	0	1		111 (7)	\downarrow	1	0	0	1	0	1	1
011 (3)	0	1	1	0	\uparrow	0	0	1	001 (1)	0	1	0	0	1	0	\downarrow	1		
001 (1)	0	1	1	0	1	0	\uparrow	1	101 (5)	0	1	\uparrow	0	1	0	0	1		
000 (0)	0	1	1	0	1	0	1	\downarrow	000 (0)	0	1	1	0	1	0	0	\downarrow		
Содержимое Contents		0	1	1	0	1	0	1	0	Содержимое Contents		0	1	1	0	1	0	1	0

На каждом шаге реализации базового элемента только одна запоминающая ячейка, выделенная символом \uparrow или \downarrow , меняет свое состояние на противоположное. Из табл. 2 видно, что после выполнения базового элемента (2) начальное состояние ЗУ осталось неизменным.

Базовый элемент на основе четных адресных последовательностей позволяет синтезировать два неразрушающих маршевых теста:

$$\begin{aligned} \text{March_pA_1: } & \{\uparrow(rb); p\uparrow(rb, w\bar{b}, rb); \uparrow(rb)\}, (2N+3pN); \\ \text{March_pA_2: } & \{\downarrow(rb); p\uparrow(rb, w\bar{b}, rb); p\downarrow(rb, w\bar{b}, rb); \downarrow(rb)\}, (2N+6pN). \end{aligned} \quad (3)$$

В обоих тестах (3) произвольный порядок адресов для первой и последней фаз, включающих только операции чтения, должен быть одинаков: возрастающий \uparrow , как показано для $March_pA_1$, либо убывающий \downarrow , как показано для $March_pA_2$. Это связано с тем, что первая фаза тестов $March_pA_1$ и $March_pA_2$ используется для сжатия исходного состояния ЗУ и получения эталонной сигнатуры S_F , а их последняя фаза – для получения реального значения сигнатуры S_R после выполнения предыдущего либо предыдущих базовых элементов. В случае проявления неисправностей в ходе выполнения базовых элементов их наличие будет определяться выполнением неравенства $S_F \neq S_R$. Выражения в круглых скобках для каждого из представленных тестов (3) показывают их временную сложность в зависимости от кратности p повторения адресов.

Базовый элемент (2) обеспечивает активизацию и обнаружение всех простейших неисправностей типа SAF и TF . Операция записи $w\bar{b}$ и четная адресация pA реализуют как чтение нуля и единицы из текущей ячейки, так и выполнение двух переходов ее состояния, что является условием активизации данных неисправностей. Их обнаружение обеспечивает вторая операция чтения rb , результат которой сравнивается со значением, полученным при выполнении первой операции чтения rb базового элемента (2). При отсутствии указанных неисправностей эти значения должны быть противоположными, а их равенство свидетельствует о наличии неисправности и ее местоположении. Таким образом, для простейших неисправностей типа SAF и TF предложенный тест $March_pA_1$ в отличие от известных неразрушающих тестов [8, 9] имеет максимальную диагностическую способность.

Аналогично максимально возможная для маршевых тестов диагностическая способность тестов $March_pA_1$ и $March_pA_2$ достигается и для случая сложных кодочувствительных неисправностей $PNPSFk$. Выполнение базового элемента в случае $PNPSFk$ позволяет идентифицировать адрес базовой ячейки, которая не может выполнить один из переходов в этой ячейке для конкретного содержимого в соседних ячейках. В данном случае обнаруживаются только два их вида: $\langle u, u, u, \dots, u, \uparrow, u, u, u, \dots, u \rangle$ и $\langle d, d, d, \dots, d, \downarrow, d, d, d, \dots, d \rangle$, где $u, d \in \{0, 1\}$. Значения содержимого u и d соседних ячеек произвольны и зависят как от их начального состояния, так и от вида последовательности адресов pA . Например, для случая ЗУ с восемью ячейками и неисправностей $PNPSF3$ в ячейках с адресами 2, 4 и 5 тест $March_pA_1$ (3), применив адресную последовательность $2A_{S_{30}}$, позволяет обнаруживать неисправности $\langle 1, 1, \uparrow \rangle$, $\langle 0, 0, \downarrow \rangle$, $\langle 1, \uparrow, 0 \rangle$, $\langle 0, \downarrow, 1 \rangle$, $\langle \uparrow, 0, 1 \rangle$ и $\langle \downarrow, 1, 0 \rangle$ (см. табл. 2). Таким образом, однократное применение теста $March_2A_1$ позволяет достичь полноты покрытия для неисправностей $PNPSF3$, равной 25 %.

Для обнаружения неисправностей взаимного влияния необходимо выполнить анализ состояния ячейки жертвы после активизации неисправности, что невозможно в рамках базового элемента (2). Невыполнение одного из переходов в ячейке жертвы как результат наличия неисправности CF приведет к нечетному количеству изменений ее состояния на противоположное. Соответственно, конечное содержимое ЗУ перед выполнением последней фазы чтения в обоих тестах $March_2A_1$ и $March_2A_2$ ЗУ будет отличаться от его исходного состояния, что приведет к выполнению неравенства $S_F \neq S_R$. Количественно полнота покрытия тестом $March_2A_1$ таких неисправностей независимо от адресной последовательности всегда постоянна, как это видно, например, из экспериментальных данных для $CFid$ (табл. 3).

В табл. 3 показано, что независимо от адресной последовательности $2A_C$ полнота покрытия FC (*faults coverage*) неисправностей $CFid$ тестом $March_2A_1$ всегда равняется 50 %. Приведенные данные получены для ЗУ емкостью $N = 256$ бит с нулевым начальным состоянием. Адресные последовательности $2A_{c_0}, \dots, 2A_{c_7}, 2A_{c_8}$ сформированы из счетчиковой последовательности $A_C = c_8c_7c_6c_5c_4c_3c_2c_1c_0$ путем удаления соответствующего бита, например $2A_{c_8} = c_7c_6c_5c_4c_3c_2c_1c_0$.

Неразрушающий маршевый тест $March_pA_2$ отличается от $March_pA_1$ наличием второго базового элемента с обратным порядком ($p\downarrow$) адресов pA , что расширяет его возможности обнаруживать сложные неисправности, сохраняя эффективность $March_pA_1$ для более простых неисправностей. Действительно, для неисправностей $CFid$ тест $March_pA_2$ обнаруживает заметно большее их количество, вплоть до 100%-го покрытия (табл. 3).

Очевидно, что на полноту покрытия FC неисправностей $CFid$ тестом $March_{pA_2}$ в значительной степени влияет выбранная адресная последовательность pA . Для примера, приведенного в табл. 3, 100%-я полнота покрытия неисправностей $CFid$ тестом $March_{2A_2}$ достигается только для одной адресной последовательности $2A_{c8}$. Тест $March_{2A_2}$ с данной адресной последовательностью для произвольной ячейки-агрессора ЗУ, состоящего из 256 ячеек, в остальных ячейках формирует условия обнаружения всех возможных неисправностей CF .

Таблица 3. Полнота покрытия неисправностей $CFid$ тестами $March_{2A_1}$ и $March_{2A_2}$, %

Table 3. Faults coverage of $CFid$ faults by $March_{2A_1}$ and $March_{2A_2}$ tests, %

$CFid$	$March_{2A_1}$							$March_{2A_2}$						
	$2A_{c0}$	$2A_{c1}$	$2A_{c3}$	$2A_{c4}$	$2A_{c6}$	$2A_{c7}$	$2A_{c8}$	$2A_{c0}$	$2A_{c1}$	$2A_{c3}$	$2A_{c4}$	$2A_{c6}$	$2A_{c7}$	$2A_{c8}$
$\wedge(\uparrow;0)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,39	2,75	5,88	24,7	49,8	100
$\wedge(\uparrow;1)$	100	100	100	100	100	100	100	100	100	100	100	100	100	100
$\wedge(\downarrow;0)$	0,00	0,39	2,75	5,88	24,7	49,8	100	0,00	0,39	2,75	5,88	24,7	49,8	100
$\wedge(\downarrow;1)$	100	99,6	97,3	94,1	75,3	50,2	0,00	100	100	100	100	100	100	100
$\vee(\uparrow;0)$	0,00	0,39	2,75	5,88	24,7	49,8	100	0,00	0,39	2,75	5,88	24,7	49,8	100
$\vee(\uparrow;1)$	100	99,6	97,3	94,1	75,3	50,2	0,00	100	100	100	100	100	100	100
$\vee(\downarrow;0)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,39	2,75	5,88	24,7	49,8	100
$\vee(\downarrow;1)$	100	100	100	100	100	100	100	100	100	100	100	100	100	100
FC	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,2	51,4	52,9	62,4	75,0	100

В общем случае для нулевого начального состояния ЗУ ($b_i = 0$, где $i \in \{0, 1, \dots, N - 1\}$) условие обнаружения состоит в формировании тестом $March_{2A_2}$ четырех состояний в остальных ячейках ЗУ, каждая из которых потенциально может быть жертвой:

$$\begin{aligned}
 &0, 0, 0, \dots, 0, \uparrow, 1, 1, 1, \dots, 1; \\
 &1, 1, 1, \dots, 1, \downarrow, 0, 0, 0, \dots, 0; \\
 &1, 1, 1, \dots, 1, \uparrow, 0, 0, 0, \dots, 0; \\
 &0, 0, 0, \dots, 0, \downarrow, 1, 1, 1, \dots, 1.
 \end{aligned} \tag{4}$$

Как видно из соотношений (4), для любой ячейки ЗУ, выступающей в роли агрессора, выполняются оба перехода \uparrow и \downarrow для двух возможных состояний в остальных $N - 1$ ячейках. Соотношения (4) являются необходимым и достаточным условием обнаружения несвязных неисправностей взаимного влияния (CF), сформулированным при разработке теста $March C$ [4]. Анализ условий (4) показывает, что первые два соотношения реализуются первым базовым элементом теста $March_{2A_2}$, а два вторых – вторым базовым элементом, для которого используется обратный порядок адресной последовательности $2A$ (3).

Ранее отмечалось, что тест $March_{2A_1}$ либо первый базовый элемент теста $March_{2A_2}$ обнаруживают две неисправности $PNPSFk$: $\langle u, u, u, \dots, u, \uparrow, u, u, u, \dots, u \rangle$ и $\langle d, d, d, \dots, d, \downarrow, d, d, d, \dots, d \rangle$. Соответственно, второй базовый элемент будет обнаруживать дополнительно еще две неисправности $\langle u, u, u, \dots, u, \downarrow, u, u, u, \dots, u \rangle$ и $\langle d, d, d, \dots, d, \uparrow, d, d, d, \dots, d \rangle$ только в случае выполнения неравенства $u, u, u, \dots, u \neq d, d, d, \dots, d$ для произвольного k . Для нулевого начального состояния ЗУ последнее неравенство выполняется для любого k в случае реализации первым базовым элементом теста $March_{2A_2}$ двух первых соотношений (4).

Приведенный выше анализ позволяет сформулировать обобщенное требование к тесту $March_{2A_2}$ и, соответственно, к адресной последовательности $2A$, используемой при его реализации для ЗУ емкостью N бит с произвольным начальным состоянием.

Для достижения максимальной способности к обнаружению неисправностей CF и $PNPSFk$ с помощью теста $March_{2A_2}$ для любой произвольной ячейки ЗУ, действующей как агрессор

(в случае *CFid*) или как базовая ячейка (в случае *PNPSFk*), должно выполняться следующее условие для всех остальных $N - 1$ ячеек памяти:

$$\begin{aligned} & b, b, b, \dots, b, \uparrow, \bar{b}, \bar{b}, \bar{b}, \dots, \bar{b}; \\ & \bar{b}, \bar{b}, \bar{b}, \dots, \bar{b}, \downarrow, b, b, b, \dots, b. \end{aligned} \quad (5)$$

Символы \uparrow и \downarrow обозначают произвольную ячейку, в то время как случайные состояния $b \in \{0,1\}$ остальных ячеек упорядоченно изменяются на противоположные значения в соответствии с базовым элементом (2).

2. Количественные характеристики адресных последовательностей pA . Задача оценки максимальной покрывающей способности при тестировании ЗУ, зависящая от используемых адресных последовательностей, рассматривалась в рамках многократных маршевых тестов ЗУ [5, 11, 19, 20]. Исследовалась эффективность манхэттенского расстояния (*Manhattan distance*) и его взвешенных модификаций, таких как *Canberra distance* и *Bray Curtis distance*, для оценки адресных последовательностей, отвечающих определению 1 [5, 11]. Указанные характеристики использовались как мера различия между адресными последовательностями при формировании максимально отличающихся последовательностей адресов для последующих итераций теста ЗУ [5, 21, 22]. Было показано, что эффективность однократных маршевых тестов не зависит от вида адресной последовательности A , отвечающей определению 1 [4, 5, 23]. В то же время применение последовательностей с повторяющимися адресами pA для однократной реализации тестов (3) в значительной степени влияет на их эффективность (см. разд. 1).

Ранее уже отмечалось, что для четного p последовательность адресов pA состоит из всех возможных 2^m адресов $A(0), A(1), A(2), \dots, A(2^m-2), A(2^m-1)$, где $A(j) \in \{0, 1, 2, \dots, 2^m - 1\}, j \in \{0, 1, 2, \dots, 2^m - 1\}$, представляющих собой двоичные комбинации $a_{m-1}a_{m-2} \dots a_2a_1a_0$, каждая из которых генерируется p раз. Это означает, что в произвольной последовательности pA , длина которой равняется $p2^m$, p адресов $A(j)$ генерируются на определенном расстоянии друг от друга, определяемом алгоритмом формирования pA . Например, для адресной последовательности $2A_{C_{e1}}$ адреса $A(2) = 010$ генерируются на расстоянии 2 друг от друга, а те же два адреса $A(2)$ в последовательности $2A_{C_{e2}}$ генерируются на расстоянии 4 (см. табл. 1). Очевидно, что каждый адрес $A(j)$ произвольной последовательности pA имеет фиксированные расстояния между его повторными формированиями. Например,

Индекс адреса	0	1	...	α	...	β	...	$p2^m - 1$
Адрес				$A(j)$		$A(j)$		

Индексы адресов $\alpha < \beta < \dots < p2^m - 1$ в последовательности длиной $p2^m$ указывают местоположение всех повторяющихся p раз адресов $A(j)$. Предположим, что впервые адрес $A(j)$ формируется с индексом α в произвольной последовательности pA . Тогда расстояние между первым и вторым появлениями адреса $A(j)$ будет определяться разностью значений индексов α и β их позиций в последовательности pA , т. е. $D(A(j), pA) = \beta - \alpha = r$. Учитывая то, что период последовательности адресов pA равен $p2^m$, а расстояние между первым и вторым адресами $A(j)$ равно r , расстояние между вторым и первым адресами $A(j)$ будет равно $p2^m - r$. Расстояние между вторым и третьим формированиями адреса $A(j)$, как и для любой пары адресов $A(j)$, в последовательности pA определяется аналогичным образом.

В общем случае произвольная пара из возможных пар повторения адреса $A(j)$ в последовательности pA имеет два значения метрики расстояния $D(A(j), pA)$, а именно r и $p2^m - r$, которые не зависят от начального адреса с индексом 0 последовательности pA . Это следует из свойств циклических последовательностей и объясняется следующим образом. Учитывая то, что период последовательности адресов pA равен $p2^m$, а расстояние между первым и вторым адресами пары $A(j)$ равно r , расстояние между вторым и первым адресами $A(j)$ пары будет равно $p2^m - r$.

В предыдущем примере начальным адресом для последовательностей $2A_{Cc1}$ и $2A_{Cc2}$ был нулевой адрес $A(0) = 000$. Изменение начального адреса приводит к тому, что расстояние $D(A(2), 2A_{Cc1})$ равняется $r = 2$ либо $2^{m+1} - r = 2^4 - 2 = 14$. Расстояние $D(A(2), 2A_{Cc2})$ также принимает одно из двух значений: 4 или 12. Общее количество значений расстояния $D(A(j), pA)$ для четного p равняется $p(p - 1)$.

Для оценки последовательности pA в дальнейшем для каждого адреса $A(j)$ будем использовать минимальное значение расстояния $MD(A(j), pA)$ из всех возможных расстояний $D(A(j), pA)$. В качестве характеристики последовательности pA применим метрику среднего значения $AD(pA)$ минимальных расстояний $MD(A(j), pA)$ между повторяющимися адресами $A(j)$ последовательности pA , которая инвариантна относительно начального адреса. Эта метрика рассчитывается по формуле

$$AD(pA) = \frac{1}{2^m} \sum_{j=0}^{2^m-1} MD(A(j), pA). \quad (6)$$

Основываясь на том, что наиболее практичным случаем адресных последовательностей pA с повторением адресов, отвечающих определению 2, является тот, когда $p = 2$, введем для таких последовательностей характеристику, которая вычисляется следующим образом:

$$AD(2A) = \frac{1}{2^m} \sum_{j=0}^{2^m-1} \min [D(A(j), 2A), (2^{m+1} - D(A(j), 2A))]. \quad (7)$$

Тогда для примера $2A_{Cc1}$ среднее расстояние между одинаковыми адресами, рассчитанное согласно (7), будет равно двум и это значение будет повторяться для любых значений начальных адресов последовательности $2A_{Cc1}$. Более того, зачастую алгоритм формирования последовательностей pA позволяет аналитически вычислить как значение $MD(A(j), pA)$ для любого адреса $A(j)$, так и величину $AD(pA)$. Например, для адресной последовательности $2A_{Cc0}$ независимо от значения m величина $MD(A(j), 2A_{Cc0})$ для любого адреса $A(j)$ всегда равняется единице, как это видно из табл. 1 для $m = 3$. Аналогично можно заметить, что для произвольного m $MD(A(j), 2A_{Cc1}) = 2$, $MD(A(j), 2A_{Cc2}) = 4$ и $MD(A(j), 2A_{Ccm}) = 2^m$. Следовательно, при формировании адресных последовательностей $2A_{Cci}$, $i \in \{0, 1, 2, \dots, m\}$, согласно утверждению из счетчиковой последовательности адресов A_C , каждый из которых состоит из $m+1$ разрядов, среднее значение $AD(2A_{Cci})$ минимальных расстояний $MD(A(j), 2A_{Cci})$ будет вычисляться в соответствии с выражением

$$AD(2A_{Cci}) = \frac{1}{2^m} \sum_{j=0}^{2^m-1} MD(A(j), 2A_{Cci}) = 2^i, i \in \{0, 1, 2, \dots, m\}. \quad (8)$$

Следует отметить, что соотношение (8) справедливо для конкретного случая исходной адресной последовательности A_C и однозначно определенного алгоритма (см. утверждение) получения из нее последовательностей $2A_C$ для $p = 2$.

Приведем основные свойства метрик $D(A(j), pA)$, $MD(A(j), pA)$ и $AD(pA)$:

1. Минимальное значение $MD(A(j), pA)$ метрики $D(A(j), pA)$ не зависит от начального адреса последовательности pA , так как для любого начального адреса и любой пары повторяющихся адресов $A(j)$ оно принимает одно из двух значений: r или $p2^m - r$. Соответственно, значение метрики $AD(pA)$ также не зависит от начального адреса последовательности pA .

2. Для любой последовательности pA с периодом, равным $p2^m$, которая удовлетворяет определению 2, характеристики $D(A(j), pA)$, $MD(A(j), pA)$ и $AD(pA)$ принимают значения

$$1 \leq D(A(j), pA) \leq p2^m - (p - 1); 1 \leq MD(A(j), pA) \leq 2^m; 1 \leq AD(pA) \leq 2^m. \quad (9)$$

3. Максимальное значение метрики $AD(pA) = 2^m$ достигается только для случая, когда $MD(A(j), pA) = 2^m$ для всех $j \in \{0, 1, 2, \dots, 2^m - 1\}$, что следует из соотношения (7). Это свойство в первую очередь присуще последовательностям, сформированным согласно утверждению.

Максимальное значение минимума $MD(A(j), 2A)$ между двумя адресами $A(j)$ ($p = 2$) достигается для случая, когда выполняется равенство $r = 2^{m+1} - r$, из которого следует, что $MD(A(j), 2A) = 2^m$. Для произвольного четного p также $MD(A(j), pA) = 2^m$. При этом все p копии адреса $A(j)$ также расположены равноудаленно друг от друга, т. е. на расстоянии $(p2^m)/p = 2^m$ (см. рис. 1).

4. Для удовлетворяющих определению 2 последовательностей pA , адреса которых модифицированы путем инвертирования произвольного числа их бит $a_{m-1}a_{m-2} \dots a_2a_1a_0$, среднее значение $AD(pA)$ минимальных расстояний $MD(A(j), pA)$ между повторяющимися адресами $A(j)$ остается неизменным.

Свойство (4) следует из того, что применение операции отрицания для разрядов адресов $A(j)$ последовательности pA приводит к перераспределению их индексов и, соответственно, характеристик $MD(A(j), pA)$. Это означает, что в выражениях (6) и (7) для вычисления $AD(pA)$ используются те же слагаемые, только в другой последовательности. Например, для последовательности адресов $2A_{Gg3} = g_2g_1g_0$, приведенной в табл. 2, характеристика $AD(2A_{Gg3})$ (7) будет вычисляться как $(15 + 13 + 9 + 11 + 1 + 3 + 7 + 5)/8 = 8$. Между тем эта же характеристика для последовательности адресов $2A_{\overline{Gg3}} = \overline{g_2g_1g_0}$, полученной в результате применения операции отрицания для разрядов $g_2g_1g_0$ последовательности $2A_{Gg3}$, вычисляется как $(5 + 7 + 3 + 1 + 11 + 9 + 13 + 15)/8 = 8$.

Рассмотренные метрики $D(A(j), pA)$ и $AD(pA)$ для последовательностей pA позволяют сформулировать условие, аналогичное условию (5) максимальной эффективности теста $March_pA_2$ в терминах этих характеристик для $p = 2$.

Для общего случая адресных последовательностей pA условие максимальной эффективности теста $March_pA_2$, аналогичное условию (5) для $2A$, заключается в обеспечении наибольшего временного интервала между последовательными инвертированиями содержимого ячейки ЗУ с произвольным адресом $A(j)$. Это требование состоит в выборе адресных последовательностей с максимальным значением $AD(pA)$. Приведенное условие хорошо иллюстрируется на примере обнаружения тестом $March_2A_2$ неисправностей взаимного влияния $CFid$. Объединяя данные, приведенные в табл. 3, и численные значения метрики $AD(2A_{Cci})$ (8), можно заметить однозначную зависимость полноты покрытия неисправностей от величины данной интегральной метрики.

Таблица 4. Полнота покрытия $FC(CFid)$ неисправностей $CFid$ тестом $March_2A_2$, %

Table 4. Faults coverage of $FC(CFid)$ of $CFid$ failures by $March_2A_2$ test, %

$2A_{Cci}$	$2A_{c0}$	$2A_{c1}$	$2A_{c2}$	$2A_{c3}$	$2A_{c4}$	$2A_{c5}$	$2A_{c6}$	$2A_{c7}$	$2A_{c8}$
$FC(CFid)$	50,00	50,20	50,59	51,37	52,94	56,08	62,35	74,99	100,00
$AD(2A_{Cci})$	1	2	4	8	16	32	64	128	256

Действительно, для $p = 2$ только в том случае, если для любой ячейки памяти с адресом $A(j)$ расстояние $D(A(j), 2A)$ будет равно $2^m = 2^8 = 256$, в ячейках памяти сформируются обратные состояния, указанные в выражении (5), и будет достигнута максимальная полнота покрытия неисправностей $CFid$.

3. Многократные неразрушающие тесты $March_pA_1$ и $March_pA_2$. Идея неразрушающих тестов ЗУ предполагает многократное их применение, которое за счет изменяющихся начальных условий, чаще всего состояний ячеек ЗУ, позволяет обнаруживать любые неисправные состояния ЗУ [8–10]. В случае новых неразрушающих тестов $March_pA_1$ и $March_pA_2$ изменяющимися являются адресные последовательности pA , которые на каждой очередной итерации теста должны обеспечивать максимальное количество вновь обнаруживаемых неисправностей ЗУ по отношению к предыдущим итерациям. Аналогичная задача формирования

максимально различных тестовых наборов и их последовательностей решалась в рамках управляемых вероятностных тестов (*control random testing*) [24–29].

Сущность управляемых вероятностных тестов заключается в том, что очередная тестовая последовательность формируется в терминах предварительно определенных и обоснованных характеристик. В качестве этих характеристик применяются меры расстояний, причем чаще всего используются расстояния Хэмминга и Евклида [24–29]. В случае управляемых вероятностных тестов принимается гипотеза о том, что для двух тестовых последовательностей, имеющих минимальное различие, количество обнаруживаемых неисправностей будет минимальным и, наоборот, для максимально различных последовательностей обнаруживающая способность будет максимальной [24–29]. При этом и различие, и подобие тестовых последовательностей оценивались с использованием ряда интегральных метрик расстояния [24–27].

В работе [29] была показана эффективность применения расстояния Евклида для выбора следующих адресных последовательностей A_i , используемых для многократного тестирования ЗУ с помощью метода формирования адресных последовательностей, который основывается на инвертировании определенных бит исходной адресной последовательности A_k . Алгоритм формирования многократных тестов состоит в применении масок в виде двоичного вектора $\Lambda = \lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0 \neq 0 \ 0 \dots 0 \ 0$, единичные значения которого определяют наличие инверсий разрядов адресов $a_{m-1}a_{m-2}\dots a_2a_1a_0$ исходной базовой последовательности A_k по отношению к формируемой последовательности A_l [5].

Предположив, что исходная адресная последовательность A_k представляется двоичным кодом $a_{m-1}a_{m-2}\dots a_2a_1a_0$, выражение для адресов $A(j)_l \neq A(j)_k$ последовательности A_l будет иметь вид [5, 29]

$$A_l = a_{m-1}^{\lambda_{m-1}} a_{m-2}^{\lambda_{m-2}} \dots a_1^{\lambda_1} a_0^{\lambda_0} = (a_{m-1} \oplus \lambda_{m-1})(a_{m-2} \oplus \lambda_{m-2}) \dots (a_1 \oplus \lambda_1)(a_0 \oplus \lambda_0). \quad (10)$$

Отметим, что метод формирования адресных последовательностей (10) справедлив и для случая последовательностей с повторяющимися адресами pA . Например, для $m = 3$ и исходной последовательности $2A_k = 2A_{C_{C1}} = c_3c_2c_0 = \{000, 001, 000, 001, 010, 011, 010, 011, 100, 101, 100, 101, 110, 111, 110, 111\}$, применив вектор масок $\Lambda = \lambda_2\lambda_1\lambda_0 \neq 1 \ 1 \ 1$, на основании соотношения (10) получим новую последовательность адресов $2A_l = \{111, 110, 111, 110, 101, 100, 101, 100, 011, 010, 011, 010, 001, 000, 001, 000\}$.

Для адресных последовательностей pA_k и pA_l , соответствующих определению 2, справедлива следующая теорема, которая аналогична теореме относительно A_k и A_l , доказанной в работе [29].

Теорема. Для последовательностей адресов pA_k и pA_l , где pA_k включает $p2^m$ m -разрядных, сгенерированных произвольным образом двоичных адресов $a_{m-1}a_{m-2} \dots a_2a_1a_0$, каждый из которых повторяется p раз, а адреса последовательности pA_l получены согласно равенству (10) на основании вектора отрицаний $\lambda_{m-1}, \lambda_{m-2}, \dots, \lambda_1, \lambda_0$, для которого g значений $\lambda_\alpha, \lambda_\beta, \dots, \lambda_\gamma, \lambda_\delta$ ($\alpha > \beta > \dots > \gamma > \delta$) равняются единице, расстояние Евклида $ED(pA_k, pA_l)$ вычисляется следующим образом:

$$ED(pA_k, pA_l) = \sqrt{p2^m(2^{2\alpha} + 2^{2\beta} + \dots + 2^{2\gamma} + 2^{2\delta})}. \quad (11)$$

Действительно, для двух последовательностей адресов $2A_k = 2A_{C_{C1}} = c_3c_2c_0$ и $2A_l$, рассмотренных в предыдущем примере, расстояние Евклида соответствует значению, полученному из равенства (11): $ED(2A_k, 2A_l) = \sqrt{2 \times 2^3 \times (2^{2 \times 2} + 2^{2 \times 1} + 2^{2 \times 0})} = \sqrt{336}$.

В табл. 5 приведены численные значения расстояний $ED(2A_k, 2A_l)$, которые так же, как и в случае последовательностей адресов A_i [29], возрастают с ростом численного значения кода применяемой маски $\Lambda = \lambda_2\lambda_1\lambda_0$. Аналогичные значения расстояния Евклида и их различимость позволяют применять данную метрику для построения многократных неразрушающих тестов *March_pA_1* и *March_pA_2*.

Таблица 5. Значения расстояния Евклида для $m = 3$
 Table 5. Values of the Euclidean distance for $m = 3$

$2A_k = a_2 a_1 a_0$	$2A_l$						
	$a_2 a_1 a_0$	$a_2 a_1 a_0$	$a_2 a_1 a_0$	$a_2 a_1 a_0$	$a_2 a_1 a_0$	$a_2 a_1 a_0$	$a_2 a_1 a_0$
$ED(2A_k, 2A_l)$	$\sqrt{16}$	$\sqrt{64}$	$\sqrt{80}$	$\sqrt{256}$	$\sqrt{272}$	$\sqrt{320}$	$\sqrt{336}$

Используем методологию построения многократных тестов, основанную на применении масок отрицаний при формировании адресных последовательностей [5, 29], для последующих итераций тестов $March_pA_1$ и $March_pA_2$.

Для общего случая многократных неразрушающих тестов $March_pA_1$ и $March_pA_2$ применим оптимальное сочетание набора двоичных векторов масок $\Lambda = \lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0$ с целью получения адресных последовательностей $pA_0, pA_1, pA_2, \dots, pA_{w-1}$ тестов $March_pA_1$ и $March_pA_2$. В случае кратности тестов w , равной восьми, оптимальное сочетание масок для произвольного значения $m \geq 3$ приведено в табл. 6 [29], где значение индекса определяет оптимальную маску для последующей итерации многократного теста. Нулевое значение индекса означает однократное применение теста, для которого маска как таковая отсутствует, что эквивалентно ее нулевому значению. Для повторного применения теста маска, соответствующая наибольшей эффективности неразрушающего теста, принимает единичное значение для всех разрядов. Последующие итерации многократного теста основаны на применении приведенных в табл. 6 масок, оптимальность которых доказана в работе [29].

Таблица 6. Оптимальное сочетание масок $\Lambda = \lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0$ для $m \geq 3$
 Table 6. Optimal combination of masks $\Lambda = \lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0$ for $m \geq 3$

Индекс Index	Λ							
	λ_{m-1}	λ_{m-2}	λ_{m-3}	λ_{m-4}	λ_{m-5}	...	λ_1	λ_0
0	0	0	0	0	0		0	0
1	1	1	1	1	1	...	1	1
2	1	0	0	0	0	...	0	0
3	0	1	1	1	1	...	1	1
4	1	1	0	0	0	...	0	0
5	0	0	1	1	1	...	1	1
6	1	0	1	1	1	...	1	1
7	0	1	0	0	0	...	0	0

Для оценки эффективности многократного применения новых неразрушающих тестов $March_pA_1$ и $March_pA_2$ были получены экспериментальные результаты для адресных последовательностей $2A_{c0}, 2A_{c8}$ и $2A_{c4}$. Указанные адресные последовательности ранее использовались для получения полноты покрытия однократного применения данных тестов (см. табл. 3). В табл. 7 представлены результаты только двух итераций новых неразрушающих тестов, так как последующие их итерации с использованием оптимальных значений масок (см. табл. 6) не приводили к увеличению полноты покрытия неисправностей $CFid$.

Таблица 7. Полнота покрытия неисправностей $CFid$ двукратными тестами $March_2A_1$ и $March_2A_2$, %
 Table 7. Faults coverage of $CFid$ faults by double tests $March_2A_1$ and $March_2A_2$ in percentage, %

$CFid$	$March_2A_{c0_1}$		$March_2A_{c0_2}$		$March_2A_{c8_1}$		$March_2A_{c8_2}$		$March_2A_{c4_1}$		$March_2A_{c4_2}$	
	0	1	0	1	0	1	0	1	0	1	0	1
$\wedge(\uparrow;0)$	0,0	0,0	0,0	0,0	0,0	100	100	0,0	0,0	5,88	5,88	0,0
$\wedge(\uparrow;1)$	100	0,0	100	0,0	100	0,0	100	0,0	100	0,0	100	0,0
$\wedge(\downarrow;0)$	0,0	0,0	0,0	0,0	100	0,0	100	0,0	5,88	0,0	5,88	0,0
$\wedge(\downarrow;1)$	100	0,0	100	0,0	0,0	100	100	0,0	94,12	5,88	100	0,0
$\vee(\uparrow;0)$	0,0	0,0	0,0	0,0	100	0,0	100	0,0	5,88	0,0	5,88	0,0
$\vee(\uparrow;1)$	100	0,0	100	0,0	0,0	100	100	0,0	94,12	5,88	100	0,0
$\vee(\downarrow;0)$	0,0	0,0	0,0	0,0	0,0	100	100	0,0	0,0	5,88	5,88	0,0
$\vee(\downarrow;1)$	100	0,0	100	0,0	100	0,0	100	0,0	100	0,0	100	0,0
FC	50	0,0	50	0,0	50	50	100	0,0	50	2,94	52,94	0,0

Полученные результаты позволяют сделать следующие выводы:

1. Многократное тестирование на основании модификаций исходной адресной последовательности (10) с применением оптимальных сочетаний масок, приведенных в табл. 6, целесообразно только на базе теста *March_pA_2*. Данный факт объясняется тем, что второй базовый элемент теста *March_pA_2* использует обратный порядок адресов. Это эквивалентно применению единичной маски ($\Lambda = 1\ 1\ 1 \dots 1$) при повторном использовании теста *March_pA_1*. Последующие маски, соответствующие индексам 2 и 3, также являются инверсными по отношению друг к другу. В то же время сложность реализации двукратного теста *March_2A_1* равняется $16N$, что заметно больше сложности $14N$ однократного теста *March_2A_2* (3).

2. При многократном использовании теста *March_pA_2* максимизация характеристики $AD(pA)$ не приводит к увеличению покрывающей способности при повторных применениях теста с оптимальным сочетанием масок. Как видно из табл. 7, на всех последующих итерациях теста независимо от адресной последовательности $2A_{c0}$, $2A_{c8}$ и $2A_{c4}$ и, соответственно, значений ее характеристик $AD(2A)$ полнота покрытия не увеличивается. Однако, как было показано ранее для случая $p = 2$, максимальные значения характеристики $AD(2A)$ однозначно определяют максимальную покрывающую способность однократного теста *March_pA_2* для неисправностей *CFid* и других сложных неисправностей ЗУ.

Отметим, что для адресной последовательности $2A_{c0}$, для которой значение $AD(2A_{c0}) = 1$ минимально, последовательности $2A_{c8}$, имеющей максимальное значение $AD(2A_{c8}) = 256$, а также последовательности адресов $2A_{c4}$ с характеристикой $AD(2A_{c4}) = 16$ общим является равенство расстояния $MD(A(j), 2A)$ для всех значений j . Действительно, для любого $j \in \{0, 1, 2, \dots, 2^m - 1\}$ $MD(A(j), 2A_{c0}) = 1$, $MD(A(j), 2A_{c8}) = 256$ и $MD(A(j), 2A_{c4}) = 16$, как это ранее отмечалось при выводе соотношения (8).

Очевидно, что обнаружение сложных неисправностей ЗУ требует новых разнообразных состояний ЗУ для каждой итерации многократного теста *March_pA_2*. Остановимся на свойстве 4 метрик расстояния, рассмотренных в разд. 2. Согласно этому свойству применение операции отрицания для разрядов адресов $A(j)$ последовательности pA приводит к перераспределению их индексов j и, соответственно, характеристик $MD(A(j), pA)$, т. е. результатом операций отрицания является преобразование конкретного адреса $A(s)$ в адрес $A(t)$. При этом адрес $A(t)$ будет иметь характеристику $MD(A(t), pA)$, которую до преобразования имел адрес $A(s)$.

Таким образом, каждый адрес $A(s)$ последовательности адресов $2A$ в результате применения различных сочетаний отрицаний (10) принимает значения расстояния $MD(A(t), pA)$ от другого адреса $A(t)$. В случае когда все адреса $A(j)$ последовательности pA имеют одинаковые значения расстояний $MD(A(j), pA)$, применение отрицаний для получения новых адресных последовательностей, очевидно, не приведет к новому более разнообразному состоянию ячеек ЗУ. И, наоборот, если все адреса будут иметь собственные уникальные характеристики $MD(A(j), pA)$, перераспределяющиеся между адресами в результате применения выражения (10), произойдет формирование новых сочетаний состояний ячеек ЗУ. В результате создаются условия активизации и обнаружения сложных неисправностей ЗУ при многократном применении теста *March_pA_2*.

Введем новую характеристику $V(pA)$, которая численно равняется количеству отличающихся значений расстояния $MD(A(j), pA)$ адресов $A(j)$ последовательности pA . Для $p = 2$ характеристика $V(2A)$ принимает значения в диапазоне от 1 до 2^m . Например, для адресных последовательностей $2A_{Gg3}$ и $2A_{Ss0}$, приведенных в табл. 2, данная характеристика принимает значения $V(2A_{Gg3}) = 4$ и $V(2A_{Ss0}) = 1$. Значение характеристики $V(2A)$ для последовательностей адресов $2A_{c0}$ и $2A_{c8}$ (см. табл. 7) равняется единице. Это позволяет объяснить невысокую обнаруживающую способность многократного неразрушающего теста *March_pA_2* с их применением.

Для адресных последовательностей pA , свойства которых близки к свойствам случайных последовательностей, значения характеристик $AD(pA)$ и $V(pA)$ численно будут близки к своим средним значениям. Это свидетельствует об эффективности использования адресных последо-

вательностей как для однократного применения теста $March_pA_2$, так и для повторных применений с модифицированными адресами согласно выражению (10).

Для оценки эффективности многократного применения неразрушающего теста $March_2A_2$ были получены экспериментальные результаты для адресной последовательности $2A_{14}$. Она представляет собой псевдослучайную последовательность, полученную на основании порождающего полинома $\varphi(x) = 1 + x^4 + x^9$ путем удаления четвертого разряда, что обеспечивает повторное формирование каждого восьмиразрядного адреса, кроме адреса, состоящего из всех нулей. Свойства такой последовательности близки к свойствам случайных последовательностей и характеризуются средними значениями характеристик $AD(pA)$ и $V(pA)$. Как в экспериментах с последовательностями адресов $2A_{c0}$, $2A_{c8}$ и $2A_{c4}$, в данном случае использовались модификации адресов в соответствии с выражением (10) на основании масок, приведенных в табл. 6. Результаты эксперимента даны в табл. 8.

Таблица 8. Полнота покрытия неисправностей $CFid$ многократным тестом $March_2A_{14_2}$, %
Table 8. Faults coverage of $CFid$ faults by multiple test $March_2A_{14_2}$, %

$\begin{matrix} Index \\ CFid \end{matrix}$	0	1	2	3	4	5	6	7	FC
$\wedge(\uparrow;0)$	47,97	27,46	12,29	6,78	2,59	1,39	0,77	0,40	99,65
$\wedge(\uparrow;1)$	85,11	12,54	1,95	0,37	0,02	0,01	0,00	0,00	100
$\wedge(\downarrow;0)$	47,97	27,46	12,29	6,78	2,59	1,39	0,77	0,40	99,65
$\wedge(\downarrow;1)$	85,11	12,54	1,95	0,37	0,02	0,01	0,00	0,00	100
$\vee(\uparrow;0)$	52,17	23,26	13,17	5,90	2,66	1,33	0,81	0,35	99,65
$\vee(\uparrow;1)$	80,91	16,74	1,99	0,33	0,02	0,00	0,00	0,00	99,99
$\vee(\downarrow;0)$	52,17	23,26	13,17	5,90	2,66	1,33	0,81	0,35	99,65
$\vee(\downarrow;1)$	80,91	16,74	1,99	0,33	0,02	0,00	0,00	0,00	99,99
FC	66,54	20,00	7,35	3,34	1,33	0,68	0,40	0,19	99,83

Для каждой из неисправностей $CFid$ в зависимости от индекса повторного применения теста $March_2A_{14_2}$ приведены значения увеличения их полноты покрытия в процентах. Например, при однократном применении теста полнота покрытия неисправности $\wedge(\uparrow;0)$ составляет 47,97 % (табл. 8). Повторное применение теста увеличивает полноту покрытия на 27,46 %, а восьмикратное его использование обеспечивает суммарную полноту покрытия неисправности $\wedge(\uparrow;0)$ 99,65 %. В последней строке (FC) табл. 8 приведены значения полноты покрытия всех восьми неисправностей $CFid$.

Полученные данные подтверждают эффективность адресной последовательности $2A_{14}$ как при однократном применении теста $March_2A_{14_2}$, так и при многократном тестировании ЗУ. Данные табл. 7 и 8 свидетельствуют о зависимости результатов тестирования ЗУ от свойств применяемых адресных последовательностей, данные табл. 3 – об эффективности использования последовательностей $2A$ с максимальным значением расстояния $AD(2A)$ для однократных тестов $March_2A_1$ и $March_2A_2$. В свою очередь, анализ результатов, представленных в табл. 7 и 8, подтверждает гипотезу о применении последовательностей $2A$ с максимальным значением характеристики $V(2A)$ для обеспечения эффективности многократного теста $March_2A_2$.

Заключение. Предложенный подход к построению неразрушающих тестов ЗУ, основанный на применении адресных последовательностей с четным повторением адресов, показал свою высокую эффективность для случая $p = 2$. Двойное повторение адресов позволило уменьшить временную сложность предлагаемых неразрушающих тестов $March_2A_1$ и $March_2A_2$ по сравнению с известными методами и увеличить их диагностическую способность при сохранении высокой обнаруживающей способности неисправностей ЗУ. Исследована зависимость эффективности новых неразрушающих тестов $March_2A_1$ и $March_2A_2$ от применяемых адресных последовательностей и обоснован их оптимальный выбор для однократной и многократной реализации новых тестов. Интересным представляется дальнейшее исследование разработанных тестов $March_pA_1$ и $March_pA_2$ для случаев большей кратности p повторения адресов.

В первую очередь это касается сложных кодочувствительных неисправностей, так как многократное повторение адресов позволит обеспечить большее количество уникальных состояний ЗУ. Однако увеличение кратности повторения адресов существенно увеличивает временную сложность новых тестов, что оправдано только для случая сложных неисправностей современных ЗУ. Достаточно глубоким является вопрос формирования разнообразных адресных последовательностей с четным повторением адресов. Направления исследований в части генерирования адресных последовательностей с повторяющимися адресами и способы получения формальных методик оценки их свойств, влияющих на обнаруживающую способность неисправностей ЗУ, остаются практически открытыми.

Вклад авторов. В. Н. Ярмолик предложил идею неразрушающих тестов на базе многократно повторяющихся адресных последовательностей, В. А. Леванцевич и Д. И. Деменковец приняли участие в обобщении и анализе полученных результатов, И. Мрозек провел экспериментальные исследования.

Список использованных источников

1. Sharma, A. K. *Advanced Semiconductor Memories: Architectures, Designs, and Applications* / A. K. Sharma. – London : John Wiley & Sons, 2003. – 652 p.
2. The new hardware development trend and the challenges in data management and analysis / W. Pan [et al] // *Data Science and Engineering*. – 2018. – Vol. 6, no. 3. – P. 263–276.
3. Prince, B. *High-Performance Memories: New Architecture DRAM's and SRAM's, Evolution and Function* / B. Prince. – London : John Wiley & Sons, 1999. – 354 p.
4. Goor, A. J. *Testing Semiconductor Memories: Theory and Practice* / A. J. Goor. – Chichester, UK : John Wiley & Sons, 1991. – 536 p.
5. Ярмолик, С. В. Маршевые тесты для самотестирования ОЗУ / С. В. Ярмолик, А. П. Занкович, А. А. Иванюк. – Минск : Бестпринт, 2009. – 270 с.
6. Du, X. *New Memory BIST and Repair Methods* / X. Du. – Iowa : University of Iowa, 2004. – 276 p.
7. Chakraborty, A. *Fault-Tolerance and Reliability Techniques for High-Density Random-Access Memories* / A. Chakraborty, P. Mazumder. – Prentice Hall, 2002. – 448 p.
8. Nicolaidis, M. *Theory of transparent BIST for RAMs* / M. Nicolaidis // *IEEE Transactions on Computers*. – 1996. – Vol. 45, no. 10. – P. 1141–1156.
9. *Неразрушающее тестирование запоминающих устройств* / В. Н. Ярмолик [и др.]. – Минск : Бестпринт, 2005. – 230 с.
10. Ярмолик, В. Н. Обзор методов неразрушающего тестирования ОЗУ / В. Н. Ярмолик, А. П. Занкович // *Доклады БГУИР*. – 2005. – № 4(12). – С. 62–72.
11. Ярмолик, С. В. Многократные неразрушающие маршевые тесты с изменяемыми адресными последовательностями / С. В. Ярмолик, В. Н. Ярмолик // *Автоматика и телемеханика*. – 2007. – № 4. – С. 126–137.
12. Ярмолик, В. Н. Адресные последовательности для многократного тестирования ОЗУ / В. Н. Ярмолик, С. В. Ярмолик // *Информатика*. – 2014. – № 2(39). – С. 124–136.
13. Yarmolik, V. N. Modified gray and counter sequences for memory test address generation / V. N. Yarmolik, S. V. Yarmolik // *Proc. of the 13th Intern. Conf. MIXDES Design of Integrated Circuits and Systems, Gdynia, Poland, 22–24 June 2006*. – Gdynia, 2006. – P. 572–576.
14. Goor, A. J. Optimizing memory BIST Address Generator implementations / A. J. Goor, H. Kukner, S. Hamdioui // *Proc. of the 2011 6th Intern. Conf. on Design & Technology of Integrated Systems in Nanoscale Era (DTIS), Athens, Greece, 6–8 Apr. 2011*. – Athens, 2011. – P. 572–576.
15. Ярмолик, С. В. Квазислучайное тестирование вычислительных систем / С. В. Ярмолик, В. Н. Ярмолик // *Информатика*. – 2013. – № 3(39). – С. 92–103.
16. Savage, C. A survey of combinatorial Gray code / C. Savage // *SIAM Review*. – 1997. – Vol. 39, no. 4. – P. 605–629.
17. Chi, H. Computational investigations of quasi-random sequences in generating test cases for specification-based tests / H. Chi, E. I. Jones // *Proc. of the Winter Simulation Conf. WSC 2006, Monterey, CA, USA, 3–6 Dec. 2006*. – Monterey, 2006. – P. 975–980.
18. Chen, T. Y. Quasi-random testing / T. Y. Chen, R. Merkel // *IEEE Transaction on Reliability*. – 2007. – Vol. 56, no. 3. – P. 562–568.

19. Yarmolik, S. V. Address sequences and backgrounds with different hamming distance for multiple run March tests / S. V. Yarmolik // *IEEE Intern. J. of Applied Mathematics and Computer Science*. – 2008. – Vol. 18, no. 3. – P. 329–339.
20. Ярмолик, С. В. Анализ количественных характеристик различия при тестировании ОЗУ / С. В. Ярмолик, А. Н. Курбацкий, В. Н. Ярмолик // *Информатика*. – 2008. – № 3(19). – С. 90–98.
21. Thompson, A. C. *Minkowski Geometry* / A. C. Thompson. – Cambridge, N. Y., 1996. – 364 p.
22. Tubbs, J. D. Note on binary template matching / J. D. Tubbs // *Pattern Recognition*. – 1989. – Vol. 22, no. 4. – P. 359–365.
23. Sokol, B. Impact of the address changing on the detection of pattern sensitive faults / B. Sokol, I. Mrozek, V. N. Yarmolik // *Information Processing and Security Systems*. – London : Springer Science + Business Media, Inc., 2005. – P. 217–226.
24. Zhou, Z. Q. Using coverage information to guide test case selection in adaptive random testing / Z. Q. Zhou // *Proc. 34th IEEE Computer Soft and Applications Conf., Seoul, Korea, 19–23 July 2010*. – Seoul, 2010. – P. 208–213.
25. Chan, K. P. Good random testing / K. P. Chan, T. Y. Chen, D. Towey // *Proc. 9th Ada-Europe Intern. Conf. on Reliable Software Technologies (LNCS), Palma de Mallorca, Spain, 14–18 June 2004*. Palma de Mallorca, 2004. – P. 200–212.
26. Kuo, F. C. An in-depth study of mirror adaptive random testing / F. C. Kuo // *Proc. 14th European Conf. on Soft Quality, Los Alamitos, CA, USA*. – Los Alamitos, 2009. – P. 51–58.
27. Shiyi, Xu. Orderly random testing for both hardware and software / Xu Shiyi // *Proc. Pacific Rim Intern. Symp. on Dependable Computing, Taipei, Taiwan, 15–17 Dec. 2008*. – Taipei, 2008. – P. 160–167.
28. Ярмолик, С. В. Управляемое случайное тестирование / С. В. Ярмолик, В. Н. Ярмолик // *Информатика*. – 2011. – № 1(29). – С. 79–88.
29. Ярмолик, В. Н. Многократные управляемые вероятностные тесты / В. Н. Ярмолик, В. А. Леванцевич, И. Мрозек // *Информатика*. – 2015. – № 2(12). – С. 63–76.

References

1. Sharma A. K. *Advanced Semiconductor Memories: Architectures, Designs, and Applications*. London, John Wiley & Sons, 2003, 652 p.
2. Pan W., Li Z., Zhang Y., Weng C. The new hardware development trend and the challenges in data management and analysis. *Data Science and Engineering*, 2018, vol. 6, no. 3, pp. 263–276.
3. Prince B. *High-Performance Memories: New Architecture DRAM's and SRAM's, Evolution and Function*. London, John Wiley & Sons, 1999, 354 p.
4. Goor A. J. *Testing Semiconductor Memories: Theory and Practice*. Chichester, UK, John Wiley & Sons, 1991, 536 p.
5. Yarmolik S. V., Zankovich A. P., Ivaniuk A. A. Marshevye testu dlya samotestirovaniya OZU. *RAM Self-Test March Tests*. Minsk, Bestprint, 2009, 270 p. (In Russ.).
6. Du X. *New Memory BIST and Repair Methods*. Iowa, University of Iowa, 2004, 276 p.
7. Chakraborty A. *Fault-Tolerance and Reliability Techniques for High-Density Random-Access Memories*. Prentice Hall, 2002, 448 p.
8. Nicolaidis M. Theory of transparent BIST for RAMs. *IEEE Transactions on Computers*, 1996, vol. 45, no. 10, pp. 1141–1156.
9. Yarmolik V. N., Murashko I. A., Kummert A., Ivaniuk A. A. Nerazrushayushee testirovanie zapominayuschih ustroystv. *Transparent Memory Testing*. Minsk, Bestprint, 2005, 230 p. (In Russ.).
10. Yarmolik V. N., Zankovich A. P. Overview of transparent testing methods for RAM. *Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki [Reports of the Belarusian State University of Informatics and Radioelectronics]*, 2005, no. 4(12), pp. 62–72 (In Russ.).
11. Yarmolik S. V., Yarmolik V. N. Multiple non-destructive marching tests with variable address sequences. *Avtomatika i telemekhanika [Automation and Remote]*, 2007, no. 4, pp. 126–137 (In Russ.).
12. Yarmolik V. N., Yarmolik S. V. Address sequences for repeated testing of RAM. *Informatika [Informatics]*, 2014, no. 2(39), pp. 124–136 (In Russ.).
13. Yarmolik V. N., Yarmolik S. V. Modified gray and counter sequences for memory test address generation. *Proceedings of the 13th International Conference MIXDES Design of Integrated Circuits and Systems, Gdynia, Poland, 22–24 June 2006*. Gdynia, 2006, pp. 572–576.

14. Goor A. J., Kukner H., Hamdioui S. Optimizing memory BIST Address Generator implementations. *Proceedings of the 2011 6th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS), Athens, Greece, 6–8 April 2011*. Athens, 2011, pp. 572–576.
15. Yarmolik S. V., Yarmolik V. N. *Quasi-random testing of computing systems*. Informatika [Informatics], 2013, no. 3(39), pp. 92–103 (In Russ.).
16. Savage C. A survey of combinatorial Gray code. *SIAM Review*, 1997, vol. 39, no. 4, pp. 605–629.
17. Chi H., Jones E. I. Computational investigations of quasi-random sequences in generating test cases for specification-based tests. *Proceedings of the Winter Simulation Conference WSC 2006, Monterey, California, USA, 3–6 December 2006*. Monterey, 2006, pp. 975–980.
18. Chen T. Y., Merkel R. Quasi-random testing. *IEEE Transaction on Reliability*, 2007, vol. 56, no. 3, pp. 562–568.
19. Yarmolik S. V. Address sequences and backgrounds with different hamming distance for multiple run March tests. *IEEE International Journal of Applied Mathematics and Computer Science*, 2008, vol. 18, no. 3, pp. 329–339.
20. Yarmolik S. V., Kurbatski A. N., Yarmolik V. N. *Dissimilarity measure analysis for random access memory testing*. Informatika [Informatics], 2008, no. 3(19), pp. 90–98 (In Russ.).
21. Thompson A. C. *Minkowski Geometry*. Cambridge, New York, 1996, 364 p.
22. Tubbs J. D. Note on binary template matching. *Pattern Recognition*, 1989, vol. 22, no 4, pp. 359–365.
23. Sokol B., Mrozek I., Yarmolik V. N. Impact of the address changing on the detection of pattern sensitive faults. *Information Processing and Security Systems*, London, Springer Science + Business Media, Inc., 2005, pp. 217–226.
24. Zhou Z. Q. Using coverage information to guide test case selection in adaptive random testing. *Proceedings 34th IEEE Computer Soft and Applications Conference, Seoul, Korea, 19–23 July 2010*. Seoul, 2010, pp. 208–213.
25. Chan K. P., Chen T. Y., Towey D. Good random testing. *Proceedings 9th Ada-Europe International Conference on Reliable Software Technologies (LNCS), Palma de Mallorca, Spain, 14–18 June 2004*. Palma de Mallorca, 2004, pp. 200–212.
26. Kuo F. C. An in-depth study of mirror adaptive random testing. *Proceedings 14th European Conference on Soft Quality, Los Alamitos, CA, USA*. Los Alamitos, 2009, pp. 51–58.
27. Shiyi Xu. Orderly random testing for both hardware and software. *Proceedings Pacific Rim International Symposium on Dependable Computing, Taipei, Taiwan, 15-17 December 2008*. Taipei, 2008, pp. 160–167.
28. Yarmolik S. V., Yarmolik V. N. *Controlled random testing*. Informatika [Informatics], 2011, no. 1(29), pp. 79–88 (In Russ.).
29. Yarmolik V. N., Levantsevich V. A., Mrozek I. *Multiple controlled probability tests*. Informatika [Informatics], 2015, no. 2(12), pp. 63–76 (In Russ.).

Информация об авторах

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

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

Леванцевич Владимир Александрович, магистр технических наук, старший преподаватель, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: lvn@bsuir.by

Деменковец Денис Викторович, магистр технических наук, старший преподаватель, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: demenkovets@bsuir.by

Information about the authors

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

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

Vladimir A. Levantsevich, M. Sci. (Eng.), Senior Lecture, Belarusian State University of Informatics and Radioelectronics.
E-mail: lvn@bsuir.by

Denis V. Demenkovets, M. Sci. (Eng.), Senior Lecture, Belarusian State University of Informatics and Radioelectronics.
E-mail: demenkovets@bsuir.by

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING



УДК 519.816, 519.226, 519.254, 519.244.3
<https://doi.org/10.37661/1816-0301-2021-18-3-36-47>

Оригинальная статья
Original Paper

Последовательное статистическое принятие решений в задачах анализа потоков данных

А. Ю. Харин

Белорусский государственный университет,
пр. Независимости, 4, Минск, 220030, Беларусь
E-mail: KharinAY@bsu.by

Аннотация. В задачах анализа потоков данных актуальны проблемы статистического принятия решений о параметрах наблюдаемых потоков. Для их решения в работе предлагается использовать последовательные статистические решающие правила. Такие правила построены в статье для трех моделей потоков наблюдений: последовательности независимых однородных наблюдений; последовательности наблюдений, образующих временной ряд с трендом; последовательности зависимых наблюдений, образующих однородную цепь Маркова. Для каждого случая рассмотрена также ситуация, когда модель описывает наблюдаемые стохастические данные с искажениями. В качестве допустимых искажений используются «выбросы» («засорения»), которые адекватно описывают наиболее часто встречающиеся на практике ситуации. Предложены семейства последовательных решающих правил, в рамках которых строятся робастные решающие правила, позволяющие снизить влияние искажений на характеристики эффективности. Для иллюстрации преимуществ построенных решающих правил приводятся результаты компьютерных экспериментов.

Ключевые слова: последовательное решающее правило, статистический тест, временной ряд с трендом, однородная цепь Маркова, искажения

Благодарности. Результаты исследования получены при выполнении НИР «Библиотека процедур для компьютерного мониторинга данных и принятия решений на основе методов последовательного анализа» в рамках ГПНИ «Цифровые и космические технологии, безопасность человека, общества и государства», а также при поддержке стипендии Президента Республики Беларусь талантливым молодым ученым на 2021 г.

Для цитирования. Харин, А. Ю. Последовательное статистическое принятие решений в задачах анализа потоков данных / А. Ю. Харин // Информатика. – 2021. – Т. 18, № 3. – С. 36–47. <https://doi.org/10.37661/1816-0301-2021-18-3-36-47>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 02.08.2021
Подписана в печать | Accepted 01.09.2021
Опубликована | Published 29.09.2021

Sequential statistical decision making in problems of data flows analysis

Alexey Y. Kharin

Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220030, Belarus
E-mail: KharinAY@bsu.by

Abstract. In the problems of data flows analysis, the problems of statistical decision making on parameters of observed data flows are important. For their solution it is proposed to use sequential statistical decision rules. The rules are constructed for three models of observation flows: sequence of independent homogeneous observations; sequence of observations forming a time series with a trend; sequence of dependent observations forming a homogeneous Markov chain. For each case the situation is considered, where the model describes the observed stochastic data with a distortion. "Outliers" ("contamination") are used as the admissible distortions that adequately describe the majority of situations appear in practice. For such situations the families of sequential decision rules are proposed, and robust decision rules are constructed that allow to reduce influence of distortion to the efficiency characteristics. The results of computer experiments are given to illustrate the constructed decision rules.

Keywords: sequential decision rule, statistical test, time series with trend, homogeneous Markov chain, distortion

Acknowledgements. The results of the paper are obtained within the Project "Procedures library for computer monitoring of data and decision making on the basis of sequential analysis methods" in frames of the State research program "Digital and space technologies, human and state safety", and with the support of the grant of the President of the Republic of Belarus of talented young researchers for the year 2021.

For citation. Kharin A. Y. Sequential statistical decision making in problems of data flows analysis. *Informatics*, 2021, vol. 18, no. 3, pp. 36–47 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-36-47>

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

Введение. В современных прикладных задачах часто возникает необходимость анализа потоков стохастических данных с целью принятия решения об одном из двух режимов функционирования системы, порождающей такой поток [1] (например, «корректная работа» и «типовой сбой»). Режим функционирования характеризуется значением параметра (вектора параметров) распределения вероятностей, моделирующего наблюдения в потоке.

Для принятия решения в пользу одного из режимов в прикладных задачах эффективным подходом представляется последовательный статистический анализ [2]. В нем делается предположение о том, что число наблюдений, необходимых для принятия решения с заданной точностью (малыми значениями вероятностей ошибок), априори не фиксируется, а определяется в зависимости от самих стохастических наблюдений и является случайной величиной. Происходит «подстраивание» числа необходимых наблюдений по мере их поступления к сложности задачи принятия решения для конкретной ситуации, сформированной реальными наблюдениями. Другими словами, после каждого наблюдения имеются две возможности: принимать окончательное решение в пользу одного из двух возможных режимов либо принимать решение о том, что при полученных наблюдениях требуемая точность не может быть обеспечена и необходимо следующее наблюдение. Такая «гибкость» позволяет минимизировать среднее число наблюдений, необходимых для обеспечения заданной точности решений, однако создает значительные сложности теоретического анализа характеристик эффективности последовательных статистических решающих правил (условных вероятностей ошибок и математических ожиданий случайного числа необходимых наблюдений) [3].

Несмотря на трудности теоретического анализа, последовательные статистические решающие правила интенсивно используются в медицине, мониторинге режимов функционирования сложных технических объектов, контроле качества производимой продукции и во многих дру-

гих областях [1]. Эти правила представляют собой естественную схему вовлечения получаемой информации в процесс принятия решений по мере поступления наблюдений и позволяют существенно экономить число необходимых наблюдений, стоимость каждого из которых может быть значительной [2, 3].

Подход, разработанный в книге [4], дает возможность вычислять характеристики эффективности последовательных тестов не только в случае полного соответствия модели тем потокам данных, для описания которых она построена, но и при наличии искажений [5] – отклонений вероятностных законов, характеризующих наблюдения, от гипотетических модельных предположений.

В настоящей работе рассмотрены три модели потока стохастических данных: последовательность независимых одинаково распределенных случайных наблюдений; последовательность неоднородных наблюдений, образующих временной ряд с трендом, и последовательность зависимых наблюдений, образующих однородную цепь Маркова. Для каждой модели представлено последовательное решающее правило. Предложены семейства последовательных решающих правил, позволяющие строить робастные [6] решающие правила, устойчивые к влиянию искажений на характеристики эффективности.

Поток данных, представляющих независимые однородные наблюдения. Рассмотрим вначале простейшую модель потока данных. Пусть наблюдается поток $x_1, x_2, \dots \in U \subseteq \mathbf{R}^N$ независимых случайных векторов с распределением вероятностей P_θ , имеющим без ограничения общности плотность распределения вероятностей $p_\theta(x)$ относительно некоторой меры $\mu(x)$, $x \in U \subseteq \mathbf{R}^N$, где $\theta \in \Theta = \{0, 1\}$ – значение параметра, характеризующего поток; наблюдателю это значение неизвестно. (В случае дискретных распределений вероятностей здесь вместо привычной плотности распределения вероятностей относительно меры Лебега имеются в виду вероятности соответствующих событий.)

Сформулированы две гипотезы H_0 и H_1 о значении параметра, которые соответствуют двум режимам работы системы, порождающей поток случайных наблюдений:

$$H_0 : \theta = 0, H_1 : \theta = 1. \quad (1)$$

Последовательное решающее правило задается двумя компонентами: моментом остановки и терминальным решением. Первая компонента представляет собой функцию от случайных наблюдений и указывает номер наблюдения (случайный вследствие этой зависимости), после которого принимается решение о завершении процесса наблюдения. До момента остановки каждый раз принимается решение о продолжении этого процесса и необходимости получения следующего наблюдения, поскольку требуемая точность в контексте малых значений вероятностей возможных ошибок не может быть обеспечена совокупностью полученных наблюдений. В момент остановки принимается терминальное решение в пользу одной из гипотез (1) в соответствии с правилом, определяемым второй компонентой, на основании всех наблюдений, полученных к моменту остановки.

Рассмотрим семейство последовательных решающих правил $\delta_\lambda = (\tau_\lambda, d_\lambda)$, представленных в виде упорядоченных пар и основанных на функции $\lambda(\cdot) : U \rightarrow \mathbf{R}$. Здесь

$$\tau_\lambda = \inf \{n : \Lambda_n \notin (C_-, C_+)\} - \quad (2)$$

зависящий от наблюдений x_1, \dots, x_n случайный момент остановки процесса наблюдения (первый момент выхода из интервала между порогами). Терминальное решение $d_\lambda = i$, $i \in \{0, 1\}$, означает выбор гипотезы H_i в качестве результата в соответствии с правилом

$$d_\lambda = \mathbf{1}_{[C_+, +\infty)}(\Lambda_n). \quad (3)$$

В задании компонент (2), (3)

$$\Lambda_n = \Lambda_n(x_1, \dots, x_n) = \sum_{i=1}^n \lambda(x_i), \quad n \in \mathbf{N} = \{1, 2, 3, \dots\}, - \quad (4)$$

критериальная статистика; $C_-, C_+ \in \mathbf{R}$ – параметры (называемые порогами) последовательного решающего правила (2)–(4), $C_- < C_+$; $\mathbf{1}_A(\cdot)$ – индикаторная функция принадлежности аргумента множеству A .

Последовательное статистическое решающее правило, основанное на статистике отношения правдоподобия, является элементом рассмотренного семейства последовательных решающих правил (2)–(4) при следующей функции:

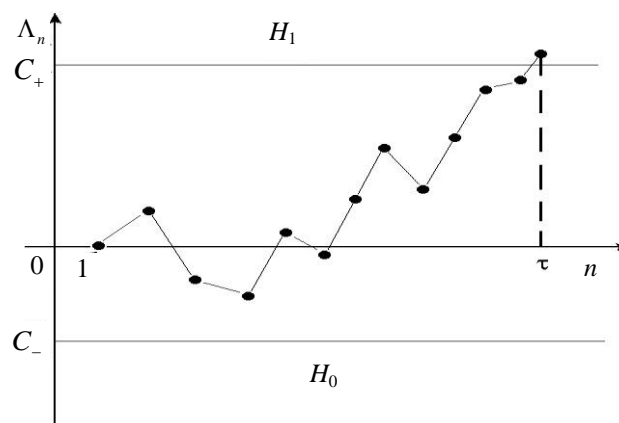
$$\lambda(u) = \lambda_w(u) = \log \frac{p_1(u)}{p_0(u)}, \quad u \in U. \quad (5)$$

На практике, как правило, значения порогов C_-, C_+ вычисляют по формулам [2]

$$C_- = \log \frac{\beta_0}{1-\alpha_0}, \quad C_+ = \log \frac{1-\beta_0}{\alpha_0}, \quad (6)$$

где $\alpha_0, \beta_0 \in (0, 1)$ – допустимые значения вероятностей ошибок первого рода (отвергается гипотеза H_0 ($d_\lambda = 1$) при условии, что она верна) и второго рода (отвергается гипотеза H_1 ($d_\lambda = 0$) при условии, что она верна) соответственно. Эти величины задаются пользователем решающего правила с учетом его представлений о допустимости соответствующих малых значений указанных характеристик. Логарифм берется по произвольному основанию. Важно лишь, чтобы при вычислении функции (5) и порогов (6) оно было одним и тем же.

Процесс изменения критериальной статистики в последовательном статистическом решающем правиле схематично изображен на рисунке.



Динамика критериальной статистики в процессе принятия решения в пользу H_1 последовательным правилом
 Test statistic dynamics in the decision making process
 in favor of H_1 with the sequential rule

Фактические значения вероятностей ошибок первого и второго рода для теста, основанного на функции $\lambda(\cdot)$, обозначим соответственно

$$\alpha = \alpha(\delta_\lambda) = E_0 \{P_0 \{d_\lambda = 1 | \tau_\lambda\}\}, \beta = \beta(\delta_\lambda) = E_1 \{P_1 \{d_\lambda = 0 | \tau_\lambda\}\}, \quad (7)$$

где $E_\theta \{\cdot\}$ – математическое ожидание (среднее значение случайного аргумента) по распределению P_θ . Обозначим условное математическое ожидание случайного числа необходимых наблюдений, когда справедлива гипотеза H_k , $k \in \Theta$:

$$t_k = t_k(\delta_\lambda) = E_k \{ \tau_\lambda \}. \quad (8)$$

В контексте данной работы рассматривается множество из четырех характеристик эффективности последовательного статистического решающего правила $\delta_\lambda = (\tau_\lambda, \delta_\lambda)$, определенных соотношениями (7), (8).

В работе [2] доказано, что для последовательного решающего правила (2)–(6) проверки гипотез (1) о параметре наблюдаемого потока независимых однородных наблюдений выполнены следующие неравенства относительно вероятностей ошибочных решений (7):

$$\alpha(\delta_{\lambda_w}) \leq \frac{\alpha_0}{1-\beta_0}, \beta(\delta_{\lambda_w}) \leq \frac{\beta_0}{1-\alpha_0}.$$

При этом сумма фактических вероятностей ошибочных решений ограничена суммой заданных значений:

$$\alpha(\delta_{\lambda_w}) + \beta(\delta_{\lambda_w}) \leq \alpha_0 + \beta_0.$$

Известно также, что для произвольных положительных значений α, β , в сумме меньших единицы, существуют значения порогов $C_-, C_+ \in \mathbf{R}$, для которых $\alpha(\delta_{\lambda_w}) = \alpha$, $\beta(\delta_{\lambda_w}) = \beta$, и при этом выполняется принцип оптимальности А. Вальда: для любого другого статистического решающего правила δ с фактическими вероятностями ошибочных решений, не превосходящими соответственно α и β , условные средние числа необходимых наблюдений не могут быть меньше, чем у δ_{λ_w} :

$$t_0(\delta_{\lambda_w}) \leq t_0(\delta), t_1(\delta_{\lambda_w}) \leq t_1(\delta).$$

Для условных математических ожиданий числа необходимых наблюдений известны приближенные выражения [14] (в предположении отличия от нуля знаменателей дробей в правых частях, $p_i(x) > 0$, $i = 0, 1$, а также $0 < \alpha_0, \beta_0 < 1$)

$$t_0(\delta_{\lambda_w}) \approx \frac{(1-\alpha_0) \log \frac{\beta_0}{1-\alpha_0} + \alpha_0 \log \frac{1-\beta_0}{\alpha_0}}{E_0 \left\{ \log \frac{p_1(x)}{p_0(x)} \right\}}, t_1(\delta_{\lambda_w}) \approx \frac{\beta_0 \log \frac{\beta_0}{1-\alpha_0} + (1-\beta_0) \log \frac{1-\beta_0}{\alpha_0}}{E_1 \left\{ \log \frac{p_1(x)}{p_0(x)} \right\}}.$$

Доказано [15], что последовательное решающее правило (2)–(6) с вероятностью единица завершается за конечное число наблюдений при выполнении условий, которые на практике интерпретируются как «различимость» соответствующих распределений вероятностей и, как правило, выполняются.

Фактические значения $t_0(\delta_{\lambda_w}), t_1(\delta_{\lambda_w})$ могут существенно отличаться от указанных выше приближений, а задаваемые значения α_0, β_0 – от фактических значений $\alpha(\delta_{\lambda_w}), \beta(\delta_{\lambda_w})$. Поэтому задача оценивания характеристик эффективности (7), (8) для последовательных решающих правил является важной [3]. Для решения этой задачи разработан подход [4], основанный на аппроксимации значений критериальной статистики последовательного решающего правила специальными цепями Маркова. Применим этот подход для простейшей модели независимых многомерных дискретных наблюдений [7], которая подвержена искажениям [8].

Пусть рассмотренная выше вероятностная модель потока наблюдений подвержена искажениям и наблюдения в потоке получены из смеси распределений вероятностей

$$\bar{P}_k(x) = (1 - \varepsilon_k)P_k(x) + \varepsilon_k\tilde{P}_k(x), \quad x \in U, \quad k \in \{0,1\}, \quad (9)$$

где $\varepsilon_k \in [0, \varepsilon_+]$ – вероятность появления «выброса», или «засорения». Как правило, ее значение неизвестно при анализе поступающего потока данных, максимальный уровень искажения ε_+ известен в конкретной решаемой задаче; $\tilde{P}_k(u)$ – произвольное «засоряющее» распределение вероятностей, $u \in U$, $\tilde{P}_k(\cdot) \neq P_k(\cdot)$. С целью снижения влияния искажений (9) на характеристики эффективности построим семейство последовательных решающих правил $\delta_g = (\tau_g, d_g)$:

$$\tau_g = \inf \left\{ n : \sum_{i=1}^n g(\lambda_w(x_i)) \notin (C_-, C_+) \right\}, \quad d_g = \mathbf{1}_{[C_+, +\infty)} \left(\sum_{i=1}^n g(\lambda_w(x_i)) \right), \quad (10)$$

где

$$g(z) = g_- \mathbf{1}_{(-\infty, g_-)}(z) + z \mathbf{1}_{[g_-, g_+]}(z) + g_+ \mathbf{1}_{(g_+, +\infty)}(z), \quad z \in \mathbf{R},$$

а $g_-, g_+ \in \mathbf{R}$ – дополнительные параметры последовательных решающих правил в рассматриваемом семействе, $g_- < g_+$.

Указанный выше подход при заданных значениях C_-, C_+, g_-, g_+ , «засоряющих» распределениях вероятностей \tilde{P}_k , $k \in \{0,1\}$, и вероятностях появления «искажений» позволяет вычислять характеристики эффективности $\tilde{\alpha}(\delta_g), \tilde{\beta}(\delta_g), \tilde{t}_0(\delta_g), \tilde{t}_1(\delta_g)$, определяемые аналогично выражениям (7), (8) для последовательных решающих правил (10) при наличии искажений (9).

Сформулируем критерий для построения робастного (устойчивого к искажениям (9)) последовательного решающего правила в рамках семейства (10):

$$\begin{cases} \sup_{\{\tilde{P}_k\}, \{\varepsilon_k\}} \left(w_0 \tilde{\alpha}(\delta_g; \tilde{P}_0, \varepsilon_0) + w_1 \tilde{\beta}(\delta_g; \tilde{P}_1, \varepsilon_1) \right) \rightarrow \min_{g_-, g_+}, \\ \tilde{t}_0(\delta_g) + \tilde{t}_1(\delta_g) \leq C \cdot \left(\tilde{t}_0(\delta_{\lambda_w}) + \tilde{t}_1(\delta_{\lambda_w}) \right), \end{cases} \quad (11)$$

где w_0, w_1 – задаваемые величины потерь, вызванных ошибками принятия решений первого и второго рода соответственно; C – коэффициент максимально допустимого увеличения среднего числа наблюдений. Алгоритмическая сложность численного решения задачи (11) составляет $O(K^2)$, где K – число состояний аппроксимирующей цепи Маркова, используемой в рас-

сма триваемом подходе для вычисления характеристик эффективности (7), (8) последовательного решающего правила.

Проиллюстрируем подход результатами компьютерных экспериментов для следующего частного случая:

$$x_i = (x_{i1}, \dots, x_{i5})', \quad x_{ii} \in A = \{1, 2, 3, 4\}, \quad U = \underbrace{A \times \dots \times A}_5, \quad |U| = 1024, \quad t = 1, 2, \dots$$

Нулевая гипотеза состоит в предположении дискретного равномерного распределения вероятностей на множестве U . Альтернативное предположение заключается в том, что при независимости компонент в первых трех из них имеется преобладание единиц над остальными тремя возможными значениями, четвертая и пятая компоненты распределены дискретно равномерно:

$$P_0(u) = \frac{1}{1024}, \quad u \in U;$$

$$P_1\{x_{ii} = 1\} = 0,4, \quad P_1\{x_{ii} = j\} = 0,2, \quad j \in \{2, 3, 4\}, \quad i \in \{1, 2, 3\},$$

$$P_1\{x_{ii} = j\} = \frac{1}{4}, \quad j \in A, \quad i \in \{4, 5\}.$$

Значения порогов C_-, C_+ выбраны в соответствии с выражениями (6), $C = 6, w_0 = w_1 = 1$. Последовательное решающее правило, основанное на функции (5), сравнивалось с построенным робастным решающим правилом при $g_- = -0,231, g_+ = 0,241$ в рамках искаженной модели, заданной следующим образом:

$$\varepsilon_- = \varepsilon_+ = 0,1,$$

$$\tilde{P}_0\{x_{ii} = 1\} = 1, \quad \tilde{P}_0\{x_{ii} \neq 1\} = 0, \quad i \in \{1, 2, 3\}, \quad \tilde{P}_0\{x_{ii} = j\} = \frac{1}{4}, \quad j \in A, \quad i \in \{4, 5\};$$

$$\tilde{P}_1\{x_{ii} = 3\} = 1, \quad \tilde{P}_1\{x_{ii} \neq 3\} = 0, \quad i \in \{1, 2, 3\}, \quad \tilde{P}_1\{x_{ii} = j\} = \frac{1}{4}, \quad j \in A, \quad i \in \{4, 5\}.$$

Результаты экспериментов (таблица) демонстрируют существенный выигрыш в малости вероятностей ошибок первого и второго рода для построенного робастного решающего правила (11) в сравнении с традиционным последовательным решающим правилом, полученный ценой приемлемого увеличения среднего числа необходимых наблюдений.

Результаты вычислительных экспериментов

Results of computer experiments

α_0	β_0	$\tilde{\alpha}$	$\tilde{\beta}$	$\tilde{\alpha}(\delta_g)$	$\tilde{\beta}(\delta_g)$	\tilde{t}_0	\tilde{t}_1	$\tilde{t}_0(\delta_g)$	$\tilde{t}_1(\delta_g)$
0,01	0,01	0,243	0,124	0,004	0,008	60,1	51,3	224,2	261,6
0,01	0,05	0,227	0,216	0,004	0,041	39,2	40,7	145,0	240,3
0,05	0,05	0,304	0,197	0,023	0,043	28,2	26,6	137,4	156,5
0,05	0,10	0,289	0,291	0,025	0,084	21,7	22,6	106,3	144,1
0,10	0,10	0,352	0,248	0,058	0,085	16,9	16,4	95,5	108,9

Поток данных, образующий временной ряд с трендом. Рассмотрим теперь более сложную модель – поток неоднородных стохастических наблюдений, образующих временной ряд с трендом. Пусть наблюдения $x_1, x_2, \dots \in \mathbf{R}$ имеют вид

$$x_t = \theta^T \psi(t) + \xi_t, \quad t \geq 1, \quad (12)$$

где $\psi(t) = (\psi_1(t), \psi_2(t), \dots, \psi_m(t))^T$, $t \geq 1$, – вектор заданных базисных функций тренда ($(\cdot)^T$ означает транспонирование); $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T \in \mathbf{R}^m$ – вектор коэффициентов (его истинное значение в процессе наблюдения неизвестно); $\{\xi_t, t \geq 1\}$ – последовательность независимых одинаково распределенных случайных величин (ошибок, погрешностей наблюдения). Обозначим плотности распределения вероятностей случайных наблюдений (12) через $\{p_n(x, \theta), x \in \mathbf{R}, n \geq 1\}$.

Рассмотрим две гипотезы относительно вектора параметров, определяющего тренд наблюдаемого потока данных:

$$H_0 : \theta = \theta^0, H_1 : \theta = \theta^1, \quad (13)$$

где $\theta^0, \theta^1 \in \mathbf{R}^m$ – заданные векторы, $\theta^0 \neq \theta^1$, m – число компонент параметра.

Обозначим статистику накопленного логарифмического отношения правдоподобия, вычисленную по n наблюдениям аналогично выражению (4):

$$\Lambda_n = \Lambda_n(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \lambda_i,$$

причем здесь логарифмическое отношение правдоподобия по наблюдению x_i вычисляется следующим образом: $\lambda_i = \ln(p_i(x_i, \theta^1) / p_i(x_i, \theta^0))$; $p_i(x, \theta) > 0$ – плотность распределения вероятностей случайного наблюдения номер i при условии, что истинное значение вектора параметров равно θ .

После $n = 1, 2, 3, \dots$ наблюдений решение принимается согласно правилу

$$d = \mathbf{1}_{[C_+, +\infty)}(\Lambda_n) + 2 \cdot \mathbf{1}_{(C_-, C_+)}(\Lambda_n), \quad (14)$$

где пороги C_- , C_+ являются параметрами и вычисляются в соответствии с формулами (6); $d = 2$ соответствует решению о том, что полученной в n наблюдениях информации недостаточно для обеспечения заданных малых уровней вероятностей ошибок первого и второго рода и требуется получить и обработать следующее наблюдение из потока; $d = 0$ и $d = 1$ означают остановку процесса наблюдения и принятие решения в пользу соответствующей гипотезы из (13).

При выполнении часто встречающегося предположения о распределении вероятностей погрешностей наблюдения $\xi_t \sim N_1(0, \sigma^2)$, $t = 1, 2, 3, \dots$, где $\sigma > 0$ – заданное среднее квадратическое отклонение, получаем

$$x_t \sim N(\theta^T \psi(t); \sigma^2), \quad t \geq 1, \quad p_t(x, \theta) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} (x - \theta^T \psi(t))^2\right\}.$$

Как следствие, логарифмическое отношение правдоподобия линейно по наблюдениям:

$$\lambda_t = \lambda_t(x_t) = -\frac{1}{2\sigma^2} \{2x_t(\theta^0 - \theta^1)^T \psi(t) + (\theta^1)^T \psi(t) \psi^T(t) \theta^1 - (\theta^0)^T \psi(t) \psi^T(t) \theta^0\}.$$

На практике часто гипотетическая модель потока наблюдений (12) оказывается подверженной «выбросам» [9]:

$$\bar{x}_t = \theta^T \psi(t) + \bar{\xi}_t, \quad t \geq 1, \quad (15)$$

где ошибки наблюдения образованы смесью $\bar{\xi}_t = (1 - \tau_t)\xi_t + \tau_t\tilde{\xi}_t, t \geq 1$, а $\{\tilde{\xi}_t, t \geq 1\}$ – последовательность независимых случайных величин, при этом $\{\tau_t, t \geq 1\}$ – независимые одинаково распределенные случайные величины с распределением Бернулли: $P(\tau_t = 0) = 1 - \delta, P(\tau_t = 1) = \delta$, $\tau_t, \xi_t, \tilde{\xi}_t$ независимы, а $\delta \in [0, 1/2)$ – уровень «засорения».

Для получения робастного решающего правила [10] при искажениях (15) построим семейство последовательных решающих правил, основанных на функции $f_{g_-}^{g_+}(\lambda_n)$ вместо λ_n в равенстве (14). Эту функцию по аналогии с потоком однородных наблюдений можно задать в виде

$$f_{g_-}^{g_+}(x) = g_- \cdot \mathbf{1}_{(-\infty, g_-]}(x) + x \cdot \mathbf{1}_{(g_-, g_+)}(x) + g_+ \cdot \mathbf{1}_{[g_+, +\infty)}(x)$$

или, для обеспечения непрерывности функции распределения приращений критериальной статистики, в виде

$$f_{g_-}^{g_+}(x) = \begin{cases} \frac{\varepsilon g_- + g_- - \varepsilon}{x}, & x \leq g_-; \\ x, & g_- < x < g_+; \\ -\frac{\varepsilon g_+}{x} + g_+ + \varepsilon, & x \geq g_+. \end{cases}$$

Оптимальные значения параметров g_-, g_+ определяются по критерию, аналогичному (11).

Поток зависимых наблюдений, образующих однородную цепь Маркова. Пусть теперь модель потока наблюдаемых данных допускает зависимость наблюдений и x_1, x_2, \dots образуют однородную цепь Маркова [11], принимая значения на множестве $V = \{0, 1, \dots, M-1\}$. Обозначим вектор вероятностей начальных состояний через $\pi = (\pi_i), i \in V$, а матрицу вероятностей одношаговых переходов – через $P = (p_{ij}), i, j \in V$:

$$P\{x_1 = i\} = \pi_i, \quad P\{x_n = j \mid x_{n-1} = i\} = p_{ij}, \quad i, j \in V.$$

Рассмотрим гипотезы, описывающие два типовых режима потока наблюдений в контексте параметров:

$$H_0: \pi = \pi^{(0)}, P = P^{(0)}; \quad H_1: \pi = \pi^{(1)}, P = P^{(1)}, \quad (16)$$

где $\pi^{(0)} = (\pi_i^{(0)})$, $\pi^{(1)} = (\pi_i^{(1)})$ – заданные значения вектора ненулевых вероятностей начальных состояний, $P^{(0)} = (p_{ij}^{(0)}) \neq P^{(1)} = (p_{ij}^{(1)})$ – заданные значения матриц ненулевых вероятностей одношаговых переходов для соответствующих гипотез.

При построении последовательного решающего правила проверки гипотез (16) обозначим для модели потока наблюдений статистики

$$\lambda_1 = \log \frac{\pi_{x_1}^{(1)}}{\pi_{x_1}^{(0)}}, \lambda_k = \log \frac{P_{x_{k-1}, x_k}^{(1)}}{P_{x_{k-1}, x_k}^{(0)}}, k > 1; \Lambda_n = \sum_{k=1}^n \lambda_k, n \in \mathbf{N}. \quad (17)$$

Последовательное решающее правило проверки гипотез (16) строится в соответствии с выражениями (14) и (17).

Исследуем теперь часто возникающую на практике ситуацию, когда гипотетическая модель, описывающая поток наблюдений, искажена, т. е. фактические значения параметров потока образованы смесью гипотетического и «засоряющего» значений (16) и порождают поток данных, отклоняющийся от постулируемого гипотетической моделью:

$$\bar{\pi}^{(k)} = (1 - \varepsilon)\pi^{(k)} + \varepsilon\tilde{\pi}^{(k)}, \bar{P}^{(k)} = (1 - \varepsilon)P^{(k)} + \varepsilon\tilde{P}^{(k)}, k = 0, 1, \quad (18)$$

где $\tilde{\pi}^{(k)}$ и $\tilde{P}^{(k)}$ – вектор вероятностей начальных состояний и матрица вероятностей одношаговых переходов для «засоряющей» цепи Маркова, $P^{(k)} \neq \tilde{P}^{(k)}$, $k = 0, 1$, а $\varepsilon \in [0, 1/2)$ – вероятность «засорения» (уровень искажения модели).

Построим семейство последовательных решающих правил вида (14), в котором вместо критериальной статистики Λ_n (17) будем использовать равенство

$$\Lambda_n^g(x_1, \dots, x_n) = g(\lambda_1, M) + \sum_{t=2}^n g(\lambda_t, x_{t-1}), n \in \mathbf{N}, \quad (19)$$

где $g: \mathbf{R} \times (V \cup \{M\}) \rightarrow \mathbf{R}$ – функция, зависящая от $2(M+1)$ дополнительных параметров $g_-(u)$, $g_+(u) \in \mathbf{R}$, $g_-(u) < g_+(u)$, $u \in V \cup \{M\}$:

$$g(y, u) = g_-(u) \cdot \mathbf{1}_{(-\infty, g_-(u)]}(y) + y \cdot \mathbf{1}_{(g_-(u), g_+(u))}(y) + g_+(u) \cdot \mathbf{1}_{[g_+(u), +\infty)}(y),$$

$$y \in \mathbf{R}, u \in V \cup \{M\}.$$

В рамках семейства (19) робастное последовательное решающее правило при наличии искажений (18) строится по критерию минимакса аналогично (11).

Заключение. В работе рассмотрены три модели потока стохастических наблюдений: независимые однородные наблюдения; неоднородные наблюдения, образующие временной ряд с трендом, и зависимые наблюдения, образующие однородную цепь Маркова. Для них построены последовательные решающие правила для различения двух заданных типовых ситуаций, описываемых в терминах параметров моделей. Отдельно рассмотрен случай, когда гипотетическая модель наблюдений подвержена искажениям, и разработан подход для построения робастных (устойчивых к искажениям) последовательных решающих правил.

Отметим, что предложенный подход позволяет вычислять характеристики эффективности и применим для построения робастных последовательных решающих правил в ситуации, когда типовая ситуация не описывается единственным возможным значением вектора параметров, а соответствует некоторому множеству значений [12]. Кроме того, для других моделей искажений (ошибки спецификации гипотетических значений параметров, уклонения фактических распределений вероятностей в рамках малых окрестностей в соответствующих метрических пространствах, искажения моделей зависимости и т. д.) разработанный в [4] подход для вычисления характеристик эффективности последовательных решающих правил также

применим и позволяет строить робастные последовательные решающие правила. Для модели временных рядов с трендом также удастся развить подход вычисления характеристик эффективности последовательных решающих правил на случай, когда каждый из типовых вариантов описывается единственным значением параметра, однако типовых вариантов больше чем два [13].

Представленные в работе последовательные решающие правила применимы также для решения задачи длительного мониторинга потоков данных с целью выявления смены режима функционирования наблюдаемой системы. В этом случае последовательное решающее правило применяется либо с перезапуском после принятия одного из двух решений, либо со сдвигом начала отсчета. При таких схемах принятия решений для обеспечения гарантий всей многоэтапной процедуры требуется дополнительное исследование вероятностей ошибочных решений.

Список использованных источников

1. Mukhopadhyay, N. *Sequential Methods and Their Applications* / N. Mukhopadhyay, B. de Silva. – N. Y. : Marcel Dekker, 2009. – 504 p.
2. Wald, A. *Sequential Analysis* / A. Wald. – N. Y. : John Wiley and Sons, 1947. – 212 p.
3. Lai, T. *Sequential analysis: Some classical problems and new challenges* / T. Lai // *Statistica Sinica*. – 2001. – Vol. 11. – P. 303–408.
4. Харин, А. Ю. Робастность байесовских и последовательных статистических решающих правил / А. Ю. Харин. – Минск : БГУ, 2013. – 207 с.
5. Huber, P. *Robust Statistics* / P. Huber, E. Ronchetti. – N. Y. : Wiley, 2009, 380 p.
6. Maevskii, V. V. *Robust regressive forecasting under functional distortions in a model* / V. V. Maevskii, Y. S. Kharin // *Automation and Remote Control*. – 2002. – Vol. 63, iss. 11. – P. 1803–1820.
7. Weiss, C. H. *Discrete-valued Time Series* / C. H. Weiss. – Oxford : John Wiley and Sons, 2018. – 284 p.
8. Kharin, A. Y. *Robust sequential test for hypotheses about discrete distributions in the presence of "outliers"* / A. Y. Kharin, D. V. Kishylau // *J. of Mathematical Sciences*. – 2015. – Vol. 205, iss. 1. – P. 68–73.
9. Kharin, A. *Performance and robustness analysis of sequential hypotheses testing for time series with trend* / A. Kharin, T. T. Tu // *Austrian J. of Statistics*. – 2017. – Vol. 46, no. 3–4. – P. 23–36.
10. Kharin, A. Y. *On error probabilities calculation for the truncated sequential probability ratio test* / A. Y. Kharin, T. T. Tu // *Журнал Бел. гос. ун-та. Математика. Информатика*. – 2018. – № 1. – P. 68–76.
11. Kemeny, J. G. *Finite Markov Chains* / J. G. Kemeny, J. L. Snell. – N. Y. : Springer, 1960. – 238 p.
12. Kharin, A. Y. *An approach to asymptotic robustness analysis of sequential tests for composite parametric hypotheses* / A. Y. Kharin // *J. of Mathematical Sciences*. – 2017. – Vol. 227, iss. 2. – P. 196–203.
13. Tu, T. T. *Sequential probability ratio test for many simple hypotheses on parameters of time series with trend* / T. T. Tu, A. Y. Kharin // *Журнал Бел. гос. ун-та. Математика. Информатика*. – 2019. – № 1. – P. 35–45.
14. Siegmund, D. *Sequential Analysis. Tests and Confidence Intervals* / D. Siegmund. – N. Y. : Springer-Verlag, 1985. – 272 p.
15. Ghosh, B. *Sequential Tests of Statistical Hypotheses* / B. Ghosh. – Reading : Addison – Wesley, 1970. – 454 p.

References

1. Mukhopadhyay N., Silva B. de. *Sequential Methods and Their Applications*. New York, Marcel Dekker, 2009, 504 p.
2. Wald A. *Sequential Analysis*. New York, John Wiley and Sons, 1947, 212 p.
3. Lai T. *Sequential analysis: Some classical problems and new challenges*. *Statistica Sinica*, 2001, vol. 11, pp. 303–408.
4. Kharin A. Y. *Robastnost bajesovskih i posledovatelnyh statisticheskikh reshayuschih pravil. Robustness of Bayesian and Sequential Statistical Decision Rules*. Minsk, Belorusskij gosudarstvennyj universitet, 2013, 207 p. (In Russ.).
5. Huber P., Ronchetti E. *Robust Statistics*. New York, Wiley, 2009, 380 p.
6. Maevskii V. V., Kharin Y. S. *Robust regressive forecasting under functional distortions in a model*. *Automation and Remote Control*, 2002, vol. 63, iss. 11, pp. 1803–1820.

7. Weiss C. H. *Discrete-valued Time Series*. Oxford, John Wiley and Sons, 2018, 284 p.
8. Kharin A. Y., Kishylau D. V. Robust sequential test for hypotheses about discrete distributions in the presence of "outliers". *Journal of Mathematical Sciences*, 2015, vol. 205, iss. 1, pp. 68–73.
9. Kharin A., Tu T. T. Performance and robustness analysis of sequential hypotheses testing for time series with trend. *Austrian Journal of Statistics*, 2017, vol. 46, no. 3–4, pp. 23–36.
10. Kharin A. Y., Tu T. T. *On error probabilities calculation for the truncated sequential probability ratio test*. Zhurnal Belorusskogo gosudarstvennogo universiteta. Matematika. Informatika [*Journal of the Belarusian State University. Mathematics and Informatics*], 2018, no. 1, pp. 68–76.
11. Kemeny J. G., Snell J. L. *Finite Markov Chains*. New York, Springer, 1960, 238 p.
12. Kharin A. Y. An approach to asymptotic robustness analysis of sequential tests for composite parametric hypotheses. *Journal of Mathematical Sciences*, 2017, vol. 227, iss. 2, pp. 196–203.
13. Tu T. T., Kharin A. Y. *Sequential probability ratio test for many simple hypotheses on parameters of time series with trend*. Zhurnal Belorusskogo gosudarstvennogo universiteta. Matematika. Informatika [*Journal of the Belarusian State University. Mathematics and Informatics*], 2019, no. 1, pp. 35–45.
14. Siegmund D. *Sequential Analysis. Tests and Confidence Intervals*. New York, Springer-Verlag, 1985, 272 p.
15. Ghosh B. *Sequential Tests of Statistical Hypotheses*. Reading, Addison – Wesley, 1970, 454 p.

Информация об авторе

Харин Алексей Юрьевич, доктор физико-математических наук, доцент, заведующий кафедрой теории вероятностей и математической статистики Белорусского государственного университета, главный научный сотрудник НИИ статистического анализа и моделирования Учреждения БГУ «НИИ прикладных проблем математики и информатики». ORCID ID: 0000-0002-5790-1956
E-mail: KharinAY@bsu.by

Information about the author

Alexey Y. Kharin, Dr. Sci. (Phys.-Math.), Associate Professor, Head of the Department of Probability Theory and Mathematical Statistics, Leading Researcher of the Research Laboratory of Statistical Analysis and Modeling of the Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University. ORCID ID: 0000-0002-5790-1956
E-mail: KharinAY@bsu.by



УДК 537.87:519.688
<https://doi.org/10.37661/1816-0301-2021-18-3-48-58>

Оригинальная статья
Original Paper

Аналитическое решение задачи экранирования низкочастотного магнитного поля тонкостенным цилиндрическим экраном в присутствии цилиндра

Г. Ч. Шушкевич

Гродненский государственный университет им. Янки Купалы,
ул. Ожешко, 22, Гродно, 230023, Беларусь
E-mail: gsys@grsu.by

Аннотация. Построено аналитическое решение граничной задачи, описывающей процесс проникновения низкочастотного магнитного поля через тонкостенный цилиндрический экран в присутствии цилиндра, с использованием приближенных граничных условий. Источником поля является тонкая нить бесконечно малой длины с бесконечно малым поперечным сечением, по которой циркулирует ток. Нить расположена в плоскости, перпендикулярной оси цилиндрического экрана, во внешней области по отношению к экрану. Первоначально потенциал исходного магнитного поля был представлен в виде сферических гармонических функций, затем с помощью теорем сложения, связывающих сферические и цилиндрические гармонические функции, стал представлен в виде суперпозиции цилиндрических гармонических функций. Вторичные потенциалы магнитного поля также представлены в виде суперпозиции цилиндрических гармонических функций в трехмерном пространстве. Показано, что решение поставленной граничной задачи сведено к решению системы линейных алгебраических уравнений относительно коэффициентов, входящих в представление вторичных полей. Численно исследовано влияние некоторых параметров задачи на значение коэффициента экранирования внешнего магнитного поля при прохождении через цилиндрический экран из меди в присутствии цилиндра. Результаты вычислений представлены в виде графиков. Полученные результаты могут быть использованы для экранирования технических устройств и биологических объектов от воздействия магнитных полей и обеспечения экологической среды вокруг работающих электроустановок и устройств.

Ключевые слова: граничная задача, потенциал, магнитное поле, тонкостенный цилиндрический экран, теоремы сложения, гармонические функции

Благодарности. Работа выполнялась в рамках подпрограммы «Математические модели и методы» ГПНИ «Конвергенция 2025».

Для цитирования. Шушкевич, Г. Ч. Аналитическое решение задачи экранирования низкочастотного магнитного поля тонкостенным цилиндрическим экраном в присутствии цилиндра / Г. Ч. Шушкевич // Информатика. – 2021. – Т. 18, № 3. – С. 48–58. <https://doi.org/10.37661/1816-0301-2021-18-3-48-58>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 16.06.2021
Подписана в печать | Accepted 19.07.2021
Опубликована | Published 29.09.2021

Analytical solution of problem of shielding low-frequency magnetic field by thin-walled cylindrical screen in presence of cylinder

Gennady Ch. Shushkevich

*Yanka Kupala State University of Grodno,
st. Ozheshko, 22, Grodno, 230023, Belarus
E-mail: gsys@grsu.by*

Abstract. The analytical solution of boundary value problem describing the process of penetration of low-frequency magnetic field through thin-walled cylindrical screen with cylindrical inclusion is constructed by use of approximate boundary conditions. The source of the field is a thin thread of infinitely small length with an infinitely small cross-section where current circulates. Thread is located in a plane which is perpendicular to axis of cylindrical screen, in outer region with respect to a screen. Initially the potential of initial magnetic field is represented as spherical harmonic functions, then using addition theorems connecting spherical and cylindrical harmonic functions, it became as cylindrical harmonic functions superposition. Secondary potentials of magnetic field are also presented as superposition of cylindrical harmonic functions in three-dimensional space. It is shown that the solution of formulated boundary value problem is reduced to the solution of linear algebraic equations system for coefficients included in the representation of secondary fields. The influence of some aspects of the problem on the value of the screening coefficient of an external magnetic field when passing through a cylindrical copper screen in the presence of a cylindrical inclusion is studied numerically. Calculation results are presented in graphs form. Obtained results can be used to shield technical devices and biological objects against the effects of magnetic fields to provide ecological surrounding of operating electrical installations and devices.

Keywords: boundary value problem, potential, magnetic field, thin-walled cylindrical screen, addition theorems, harmonic functions

Acknowledgements. The work was carried out within the framework of the "Mathematical Models and Methods" subprogram of the State Pedagogical Research Institute "Convergence 2025".

For citation. Shushkevich G. Ch. Analytical solution of problem of shielding low-frequency magnetic field by thin-walled cylindrical screen in presence of cylinder. *Informatics*, 2021, vol. 18, no. 3, pp. 48–58 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-48-58>

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

Введение. В условиях широкого использования во всех сферах человеческой деятельности электротехнических, электронных и радиоэлектронных приборов и оборудования актуальной задачей является формирование электромагнитной обстановки. Электромагнитная обстановка представляет собой совокупность электромагнитных полей в заданной области пространства, которая может влиять как на биологические объекты, так и на конкретные электронные устройства. Для создания благоприятной электромагнитной обстановки осуществляется электромагнитное экранирование^{1,2} [1].

Источниками низкочастотных электромагнитных излучений являются линии электропередач, силовые и трансформаторные подстанции, приборы и электроустановки на предприятиях, в научно-исследовательских лабораториях, лечебных и профилактических заведениях [2].

Универсальными методами расчета электростатических, магнитных и электромагнитных полей являются численные методы: конечных разностей, конечных элементов, интегральных уравнений³ [3–5]. Между тем актуальность разработки новых аналитических и численно-

¹Electromagnetic Shielding [Electronic resource]. – Mode of access: https://en.wikipedia.org/wiki/Electromagnetic_shielding. – Date of access: 10.04.2021.

²Practical EM Shielding [Electronic resource]. – Mode of access: <https://learnemc.com/practical-em-shielding>. – Date of access: 10.04.2021.

³Modeling and Optimizing Electromagnetic Shielding with COMSOL Multiphysics [Electronic resource]. – Mode of access: <https://www.simtcsolution.fr/en/client-cases/security-strength-of-materials/modeling-and-optimizing-electromagnetic-shielding-with-comsol-multiphysics>. – Date of access: 10.04.2021.

аналитических методов решения краевых задач математической физики не уменьшилась и в наши дни. Аналитические методы по-прежнему остаются главными средствами решения фундаментальных проблем, создают основу для тестирования решений краевых задач, полученных численными методами.

Метод разделения переменных и аппарат функций комплексного переменного наиболее часто используются для аналитического решения граничных задач математической физики [5]. При решении граничных задач математической физики для многосвязных областей успешно применяется метод теорем сложения⁴ [6]. Для моделирования процесса проникновения магнитных полей через тонкостенные экраны используют усредненные граничные условия, которые связывают магнитное поле по обе стороны тонкостенного проводящего экрана и исключают рассмотрение поля внутри экрана [7–11]. В работе [12] численно показано, что усредненные граничные условия достаточно точно описывают процесс проникновения поля через тонкостенный цилиндрический экран в широком диапазоне частот (плоскопараллельная задача).

В статье [13] разработана математическая модель экранирования магнитного поля цилиндрическим тонкостенным экраном, выполненным из материала с магнитной проницаемостью, которая нелинейно зависит от напряженности магнитного поля. Трехмерная математическая модель сплошного цилиндрического экрана для защиты от действия переменного электромагнитного поля построена в программном продукте ANSYS NFSS [14].

В настоящей статье представлена методика аналитико-численного решения граничной трехмерной задачи, описывающей процесс проникновения низкочастотного магнитного поля через тонкостенный цилиндрический экран с цилиндрическим включением, с использованием теории сферических, цилиндрических функций и Python-библиотек Numpy, Scipy, Matplotlib [15].

Постановка задачи. Пусть в пространстве R^3 находится тонкостенный цилиндрический экран Γ толщиной Δ и бесконечный круговой металлический цилиндр радиуса b , ограниченный поверхностью S . Область пространства между экраном Γ и поверхностью S обозначим через D_1 , внешнюю область по отношению к экрану Γ – D_2 , область оболочки Γ – D . Плоскостное сечение геометрии задачи показано на рис. 1.

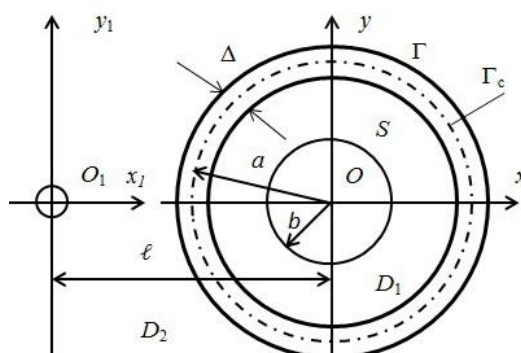


Рис. 1. Сечение цилиндрического экрана

Fig. 1. Section of cylindrical screen

Тонкостенный экран Γ выполнен из материала с тремя электромагнитными параметрами: ε – диэлектрическая проницаемость, μ – абсолютная магнитная проницаемость, γ – удельная электрическая проводимость. Область D_m , $m = 1, 2$, заполнена средой с магнитной проницаемостью μ_m .

⁴Ерофеевко, В. Т. Метод теорем сложения и теория усредненных граничных условий в краевых задачах электродинамики : автореф. дис. ... д-ра физ.-мат. наук : 01.04.03 / В. Т. Ерофеевко, БГУ. – Минск, 1993. – 29 с.

В точках O, O_1 введем декартовы координаты с одинаково направленными осями координат и соответствующие цилиндрические координаты $\{\rho, z, \varphi\}$ в точке O (в точке O_1 – аналогично):

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi, \quad z = z,$$

где $0 \leq \rho < \infty, -\pi \leq \varphi \leq \pi, -\infty < z < \infty$. Точка O_1 лежит на оси Ox_1 , расстояние между точками O, O_1 обозначим через ℓ .

На плоскости $O_1x_1y_1$ расположен источник поля – тонкая нить с бесконечно малым поперечным сечением, по которой циркулирует ток I . В результате взаимодействия первичного магнитного поля с экраном Γ образуются вторичные магнитные поля. Обозначим потенциал вторичного магнитного поля в области D_2 через U_2 , в области D_1 – через U_1 , потенциал исходного магнитного поля – через U_0 . Реальный магнитный потенциал и вектор напряженности магнитного поля определяются формулами

$$U_m = \operatorname{Re}(e^{-i\omega t} U_m), \quad \vec{H}_m = -\operatorname{grad} U_m, \quad m = 0, 1, 2.$$

В случае тонкостенных экранов магнитное поле внутри экрана не исследуется, а сшивается с помощью специальных граничных условий на срединной поверхности экрана Γ_c .

Срединная поверхность экрана Γ_c и поверхность S в цилиндрической системе координат описываются следующим образом:

$$\Gamma_c = \{\rho = a, -\pi \leq \varphi \leq \pi, -\infty < z < \infty\}, \quad S = \{\rho = b, 0 \leq \varphi \leq 2\pi, -\infty < z < \infty\}.$$

Постановка задачи. Требуется найти скалярные магнитные потенциалы U_m в области $D_m, m = 1, 2$, которые удовлетворяют:

– уравнению Лапласа

$$\Delta U_m \equiv \frac{\partial^2}{\partial x^2} U_m + \frac{\partial^2}{\partial y^2} U_m + \frac{\partial^2}{\partial z^2} U_m = 0;$$

– граничным условиям на срединной поверхности Γ_c [16]

$$\mu_2 \frac{\partial(U_0 + U_2)}{\partial \rho} - \mu_1 \frac{\partial U_1}{\partial \rho} = -pF(U_0 + U_1 + U_2), \quad (1)$$

$$\mu_2 \frac{\partial(U_0 + U_2)}{\partial \rho} + \mu_1 \frac{\partial U_1}{\partial \rho} = qF(U_0 + U_2 - U_1), \quad (2)$$

где $p = \mu\delta/2, q = 2/i\omega\delta\gamma, \delta = 2\operatorname{tg}(k\Delta/2)/k, k = \omega\sqrt{i\omega\mu\gamma}, 0 \leq \arg k < \pi;$

– граничному условию на поверхности S

$$\frac{\partial}{\partial \rho} U_1 = 0, \quad \rho = b; \quad (3)$$

– условию на бесконечности

$$U_2(M) \rightarrow 0, \quad M \rightarrow \infty, \quad (4)$$

где M – произвольная точка пространства.

Граничный оператор в цилиндрических координатах имеет следующий вид:

$$F(u) = (\vec{n}, \text{rot}[\vec{n}, \text{grad} u]) = \frac{1}{\rho^2} \frac{\partial^2 U}{\partial \varphi^2} + \frac{\partial^2 U}{\partial z^2}. \quad (5)$$

Решение задачи. Потенциал исходного магнитного поля можно представить равенством [17]

$$U_0 = P \left(\frac{\ell}{r_1} \right)^2 P_1(\cos \theta_1), \quad P = \frac{M_z}{4\pi \ell^2}, \quad (6)$$

где M_z – магнитный момент, $P_n(x)$ – полиномы Лежандра [18].

Далее воспользуемся следующими теоремами сложения (формулами), которые связывают сферические и цилиндрические гармонические функции в системе координат с началом в точке O_1 и цилиндрические гармонические функции в системе координат с началом в точках O и O_1 [19]:

$$r_1^{-n-1} P_n^m(\cos \theta_1) e^{im\varphi} = \frac{(-j)^{n+m}}{\pi(n-m)!} \int_{-\infty}^{\infty} \lambda^n \tilde{K}_n(\lambda \rho_1) e^{i(\lambda z_1 + m\varphi_1)} d\lambda, \quad \rho_1 > 0,$$

$$\tilde{K}_n(\lambda \rho_1) e^{i(\lambda z_1 + m\varphi_1)} = \sum_{k=-\infty}^{\infty} (-1)^k \tilde{K}_{m-n}(\lambda \ell) e^{i(m-k)\varphi_0} I_n(\lambda \rho) e^{i\lambda z + ik\varphi} d\lambda, \quad \rho < \ell,$$

где $\{\ell, 0, \varphi_0\}$ – цилиндрические координаты точки O в системе координат с началом в точке O_1 .

На основании приведенных выше формул ($\varphi_0 = 0$) представим потенциал исходного магнитного поля в цилиндрической системе координат с началом в точке O :

$$U_0(\rho, z, \varphi) = P \sum_{n=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g_n(\lambda) I_n(\lambda \rho) e^{i\lambda z} d\lambda \right] e^{in\varphi}, \quad (7)$$

где $g_n(\lambda) = i(-1)^{n+1} \frac{\ell^2}{\pi} \lambda \tilde{K}_n(\lambda \ell)$.

В связи с отсутствием осевой симметрии в представлении исходного магнитного поля потенциалы вторичного магнитного поля U_m , $m = 1, 2$, запишем в виде рядов Фурье так, чтобы выполнялись условия на бесконечности (4):

$$U_1(\rho, z, \varphi) = P \sum_{n=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} (X_n(\lambda) \tilde{K}_n(\lambda \rho) + Y_n(\lambda) I_n(\lambda \rho)) e^{i\lambda z} d\lambda \right] e^{in\varphi}, \quad (8)$$

$$U_2(\rho, z, \varphi) = P \sum_{n=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} Z_n(\lambda) \tilde{K}_n(\lambda \rho) e^{i\lambda z} d\lambda \right] e^{in\varphi}, \quad (9)$$

где $\tilde{K}_n(x) = (\text{sgn})^n K_n(|x|)$, $I_n(x)$ – модифицированная функция Бесселя первого рода, $K_n(x)$ – функция Макдональда [18].

Физическая размерность потенциалов в системе СИ – А.

Неизвестные функции $X_n(\lambda)$, $Y_n(\lambda)$, $Z_n(\lambda)$ подлежат определению из граничных условий (1)–(3).

Вычислим частные производные $\frac{\partial}{\partial \rho} U_0(\rho, z, \varphi)$, $\frac{\partial}{\partial \rho} U_1(\rho, z, \varphi)$, $\frac{\partial}{\partial \rho} U_2(\rho, z, \varphi)$, принимая во

внимание представления (7)–(9), и выполним граничные условия (1)–(3), учитывая граничный оператор (5). На основании ортогональности функций $\exp(in\varphi)$ на отрезке $[-\pi, \pi]$ и преобразования Фурье получим систему линейных алгебраических уравнений для нахождения функций $X_n(\lambda)$, $Y_n(\lambda)$, $Z_n(\lambda)$:

$$\begin{cases} a_{11}(n, \lambda)Z_n(\lambda) + a_{12}(n, \lambda)Y_n(\lambda) + a_{13}(\lambda)X_n(\lambda) = f_1(n, \lambda), \\ a_{21}(n, \lambda)Z_n(\lambda) + a_{22}(n, \lambda)Y_n(\lambda) + a_{23}(\lambda)X_n(\lambda) = f_2(n, \lambda), \\ a_{32}(n, \lambda)Y_n(\lambda) + a_{33}(n, \lambda)X_n(\lambda) = 0, \end{cases}$$

где

$$\begin{aligned} a_{11}(n, \lambda) &= \mu_2 \lambda \frac{d}{d\xi} \tilde{K}_n(\xi) - pG_n(\lambda) \tilde{K}_n(\xi), \quad a_{12}(n, \lambda) = -\mu_1 \lambda \frac{d}{d\xi} I_n(\xi) - pG_n(\lambda) I_n(\xi), \\ a_{13}(n, \lambda) &= -\mu_1 \lambda \frac{d}{d\xi} \tilde{K}_n(\xi) - pG_n(\lambda) \tilde{K}_n(\xi), \quad a_{21}(\lambda) = \mu_2 \lambda \frac{d}{d\xi} \tilde{K}_n(\xi) + qG_n(\lambda) \tilde{K}_n(\xi), \\ a_{22}(\lambda) &= \mu_1 \lambda \frac{d}{d\xi} I_n(\xi) - qG_n(\lambda) I_n(\xi), \quad a_{23}(n, \lambda) = \mu_1 \lambda \frac{d}{d\xi} \tilde{K}_n(\xi) - qG_n(\lambda) \tilde{K}_n(\xi), \\ a_{32}(\lambda) &= \frac{d}{d\zeta} I_n(\zeta), \quad a_{33}(\lambda) = \frac{d}{d\zeta} \tilde{K}_n(\zeta), \quad f_1(n, \lambda) = \left(-\mu_2 \lambda \frac{d}{d\xi} I_n(\xi) + pG_n(\lambda) I_n(\xi) \right) g_n(\lambda), \\ f_2(n, \lambda) &= \left(-\mu_2 \lambda \frac{d}{d\xi} I_n(\xi) - qG_n(\lambda) I_n(\xi) \right) g_n(\lambda), \quad \xi = \lambda a, \quad \zeta = \lambda b, \quad G_n(\lambda) = n^2 / a^2 + \lambda^2. \end{aligned}$$

Вычисление коэффициента экранирования. Коэффициент экранирования исходного магнитного поля тонкостенным цилиндрическим экраном с цилиндрическим включением в любой точке области D_1 определим как отношение значения вторичного магнитного потенциала U_1 в области D_1 к исходному значению потенциала U_0 :

$$Ke(\rho, z, \varphi) = \frac{|U_1(\rho, z, \varphi)|}{|U_0(\rho, z, \varphi)|}, \quad a + \Delta < \rho < b. \quad (10)$$

Используя Python-библиотеки Numpy, Scipy и Matplotlib [15], проведем вычисления коэффициента экранирования (10) для некоторых параметров задачи. Все сходящиеся бесконечные суммы определим с точностью 10^{-5} . Несобственные интегралы вычислим по методике, предложенной в работе [20, с. 503].

Представим некоторые результаты расчетов. Пусть область D_2 заполнена воздухом ($\mu_2 = 1,0000004\mu_0$, $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м), область D_1 – стеклом ($\mu_1 = 0,999987\mu_0$), материал тонкостенного экрана – медь ($\mu = 0,9999904\mu_0$, $\gamma = 5,8 \cdot 10^7$ См/м). Возьмем следующие геометрические параметры задачи: $a = 0,05$ м, $b = 0,02$ м, $\Delta = 0,0025$ м, $\ell = 0,2$ м. На рис. 2 показаны графики значений коэффициента экранирования $Ke(0,035, z, \pi)$ в области D_1 для $0,01$ м $< z < 1$ м и некоторых значений $f = 2\pi\omega$. Из приведенных вычислений следует, что с увеличением частоты

ты исходного поля коэффициент экранирования уменьшается для данных геометрических параметров задачи, т. е. поле практически не проникает через экранирующую систему при увеличении частоты исходного поля.

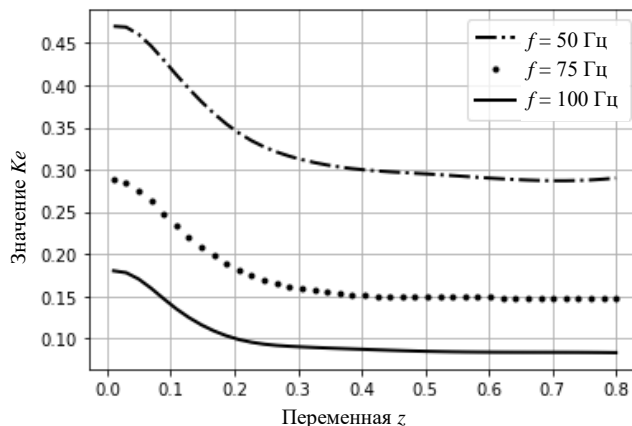


Рис. 2. Графики значений коэффициента экранирования $Ke(0,035, z, \pi)$ для некоторых частот исходного поля

Fig. 2. Shielding coefficient values $Ke(0,035, z, \pi)$ plot for some frequencies of original field

На рис. 3 показано, как изменяются значения коэффициента экранирования $Ke(0,04, z, \pi)$ для различных значений радиуса внутреннего цилиндра при $f = 50$ Гц и тех же значениях геометрических параметров, что и в приведенном выше примере расчетов. Из вычислений следует, что с уменьшением радиуса внутреннего цилиндра коэффициент экранирования уменьшается ($0,01 < z < 0,8$ м). С увеличением значений переменной z ($z > 0,6$ м) различия в значениях коэффициента экранирования практически отсутствуют.

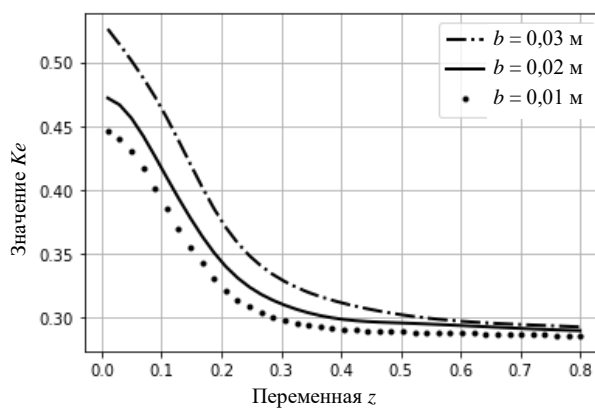


Рис. 3. Графики значений коэффициента экранирования $Ke(0,04, z, \pi)$ для некоторых значений радиуса цилиндра b и $f = 50$ Гц

Fig. 3. Shielding coefficient $Ke(0,04, z, \pi)$ plot for some values of cylinder radius b and $f = 50$ Hz

Пусть $\mu_2 = \mu_1 = \mu_0$, материал тонкостенного экрана – медь и геометрические параметры задачи: $a = 0,1$ м, $b = 0,02$ м, $\Delta = 0,0025$ м, $\ell = 0,2$ м. На рис. 4 изображены графики значений коэффициента экранирования $Ke(\rho, z, \pi)$ для $0,03 < \rho < 0,09$ м, $z = 0,01; 0,05; 0,1$ м и $f = 50$ Гц. Видно, что с увеличением значений переменной z коэффициент экранирования уменьшается.

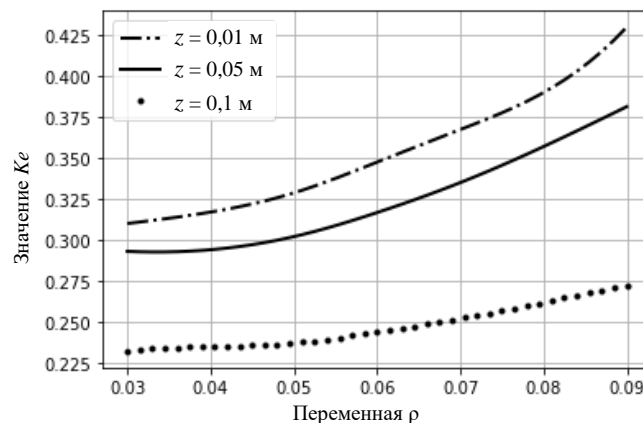


Рис. 4. Графики значений коэффициента экранирования $Ke(\rho, z, \pi)$ для некоторых значений переменной z и $f = 50$ Гц

Fig. 4. Shielding coefficient $Ke(\rho, z, \pi)$ plot for some values of variable z and $f = 50$ Hz

Пусть $\mu_2 = \mu_1 = \mu_0$ и геометрические параметры задачи $a = 0,1$ м, $b = 0,02$ м, $\Delta = 0,0025$ м, $\ell = 0,2$ м. Материал тонкостенного экрана имеет удельную электрическую проводимость $\gamma = 5,5 \cdot 10^7$ См/м и различные значения относительной магнитной проницаемости: $\mu_r = 100, 150, 200, 500$. На рис. 5 изображены графики коэффициента экранирования $Ke(\rho, 0,01, \pi)$ для $0,03 < \rho < 0,09$ м и $f = 50$ Гц. Из приведенных вычислений следует, что с увеличением значения относительной магнитной проницаемости материала экрана магнитное поле не проникает через тонкостенный экран.

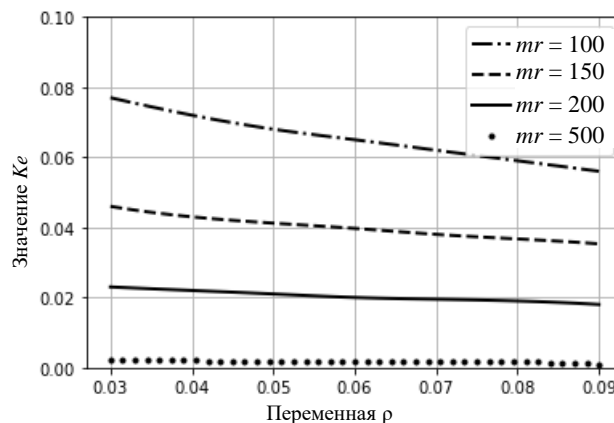


Рис. 5. Графики значений коэффициента экранирования $Ke(\rho, 0,01, \pi)$ для тонкостенного экрана с разными значениями относительной магнитной проницаемости экрана

Fig. 5. Shielding coefficient $Ke(\rho, 0,01, \pi)$ plot for thin-walled shield with different values of relative magnetic permeability of shield

Заключение. В статье разработан аналитико-численный алгоритм решения граничной трехмерной задачи, описывающей процесс проникновения низкочастотного магнитного поля через тонкостенный цилиндрический экран с цилиндрическим включением. В качестве источника поля рассмотрена тонкая нить с бесконечно малым поперечным сечением, по которой циркулирует ток. Нить расположена в плоскости, перпендикулярной оси цилиндрического экрана, во внешней области по отношению к экрану. С использованием соответствующих теорем сложения (формул) для гармонических функций решение поставленной задачи сведено к решению системы линейных алгебраических уравнений относительно коэффициентов, входящих в пред-

ставление вторичных полей. Численно исследовано влияние некоторых параметров задачи на значение коэффициента экранирования внешнего магнитного поля при прохождении через цилиндрический экран в присутствии цилиндрического включения.

Вычислительные эксперименты показали следующие результаты. Если область D_2 заполнена воздухом, область D_1 – стеклом, а материал тонкостенного экрана – медь, то при геометрических параметрах задачи $a = 0,05$ м, $b = 0,02$ м, $\Delta = 0,0025$ м, $\ell = 0,2$ м коэффициент экранирования $Ke(0,035, z, \pi)$ уменьшается с увеличением частоты исходного поля и при $f > 100$ Гц поле практически не проникает через данную экранирующую систему (см. рис. 2). При $f = 50$ Гц и тех же значениях параметров задачи коэффициент экранирования $Ke(0,04, z, \pi)$ уменьшается с уменьшением радиуса внутреннего цилиндра (см. рис. 3). С увеличением значения относительной магнитной проницаемости материала экрана μ_r при $\mu_2 = \mu_1 = \mu_0$, $a = 0,1$ м, $b = 0,02$ м, $\Delta = 0,0025$ м, $\ell = 0,2$ м, $f = 50$ Гц, $\gamma = 5,5 \cdot 10^7$ См/м коэффициент экранирования $Ke(\rho, 0,01, \pi)$, $0,03 < \rho < 0,09$ м, уменьшается и магнитное поле практически не проникает через цилиндрический экран при $\mu_r > 500$ (см. рис. 5).

Полученные результаты могут быть использованы для экранирования технических устройств и биологических объектов от воздействия магнитных полей и обеспечения экологической среды вокруг работающих электроустановок и устройств.

Список использованных источников

1. Кечиев, Л. Н. Экранирование радиоэлектронной аппаратуры. Инженерное пособие / Л. Н. Кечиев. – М. : Грифон, 2019. – 720 с.
2. Makarov, S. N. Low-Frequency Electromagnetic Modeling for Electrical and Biological Systems Using MATLAB / S. N. Makarov, G. M. Noetscher, A. Nazarian. – New Jersey : Wiley, 2015. – 616 p.
3. Дмитриев, В. И. Метод интегральных уравнений в вычислительной электродинамике / В. И. Дмитриев, Е. В. Захаров. – М. : МАКС Пресс, 2008. – 316 с.
4. Ильин, В. П. Методы конечных разностей и конечных объемов для эллиптических уравнений / В. П. Ильин. – Новосибирск : Ин-т математики им. С. Л. Соболева СО РАН, 2000. – 345 с.
5. Pierrus, J. Solved Problems in Classical Electromagnetism: Analytical and Numerical Solutions with Comments / J. Pierrus. – Oxford : Oxford University Press, 2018. – 638 p. <https://doi.org/10.1093/oso/9780198821915.001.0001>
6. Шушкевич, Г. Ч. Моделирование полей в многосвязных областях в задачах электростатики / Г. Ч. Шушкевич. – Saarbruchen : LAP LAMBERT Academic Publishing, 2015. – 228 с.
7. Erofeenko, V. T. Shielding of a low-frequency electric field by a multilayer circular disk / V. T. Erofeenko, G. H. Shushkevich // Technical Physics. – 2013. – Vol. 58, no. 6. – P. 866–871. <https://doi.org/10.1134/1.1562255>
8. Шушкевич, Г. Ч. Проникновение низкочастотного магнитного поля через плоский слой со сфероидальным включением, тонкостенными слоями / Г. Ч. Шушкевич, А. И. Куц // Веснік Гродз. дзярж. універсітэта. Сер. 2. Матэматыка. Фізіка. Інфарматыка, вылічальная тэхніка і кіраванне. – 2012. – № 3. – С. 45–52.
9. Ерофеенко, В. Т. Экранирование низкочастотного электрического поля тонкостенной незамкнутой сферической оболочкой с учетом емкостных свойств / В. Т. Ерофеенко, Г. Ч. Шушкевич // Электричество. – 2011. – № 6. – С. 57–61.
10. Erofeenko, V. T. Screening of low-frequency magnetic fields by an open thin-wall spherical shell / V. T. Erofeenko, I. S. Kozlovskaya, G. Ch. Shushkevich // Technical Physics. – 2010. – Vol. 55, no. 9. – P. 1240–1247. <https://doi.org/10.1134/S1063784210090021>
11. Аполлонский, С. М. Экранирование низкочастотного магнитного поля / С. М. Аполлонский, Г. Ч. Шушкевич // Электричество. – 2005. – № 4. – С. 57–61.
12. Глушцов, А. И. Численное исследование усредненных граничных условий для тонких цилиндрических электромагнитных оболочек / А. И. Глушцов, В. Т. Ерофеенко // Вес. акад. навук Беларусі. Сер. фіз.-мат. навук. – 1996. – № 4. – С. 110–114.
13. Ерофеенко, В. Т. Краевые задачи с интегральными граничными условиями для моделирования магнитных полей в цилиндрических пленочных оболочках / В. Т. Ерофеенко, Г. Ф. Громько, Г. М. Заяц // Дифференциальные уравнения. – 2017. – Т. 53, № 7. – С. 962–975. <https://doi.org/10.1134/S037406411707010X>

14. Щербинин, А. Г. Численное исследование эффективности цилиндрического электромагнитного экрана / А. Г. Щербинин, А. С. Мансуров // *Электротехника*. – 2017. – № 11. – С. 33–37.
15. Johansson, R. *Numerical Python: Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib* / R. Johansson. – N. Y. : Apress, 2019. – 700 p. <https://doi.org/10.1007/978-1-4842-4246-9>
16. Аполлонский, С. М. Эквивалентные граничные условия в электродинамике / С. М. Аполлонский, В. Т. Ерофеев. – СПб. : Безопасность, 1999. – 416 с.
17. Теоретические основы электротехники : в 3 т. / К. С. Демирчян [и др.]. – СПб. : Питер, 2003. – Т. 3. – 377 с.
18. Справочник по специальным функциям с формулами, графиками и таблицами / под ред. М. Абрамовица, И. Стиган. – М. : Наука, 1979. – 830 с.
19. Ерофеев, В. Т. Теоремы сложения / В. Т. Ерофеев. – Минск : Наука и техника, 1989. – 240 с.
20. Вержицкий, В. М. Основы численных методов / В. М. Вержицкий. – М. : Высшая школа, 2002. – 848 с.

References

1. Kechiev L. N. *Ekranirovanie radioelektronnoi apparatury. Inzhenernoe posobie. Shielding of Electronic Equipment. Engineering Manual*. Moscow, Grifon, 2019, 720 p. (In Russ.).
2. Makarov S. N., Noetscher G. M., Nazarian A. *Low-Frequency Electromagnetic Modeling for Electrical and Biological Systems Using MATLAB*. New Jersey, Wiley, 2015, 616 p.
3. Dmitriev V. I., Zakharov E. V. *Metod integral'nykh uravnenii v vychislitel'noi elektrodinamike. The Method of Integral Equations in Computational Electrodynamics*. Moscow, MAKS Press, 2008, 316 p. (In Russ.).
4. Piyin V. P. *Metody konechnykh raznostei i konechnykh ob'emov dlia ellipticheskikh uravnenii. Finite Difference and Finite Volume Methods for Elliptic Equations*. Novosibirsk, Institut matematiki imeni S. L. Soboleva Sibirskogo otdelenija Rossijskoj akademii nauk, 2000, 345 p. (In Russ.).
5. Pierrus J. *Solved Problems in Classical Electromagnetism: Analytical and Numerical Solutions with Comments*. Oxford, Oxford University Press, 2018, 638 p. <https://doi.org/10.1093/oso/9780198821915.001.0001>
6. Shushkevich G. Ch. *Modelirovanie polei v mnogosvyaznykh oblastiakh v zadachakh elektrostati. Modeling Fields in Multiply Connected Domains in Electrostatic Problems*. Saarbruchen, LAP LAMBERT Academic Publishing, 2015, 228 p. (In Russ.).
7. Erofeenko V. T., Shushkevich G. H. *Shielding of a low-frequency electric field by a multilayer circular disk. Technical Physics*, 2013, vol. 58, no. 6, pp. 866–871. <https://doi.org/10.1134/1.1562255>
8. Shushkevich G. Ch., Kuts A. I. *Penetration of a low-frequency magnetic field through a flat layer with a spheroidal inclusion, thin-walled layers. Vesnik Grodzenskaga dzyarzhaynaga Źniversiteta. Ser. 2. Matematyka. Fizika. Infarmatyka, vylichal'naya tekhnika i kiravanne [Bulletin of Grodno State University. Ser. 2. Mathematics. Physics. Informatics, Computer Engineering and Management]*, 2012, no. 3. pp. 45–52 (In Russ.).
9. Erofeenko V. T., Shushkevich G. Ch. *Shielding of a low-frequency electric field by a thin-walled open spherical cover taking into account capacitive properties. Elektrichestvo [Electricity]*, 2011, no. 6, pp. 57–61 (In Russ.).
10. Erofeenko V. T., Kozlovskaya I. S., Shushkevich G. Ch. *Screening of low-frequency magnetic fields by an open thin-wall spherical shell. Technical Physics*, 2010, vol. 55, no 9, pp. 1240–1247. <https://doi.org/10.1134/S1063784210090021>
11. Apollonskii, S. M., Shushkevich G. Ch. *Low frequency magnetic field shielding. Elektrichestvo [Electricity]*, 2005, no. 4, pp. 57–61 (In Russ.).
12. Glushtsov A. I., Erofeenko V. T. *Numerical study of averaged boundary conditions for thin cylindrical electromagnetic shells. Vestsi akademii navuk Belarusi. Seryia fizika-matematychnykh navuk [Proceedings of the Academy of Sciences of Belarus. Physics and Mathematics Series]*, 1996, no. 4, pp. 110–114 (In Russ.).
13. Erofeenko V. T., Gromyko G. F., Zayats G. M. *Boundary value problems with integral boundary conditions for modeling magnetic fields in cylindrical film shells. Differentsial'nye uravneniya [Differential Equations]*, 2017, vol. 53, no. 7, pp. 962–975 (In Russ.). <https://doi.org/10.1134/S037406411707010X>
14. Shcherbinin A. G., Mansurov A. S. *Numerical study of the efficiency of a cylindrical electromagnetic shield. Elektrotehnika [Electrical engineering]*, 2017, no. 11, pp. 33–37 (In Russ.).
15. Johansson R. *Numerical Python: Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib*. New York, Apress, 2019, 700 p. <https://doi.org/10.1007/978-1-4842-4246-9>

16. Apollonsky S. M., Erofeenko V. T. Ekvivalentnye granichnye usloviia v elektrodinamike. *Equivalent Boundary Conditions in Electrodynamics*. Saint Petersburg, Bezopasnost', 1999, 416 p. (In Russ.).
17. Demirchyan K. S., Neiman L. R., Korovkin N. V., Chechurin V. L. Teoreticheskie osnovy elektrotekhniki : v 3 tomah. *Theoretical Foundations of Electrical Engineering : in 3 volumes*. Saint Petersburg, Piter, 2003, vol. 3, 377 p. (In Russ.).
18. In Abramovitz M., Stigan I. (eds.). Spravochnik po spetsial'nym funktsiiam s formulami, grafikami i tablitsami. *Handbook of Special Functions with Formulas, Graphs and Tables*. Moscow, Nauka, 1979, 830 p. (In Russ.).
19. Erofeenko V. T. Teoremy slozheniya. *Addition Theorems*. Minsk, Nauka i tekhnika, 1989, 240 p. (In Russ.).
20. Verzhbitsky V. M. Osnovy chislennykh metodov. *Fundamentals of Numerical Methods*. Moscow, Vysshaya shkola, 2002, 848 p. (In Russ.).

Информация об авторе

Шушкевич Геннадий Чеславович, доктор физико-математических наук, профессор кафедры современных технологий программирования, Гродненский государственный университет им. Янки Купалы.
E-mail: gsys@grsu.by

Information about the author

Gennady Ch. Shushkevich, Dr. Sci. (Phys.-Math.), Professor of Modern Programming Technologies Department, Yanka Kupala State University of Grodno.
E-mail: gsys@grsu.by



УДК 62-503.51:62-533.65
<https://doi.org/10.37661/1816-0301-2021-18-3-59-67>

Оригинальная статья
Original Paper

Автоматическое регулирование температуры воздуха в теплице с учетом данных прогноза погоды

А. Г. Сеньков

*Белорусский государственный аграрный
технический университет,
пр. Независимости, 99, Минск, 220023, Беларусь
E-mail: sencov_1981@tut.by*

Аннотация. Рассматриваются вопросы повышения энергоэффективности и улучшения качества автоматического регулирования температуры воздуха в промышленных теплицах. Предлагается метод управления температурой воздуха в промышленной теплице на основе принципа компенсации возмущения с использованием данных прогноза погоды. Динамические свойства объекта управления приближенно описываются линейным апериодическим звеном первого порядка с запаздыванием. Для решения проблемы наличия транспортного запаздывания в объекте управления при выработке управляющего сигнала в цепи компенсации возмущения помимо сигнала датчика наружной температуры учитываются также данные прогноза изменения наружной температуры на определенный предстоящий интервал времени, что позволяет заблаговременно осуществлять соответствующее компенсирующее управляющее воздействие. Приводятся структурная схема системы автоматического регулирования температуры в помещении, теоретические математические выражения для синтеза регулятора и анализа качества регулирования, а также результаты численного моделирования процесса регулирования температуры. Предлагаемый метод управления может быть использован при регулировании температуры воздуха в зданиях и сооружениях с повышенными требованиями к параметрам микроклимата.

Ключевые слова: микроклимат, теплица, автоматическое регулирование, принцип компенсации возмущения, предиктивное управление

Для цитирования. Сеньков, А. Г. Автоматическое регулирование температуры воздуха в теплице с учетом данных прогноза погоды / А. Г. Сеньков // Информатика. – 2021. – Т. 18, № 3. – С. 59–67. <https://doi.org/10.37661/1816-0301-2021-18-3-59-67>

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию | Received 18.06.2021
Подписана в печать | Accepted 22.07.2021
Опубликована | Published 29.09.2021

Automatic control of air temperature in the greenhouse considering the weather forecast data

Andrey G. Senkov

Belarusian State Agrarian Technical University,
av. Nezavisimosti, 99, Minsk, 220023, Belarus
E-mail: sencov_1981@tut.by

Abstract. The paper concerns the problem of improving energy efficiency and the quality of air temperature automatic control in industrial greenhouses. The article proposes a method based on the principle of disturbance compensation using weather forecast data. The plant dynamic model structure is described by the linear 1st order aperiodic link with delay both for the control circuit and for the disturbance circuit. To solve the problem of time delays in the plant mathematical model it is suggested along with the use of outside temperature sensor signal to take into account the weather forecast data for the outdoor temperature changes over the specific forthcoming interval of time. It allows the controller to generate appropriate compensation control signal in advance. The block diagram of the automatic control system of the greenhouse temperature is given. Theoretical mathematical expressions for the regulation algorithm and analysis of regulation quality are obtained and the results of numerical modeling of the temperature control process are presented. The proposed automatic control method can be used to regulate the air temperature in buildings with high requirements for microclimate parameters.

Keywords: microclimate, greenhouse, automatic control, disturbance compensation control, predictive control

For citation. Senkov A. G. Automatic control of air temperature in the greenhouse considering the weather forecast data. *Informatics*, 2021, vol. 18, no. 3, pp. 59–67 (In Russ.).
<https://doi.org/10.37661/1816-0301-2021-18-3-59-67>

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

Введение. Поддержание оптимальных параметров микроклимата внутри зданий и сооружений является одной из основных задач систем автоматизации современных тепличных комплексов. Комфортные условия для растений должны обеспечиваться даже при резких колебаниях температуры наружного воздуха в ночное и дневное время холодного периода года. При этом большое значение имеет энергоэффективность систем отопления. Сокращение тепловых энергетических затрат на обогрев промышленных теплиц особенно актуально для местностей и регионов с умеренным и прохладным климатом.

Целью проводимых исследований является улучшение энергоэффективности систем надпочвенного обогрева промышленных теплиц путем совершенствования известных линейных методов автоматического регулирования по отклонению и возмущению за счет дополнительного использования данных краткосрочного метеорологического прогноза по предстоящему изменению наружной температуры воздуха.

Исходная математическая модель автоматической системы трубного надпочвенного обогрева теплицы. В настоящее время одним из лучших по энергоэффективности считается комбинированное управление, сочетающее в себе принцип регулирования по отклонению температуры внутри сооружения от заданного оптимального значения и принцип компенсации основного возмущения – изменений наружной температуры воздуха [1]. Структура комбинированной односвязной системы автоматического регулирования (САР) температуры воздуха в помещении показана на рис. 1. Вследствие достаточно медленных по времени процессов изменения температуры воздуха внутри и снаружи сооружения можно пренебречь инерционностью соответствующих датчиков температуры по сравнению с тепловой инерцией самой теплицы.

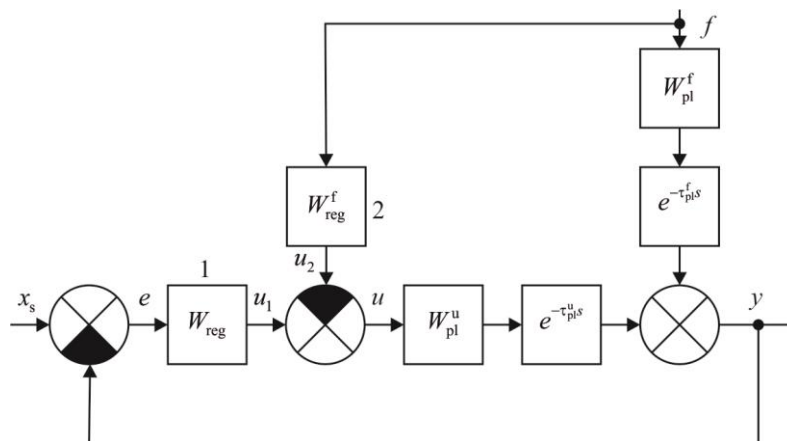


Рис. 1. Структурная схема комбинированной односвязной системы автоматического регулирования температуры воздуха в помещении [2]:
 x_s – задающее воздействие (требуемая температура воздуха) внутри сооружения; $e = x_s - y$ – ошибка регулирования

Fig. 1. Block diagram of combined single-connected system for automatic control of indoor air temperature in the industrial greenhouse [2]:

x_s – preset reference signal (required value of indoor air temperature); $e = x_s - y$ – control error

Сначала выполняется синтез передаточной функции $W_{\text{reg}}(s)$ регулятора в цепи управления по отклонению на один из типовых переходных процессов с использованием известных методов настройки, обеспечивающих требуемые значения запасов устойчивости¹. Затем синтезируется передаточная функция компенсатора возмущения $W_{\text{reg}}^f(s)$. Передаточная функция объекта управления (ОУ) по возмущающему воздействию определяется выражением [2]

$$W^f(s) = \frac{y(s)}{f(s)} = \frac{W_{\text{pl}}^f(s)e^{-\tau_{\text{pl}}^f s} - W_{\text{pl}}^u(s)e^{-\tau_{\text{pl}}^u s}W_{\text{reg}}^f(s)}{1 + W_{\text{pl}}^u(s)e^{-\tau_{\text{pl}}^u s}W_{\text{reg}}(s)}, \quad (1)$$

где y – действительное значение температуры воздуха внутри сооружения (управляемая величина); f – изменение наружной температуры воздуха (возмущающее воздействие); s – комплексная переменная; $W_{\text{pl}}^f = W_{\text{pl}}^f(s)$ и $e^{-\tau_{\text{pl}}^f s}$ – передаточная функция и транспортное запаздывание ОУ по возмущающему воздействию; $W_{\text{pl}}^u = W_{\text{pl}}^u(s)$ и $e^{-\tau_{\text{pl}}^u s}$ – передаточная функция и транспортное запаздывание ОУ по управляющему воздействию.

Из равенства (1) следует известное условие инвариантности рассматриваемой САУ к изменениям возмущающего воздействия¹, заключающееся в данном случае в необходимости синтеза компенсатора возмущения в соответствии с формулой [2]:

$$W_{\text{reg}}^f(s) = \frac{W_{\text{pl}}^f(s)e^{-\tau_{\text{pl}}^f s}}{W_{\text{pl}}^u(s)e^{-\tau_{\text{pl}}^u s}}. \quad (2)$$

Синтез компенсатора возмущения $W_{\text{reg}}^f(s)$ по формуле (2) может существенно уменьшить влияние основного внешнего возмущения на процесс управления. Так, например, при отсут-

¹Ротач, В. Я. Теория автоматического управления : учеб. для вузов / В. Я. Ротач. – М. : Изд. дом МЭИ, 2008. – 396 с.

ствии транспортного запаздывания в математической модели ОУ и аппроксимации динамики ОУ типовыми аperiodическими звеньями первого порядка

$$W_{pl}^u(s) = \frac{K_{pl}^u}{T_{pl}^u s + 1}, \quad W_{pl}^f(s) = \frac{K_{pl}^f}{T_{pl}^f s + 1} \quad (3)$$

можно добиться пятикратного уменьшения перерегулирования переходной характеристики САР по возмущающему воздействию по сравнению с САР только по отклонению, т. е. без контура компенсации возмущения [2]. Полная инвариантность не достигается вследствие накладываемых дополнительных ограничений по типу «насыщение» на выходные сигналы используемых в схеме реальных пропорционально-интегрально-дифференциальных регуляторов (ПИД-регуляторов).

В то же время при математическом описании переходных тепловых процессов в промышленных теплицах больших размеров нельзя пренебрегать транспортным запаздыванием в цепи управления, физически обусловленным большой протяженностью подающих и обратных трубопроводов в системе трубного надпочвенного обогрева. Поэтому в качестве математической модели промышленной теплицы следует использовать известную модель «серого ящика» с передаточными функциями по управляющему и возмущающему воздействиям в виде типового аperiodического звена первого порядка с транспортным запаздыванием [3–5]:

$$W_{pl}^u(s) = \frac{K_{pl}^u}{T_{pl}^u s + 1} e^{-\tau_{pl}^u s}, \quad W_{pl}^f(s) = \frac{K_{pl}^f}{T_{pl}^f s + 1} e^{-\tau_{pl}^f s}. \quad (4)$$

При синтезе компенсирующего регулятора в цепи возмущения $W_{reg}^f(s)$ по формуле (2) можно столкнуться с невозможностью его физической реализации¹ при $\tau_{pl}^u > \tau_{pl}^f$, так как в этом случае

$$W_{reg}^f(s) = \frac{K_{pl}^f}{(T_{pl}^f s + 1)} \frac{(T_{pl}^u s + 1)}{K_{pl}^u} e^{\Delta\tau s},$$

где $\Delta\tau = \tau_{pl}^u - \tau_{pl}^f > 0$.

Кроме того, наличие значительного транспортного запаздывания в цепи управления объектом, как известно, значительно ухудшает качество регулирования по отклонению и устойчивость САР при использовании ПИД-регулятора¹. Для решения проблемы транспортного запаздывания в цепи управления существуют методы управления и структурные схемы САР с использованием предиктора Смита². Применение же компенсации возмущения для объекта с транспортным запаздыванием затруднено, так как оно связано с отмеченной проблемой физической реализуемости регулятора $W_{reg}^f(s)$.

Для преодоления указанной проблемы предлагается использовать данные краткосрочного почасового прогноза погоды, позволяющие с достаточно большой точностью прогнозировать изменения наружной температуры воздуха на сутки вперед. Современные серверы погоды предоставляют в распоряжение пользователей программный интерфейс³ (англ. application programming interface, API), с помощью которого подключенный к сети Интернет пользователь (микропроцессорное устройство управления САР) может сформировать программный запрос

²Системы автоматического управления с запаздыванием : учеб. пособие / Ю. Ю. Громов [и др.]. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2007. – 76 с.

³Работа с погодой в PHP с помощью OpenWeatherMap API [Электронный ресурс] – Режим доступа: <https://prognose.ru/web-dev/beck-end/working-with-weather-in-php-using-the-openweathermap-api/>. – Дата доступа: 12.04.2021.

к серверу погоды и получить от него подробный прогноз (данные о температуре, влажности, скорости ветра, облачности, осадках) на ближайшие сутки⁴.

Предлагаемая в настоящей статье идея состоит в том, чтобы в комбинированной САР температуры в теплице в цепи компенсации возмущения в качестве входного сигнала для регулятора 2 (см. рис. 1) использовать помимо датчика наружной температуры и данные прогноза изменения наружной температуры на определенный предстоящий интервал времени τ^* для заблаговременной выработки соответствующего компенсирующего управляющего воздействия на ОУ. В итоге это позволит улучшить качество компенсации изменений наружной температуры воздуха и уменьшить энергозатраты на систему электрокалориферного надпочвенного обогрева, используемую обычно дополнительно с системой трубного надпочвенного обогрева для быстрой компенсации резких по времени изменений наружной температуры воздуха.

Математическая модель автоматической системы обогрева с учетом метеорологического прогноза по наружной температуре. Для реализации указанной идеи исходная структурная схема САР на рис. 1 должна быть преобразована к схеме, представленной на рис. 2. Непосредственно на ОУ возмущающим воздействием служит изменение текущего действительного значения наружной температуры воздуха $f(t)$, тогда как на регулятор 2 (рис. 2) в цепи компенсации подается изменение наружной температуры $f^*(t)$, прогнозируемое на некоторое ближайшее время τ^* вперед.

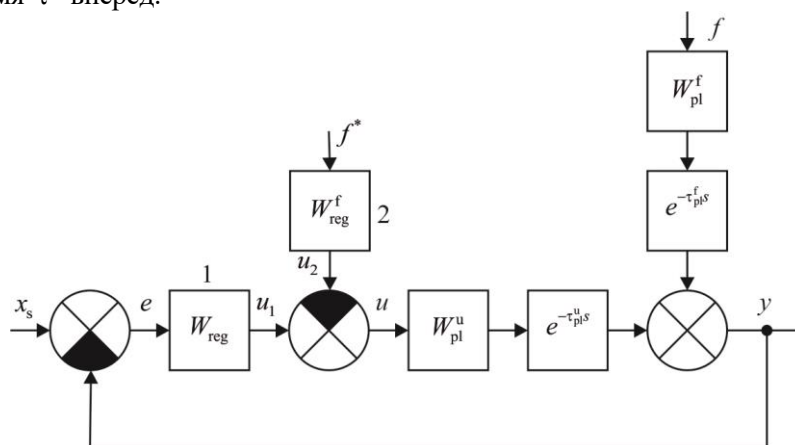


Рис. 2. Структурная схема САР с компенсацией изменения наружной температуры воздуха с учетом данных метеопрогноза

Fig. 2. Block diagram of automatic control system with compensation of outdoor air temperature changes taking into account weather forecast data

Соответствующее изменение управляемой величины $y(s)$ вместо выражения (1) приобретает вид

$$y(s)\Big|_{x_s=0} = \frac{W_{pl}^f(s)e^{-\tau_{pl}^f s} f(s) - W_{pl}^u(s)e^{-\tau_{pl}^u s} W_{reg}^f(s) f^*(s)}{1 + W_{pl}^u(s)e^{-\tau_{pl}^u s} W_{reg}(s)}, \quad (5)$$

а условие инвариантности рассматриваемой САР к изменениям возмущающего воздействия вместо (2) преобразуется в выражение

$$W_{reg}^f(s) = \frac{W_{pl}^f(s)e^{-\tau_{pl}^f s}}{W_{pl}^u(s)e^{-\tau_{pl}^u s}} \cdot \frac{f(s)}{f^*(s)}. \quad (6)$$

⁴OpenWeatherMap – как энтузиасты делают погоду [Электронный ресурс] – Режим доступа: <https://habr.com/ru/post/164045/>. – Дата доступа: 12.04.2021.

Таким образом, если прогноз изменения температуры воздуха абсолютно точен, по крайней мере на время τ^* вперед, т. е. изображение по Лапласу $f^*(s)$ «опережает» изображение $f(s)$ на время τ^* :

$$f(s) = f^*(s)e^{-\tau^*s}, \quad (7)$$

то выражение (6) для передаточной функции регулятора 2 в цепи компенсации принимает вид

$$W_{\text{reg}}^f(s) = \frac{W_{\text{pl}}^f(s)}{W_{\text{pl}}^u(s)} e^{-(\tau_{\text{pl}}^f + \tau_{\text{pl}}^u - \tau^*)s}. \quad (8)$$

Следовательно, при точно известных значениях τ_{pl}^u , τ_{pl}^f времени транспортного запаздывания ОУ и при соответствующем значении времени τ^* метеорологического прогноза, удовлетворяющем соотношению

$$\tau_{\text{pl}}^f + \tau_{\text{pl}}^u - \tau^* \geq 0, \quad (9)$$

положительно решается проблема физической реализуемости регулятора $W_{\text{reg}}^f(s)$ компенсации возмущения и теоретически достигается инвариантность рассматриваемой САР к изменениям возмущающего воздействия.

В ходе практической реализации рассматриваемого подхода неизбежно следует учитывать наличие погрешности данных краткосрочного метеопрогноза

$$f^*(t - \tau^*) = f(t) + \Delta f(t),$$

которую по аналогии с «опережением» (7) можно записать следующим образом:

$$f^*(s)e^{-\tau^*s} = f(s) + \Delta f(s), \quad (10)$$

где $\Delta f(s)$ – изображение по Лапласу температурной погрешности метеопрогноза.

С учетом равенства (10) выражение (5) можно преобразовать к виду

$$y(s)\Big|_{x_s=0} = \frac{\left(W_{\text{pl}}^f(s)e^{-\tau_{\text{pl}}^f s} - W_{\text{pl}}^u(s)e^{-(\tau_{\text{pl}}^u - \tau^*)s}W_{\text{reg}}^f(s)\right)}{1 + W_{\text{pl}}^u(s)e^{-\tau_{\text{pl}}^u s}W_{\text{reg}}(s)} f(s) - \frac{W_{\text{pl}}^u(s)e^{-(\tau_{\text{pl}}^u - \tau^*)s}W_{\text{reg}}^f(s)}{1 + W_{\text{pl}}^u(s)e^{-\tau_{\text{pl}}^u s}W_{\text{reg}}(s)} \Delta f(s). \quad (11)$$

Рассматривая температурную погрешность краткосрочного метеопрогноза $\Delta f(t)$ в качестве случайной величины и не имея при этом достоверных данных о плотности вероятности ее распределения⁵, как и в случае использования предиктора Смита [6], для уменьшения влияния погрешности $\Delta f(t)$ можно предложить модификацию структурной схемы на рис. 2 (рис. 3). На рис. 3 видно, что регулятор 2, будучи по-прежнему синтезированным в соответствии с выражением (8), обеспечивает равенство нулю числителя первой из дробей в выражении (11):

⁵Хабутдинов, Ю. Г. Краткосрочные метеорологические прогнозы : материалы по курсам «Синоптическая метеорология» и «Гидрометеорологическое обеспечение народного хозяйства» / Ю. Г. Хабутдинов, К. М. Шанталинский. – Казань : КГУ, 2008. – 52 с.

$W_{pl}^f(s)e^{-\tau_{pl}^f s} - W_{pl}^u(s)e^{-(\tau_{pl}^u - \tau^*)s}W_{reg}^f(s) = 0$. В то же время значение ошибки метеопрогноза $\Delta f(t)$ оценивается сравнивающим устройством 3 в соответствии с выражением (10) и сигнал рассогласования $\Delta f(t)$ также вводится в основной контур регулирования по отклонению, где его подавление выполняется с помощью регулятора 1.

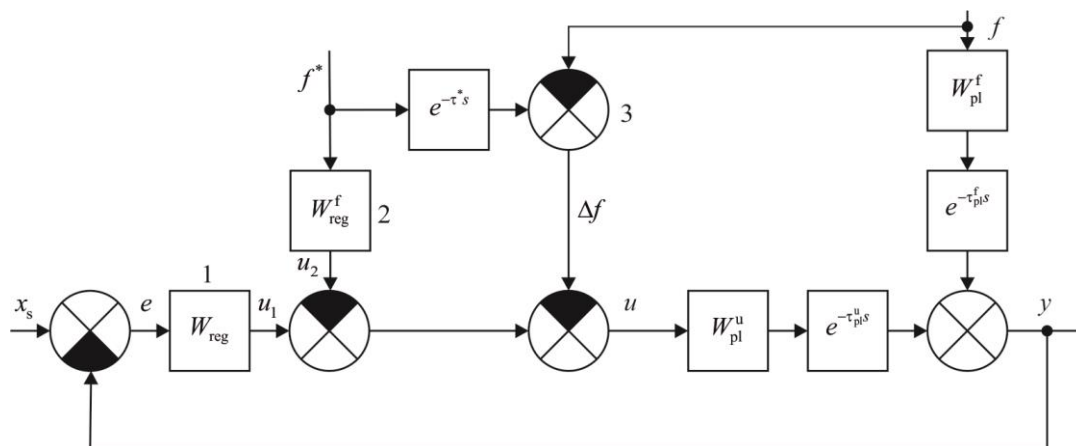


Рис. 3. Структурная схема АСР с компенсацией изменения наружной температуры воздуха с учетом возможных погрешностей данных метеопрогноза

Fig. 3. Block diagram of automatic control system with compensation of outdoor air temperature changes taking into account weather forecast data errors

Проверка реализуемости предложенной схемы АСР, а также оценка качества регулирования выполнялись методом компьютерного моделирования с использованием программного пакета динамического моделирования Simulink [7]. В качестве ОУ рассматривается блочная теплица, в которой для регулирования температуры воздуха в надпочвенной зоне имеется контур трубного водяного обогрева⁶. Управляющим воздействием на ОУ является изменение температуры теплоносителя путем изменения степени открытия трехходового смесительного клапана, установленного в контуре водяного обогрева. Для параметров динамической модели (4) ОУ приняты следующие числовые значения⁶: $T_{pl}^u = 35$ мин, $\tau_{pl}^u = 30$ мин, $K_{pl}^u = 1$ °С, $T_{pl}^f = 10$ мин, $\tau_{pl}^f = 0$ мин, $K_{pl}^f = 1$. Погрешность данных метеопрогноза моделировалась постоянным значением $\Delta f(t) = 0,25$ °С. Полученный график переходной функции АСР по возмущающему воздействию изображен на рис. 4. Как видно из рис. 4, а, значение перерегулирования при использовании предлагаемого комбинированного метода регулирования с учетом метеопрогноза (кривая 1) может быть уменьшено в четыре раза по сравнению с базовым вариантом схемы с пропорционально-интегральным регулированием по отклонению (кривая 2).

На рис. 4, б показан график соответствующего изменения во времени управляющего сигнала $u(t)$ блока регуляторов. В рассматриваемой комбинированной схеме изменение управляющего сигнала $u_{(1)}$ (кривая 1) вследствие учета данных метеопрогноза начинается заблаговременно, т. е. еще до момента времени $t = 0$ наступления ступенчатого возмущения. В то же время при использовании пропорционально-интегрального регулирования только по отклонению реакция АСР $u_{(2)}$ (кривая 2) на внешнее возмущение начинается после его наступления. Таким образом, помимо улучшения качества переходного процесса в системе за счет предусмотрительного снижения температуры теплоносителя при повышении наружной темпе-

⁶Ерков, А. А. Система управления микроклиматом в отделении блочных теплиц : автореф. дис. ... кан. тех. наук / А. А. Ерков. – М., 1995. – 20 с.

ратуры достигается также экономия тепловой энергии, пропорциональная величине

$$\Delta Q = \int_{-\tau}^{+\infty} (u_{(2)} - u_{(1)}) dt.$$

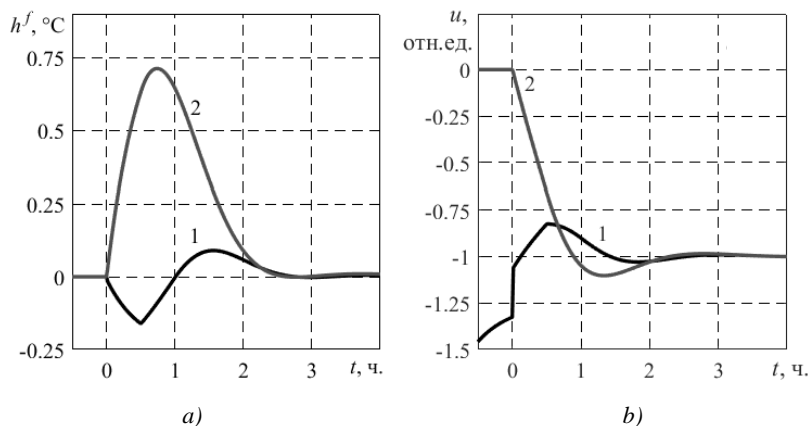


Рис. 4. Реакция САР на ступенчатое возмущающее воздействие – увеличение наружной температуры воздуха на 1°C: а) переходная функция САР по возмущающему воздействию; б) изменение управляющего сигнала регулятора;

1 – при наличии контура компенсации возмущения с учетом данных прогноза погоды; 2 – при регулировании по отклонению и отсутствии контура компенсации

Fig. 4. Automatic control system response to step increase of outdoor air temperature by 1°C: a) automatic control system time response – indoor air temperature change at time; b) controller output signal change over time;

1 – a considered system with main disturbance compensation circuit taking into account weather forecast data; 2 – a basic automatic control system with main feedback circuit and without weather forecast data

Заключение. Использование комбинированного управления позволяет улучшить качество регулирования системы при изменении возмущающего воздействия, но при этом возникает дополнительная задача синтеза параметров регулятора в цепи компенсации основного внешнего возмущения. Предложенные в настоящей работе способ компенсации изменений наружной температуры воздуха и структурная схема САР, основанные на дополнительном применении данных метеопрогноза по ожидающимся изменениям наружной температуры, решают проблему практического использования линейных регуляторов, например пропорционально-интегрального и пропорционального регуляторов, в цепи компенсации внешнего возмущения при наличии значительного транспортного запаздывания в математической динамической модели ОУ, которым в данном случае является блочная теплица с контуром водяного надпочвенного обогрева. В случае настройки регулятора в цепи компенсации возмущения исходя из критерия инвариантности САР к изменениям возмущающего воздействия, как показало численное моделирование, может быть достигнуто значительное улучшение показателей качества регулирования, например четырехкратное снижение значения перерегулирования переходного процесса по сравнению с базовым вариантом САР по отклонению.

Предложенный подход может быть применим также при разработке интеллектуальных систем отопления производственных, жилых и сельскохозяйственных помещений, характеризующихся значительной тепловой инерцией.

Список использованных источников

1. Панферов, В. И. Адаптация погодного графика регулирования отопления / В. И. Панферов, С. В. Панферов // Вестник МГСУ. – 2011. – № 7. – С. 257–261.
2. Шейх Эль Нажжарин, М. Х. Синтез компенсатора возмущения в комбинированной системе автоматического управления с ПИД-регулятором в цепи основного канала / М. Х. Шейх Эль Нажжарин, А. Г. Сеньков, Н. Н. Гурский // Доклады БГУИР. – 2015. – № 4(90). – С. 98–102.

3. Панферов, В. И. Динамическая модель отопительных приборов и систем / В. И. Панферов, С. В. Панферов // Вестник ЮУрГУ. Сер. «Компьютерные технологии, управление, радиоэлектроника». – 2015. – Т. 15, № 2. – С. 75–82.
4. Сотников, А. Г. Автоматизация систем кондиционирования воздуха и вентиляции / А. Г. Сотников. – Л. : Машиностроение, 1984. – 240 с.
5. Сканави, А. Н. Переходные тепловые процессы в отопительных приборах / А. Н. Сканави, Л. М. Махов, В. Э. Сварич // Изв. вузов. Строительство и архитектура. – 1986. – № 4. – С. 86–88.
6. Бобиков, А. И. Настройка параметров предиктора Смита с фильтром с помощью Simulink Response Optimization / А. И. Бобиков, Т. С. Бубнова // Вестник РГРТУ. – 2017. – № 61. – С. 96–104.
7. Ощепков, А. Ю. Системы автоматического управления: теория, применение, моделирование в MATLAB / А. Ю. Ощепков. – СПб. : Лань, 2013. – 208 с.

References

1. Panferov V. I., Panferov S. V. *Adaptation of the weather heating control schedule*. Vestnik Moskovskogo gosudarstvennogo stroitel'nogo universiteta [Moscow State University of Civil Engineering Bulletin], 2011, no. 7, pp. 257–261 (In Russ.).
2. Cheikh El Najjarine M. H., Senkov A. G., Hurski N. N. *Synthesis of feedforward compensator in the combined system of automatic control with PID-regulator on the main channel circuit*. Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki [Reports of the Belarusian State University of Informatics and Radioelectronics], 2015, no. 4(90), pp. 98–102 (In Russ.).
3. Panferov V. I., Panferov S. V. *Dynamic model of heating devices and systems*. Vestnik Juzhno-Ural'skogo gosudarstvennogo universiteta. Serija "Komp'yuternye tehnologii, upravlenie, radioelektronika" [Bulletin of the South Ural State University. Series "Computer technology, Control, Electronics"], 2015, vol. 15, no. 2, pp. 75–82 (In Russ.).
4. Sotnikov A. G. *Avtomatizacija sistem kondicionirovanija vozduha i ventiljacii. Automation of Air Conditioning and Ventilation Systems*. Leningrad, Mashinostroenie, 1984, 240 p. (In Russ.).
5. Skanavi A. N., Mahov L. M., Svaritch V. E. *Transient thermal processes in heating devices*. Izvestija vuzov. Stroitel'stvo i arhitektura [Proceedings of Universities. Construction and Architecture], 1986, no. 4, pp. 86–88 (In Russ.).
6. Bobikov A. I., Bubnova T. S. *Tuning filtered Smith predictor parameters with Simulink Response optimization*. Vestnik Rjazanskogo gosudarstvennogo radiotekhnicheskogo universiteta [Bulletin of the Ryazan State Radio Engineering University], 2017, no. 61, pp. 96–104 (In Russ.).
7. Oshhepkov, A. Ju. *Sistemy avtomaticheskogo upravlenija: teorija, primenenie, modelirovanie v MATLAB. Automatic Control Systems: Theory, Application, Modeling in MATLAB*. Saint Petersburg, Lan', 2013, 208 p. (In Russ.).

Информация об авторе

Сеньков Андрей Григорьевич, кандидат технических наук, доцент, заведующий кафедрой автоматизированных систем управления производством, Белорусский государственный аграрный технический университет.

Information about the author

Andrey G. Senkov, Cand. Sci. (Eng.), Assistant Professor, Head of the Department of Automated Production Control Systems, Belarusian State Agrarian Technical University.

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION



УДК 004.94 + 534.77
<https://doi.org/10.37661/1816-0301-2021-18-3-68-82>

Оригинальная статья
Original Paper

Метод моделирования эффекта ухудшения частотного разрешения слуха у больных нейросенсорной тугоухостью

М. И. Порхун[✉], М. И. Вашкевич

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

Аннотация. Предложен метод моделирования эффекта ухудшения частотного разрешения уха у больных нейросенсорной тугоухостью, особенностью которого является возможность его настройки по аудиограмме конкретного человека. Метод основан на пофреймовой обработке сигнала в частотной области. Моделирование эффекта ухудшения частотного разрешения уха достигается за счет обработки составляющих амплитудного спектра исходного звукового сигнала «размывающей» функцией. «Размывающая» функция формируется из амплитудно-частотных характеристик слуховых фильтров, ширина полос которых определяется исходя из аудиограммы тугоухого человека. Предложенный метод реализован в среде MATLAB. Проведено экспериментальное исследование влияния эффекта ухудшения частотного разрешения уха с использованием теста на разборчивость речи. В эксперименте участвовало 15 человек, которым давали прослушивать записи, обработанные предложенным методом с различными настройками параметрами, в том числе с добавлением белого шума и без добавления. Экспериментальные данные показали, что ухудшение частотного разрешения уха приводит к ухудшению разборчивости речи, особенно сильному при наличии фонового шума. На основании ответов участников эксперимента составлены таблицы спутывания звуков, отражающие факт неразличимости схожих по частоте звуков, что подтверждает корректность работы предложенного метода.

Ключевые слова: слуховые патологии, частотное разрешение уха, моделирование потери слуха, гоэфильтер, разборчивость речи

Для цитирования. Порхун, М. И. Метод моделирования эффекта ухудшения частотного разрешения слуха у больных нейросенсорной тугоухостью / М. И. Порхун, М. И. Вашкевич // Информатика. – 2021. – Т. 18, № 3. – С. 68–82. <https://doi.org/10.37661/1816-0301-2021-18-3-68-82>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 09.04.2021
Подписана в печать | Accepted 07.06.2021
Опубликована | Published 29.09.2021

A method for simulation the effect of the reduced frequency resolution of the ear in patients with sensorineural hearing loss

Maxim I. Porhun[✉], Maxim I. Vashkevich

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

[✉]E-mail: porhun@bsuir.by

Abstract. A method for the simulation of reduced frequency resolution of the ear in patients with sensorineural hearing loss is proposed. The method is based upon the ability to adjust it according to the audiogram of a concrete person by frame-by-frame signal processing in the frequency domain. Simulation of the effect of the reduced frequency resolution of the ear is achieved by processing the components of amplitude spectrum of the original sound signal by the "smearing" function. The "smearing" function is formed from the amplitude-frequency characteristics of the auditory filters, which bandwidth is determined by the audiogram of the deaf person. The proposed method is implemented in the MATLAB. An experimental study of the effect of the reduced frequency resolution of the ear using the speech intelligibility test was conducted. The experiment involved 15 people who listened the records processed by the proposed method with various settings and noise conditions. Experimental data have shown that reduced frequency resolution of the ear leads to the deterioration in speech intelligibility, especially in the presence of background noise. Based on the answers of the participants of the experiment, the confusion tables of sounds were compiled, reflecting the fact of indistinguishability of sounds similar in frequency, that confirms the correctness of the proposed method.

Keywords: hearing pathologies, ear frequency resolution, hearing loss simulation, roex-filter, speech intelligibility

For citation. Porhun M. I., Vashkevich M. I. A method for simulation the effect of the reduced frequency resolution of the ear in patients with sensorineural hearing loss. *Informatics*, 2021, vol. 18, no. 3, pp. 68–82 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-68-82>

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

Введение. Согласно данным Всемирной организации здравоохранения, нейросенсорная тугоухость является одним из самых распространенных заболеваний слуховой системы человека. Она характеризуется рядом изменений в органах слуха: ухудшением частотного разрешения (частотной избирательности) уха, феноменом усиленного нарастания громкости (рекруитмент), сужением слышимого динамического диапазона, а также снижением порогов слышимости [1].

Разработка методов моделирования потери слуха преследует несколько целей. Во-первых, модели потери слуха применяются для демонстрации перцептуальных, коммуникативных и социальных аспектов тугоухости для нормально слышащих людей. Во-вторых, такие модели используются для обучения нормально слышащих людей восприятию звука в том виде, в котором тугоухие люди его слышат. Это, в свою очередь, помогает людям с нормальным слухом проще контактировать со слабослышащими [2]. В-третьих, моделирование потери слуха применяется при проверке эффективности методов коррекции слуха без непосредственного участия слабослышащих людей [3, 4].

Тугоухость включает в себя целый ряд аспектов [1], поэтому метод моделирования должен отражать наиболее значимые из них. Существующие методы зачастую не учитывают эффект ухудшения частотного разрешения уха либо не могут автоматически настраиваться согласно параметрам слуха конкретного слабослышащего человека [5, 6]. В настоящей работе предлагается метод моделирования потери слуха, учитывающий ухудшение частотного разрешения уха и имеющий возможность настройки по аудиограмме слабослышащего человека.

Моделирование эффекта ухудшения частотного разрешения уха. Система, выполняющая моделирование эффекта ухудшения частотного разрешения уха, изображена на рис. 1.

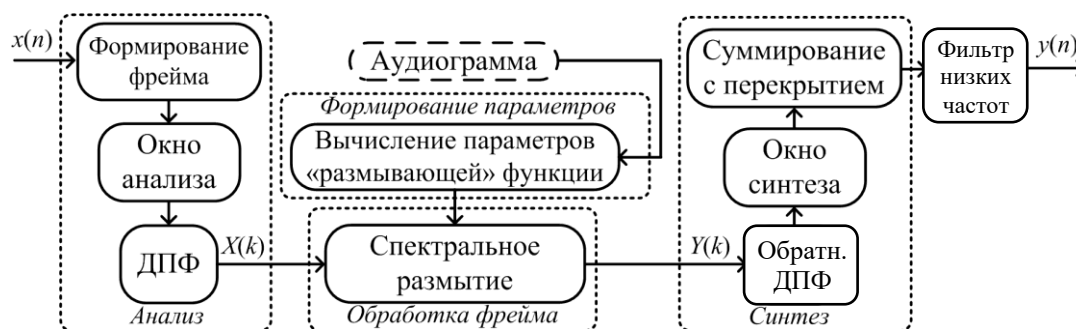


Рис. 1. Структура метода моделирования эффекта ухудшения частотного разрешения уха

Fig. 1. Structure of the method for simulation the effect of reduced frequency resolution of ear

Входной сигнал $x(n)$ разделяется на перекрывающиеся фреймы, которые умножаются на окно анализа, и переводится в частотную область при помощи дискретного преобразования Фурье (ДПФ). Далее Фурье-образ входного фрейма $X(k)$ поступает в блок спектрального размытия, где выполняется моделирование эффекта ухудшения частотного разрешения уха. Общий принцип моделирования эффекта спектрального размытия описан в работе [7]. Полученный в результате обработки Фурье-образ $Y(k)$ переводится обратно во временную область при помощи обратного ДПФ и умножается на окно синтеза. Обработанный сигнал синтезируется при помощи метода суммирования с перекрытием [8]. Последним этапом формирования выходного сигнала $y(n)$ является пропускание его через фильтр низких частот, задача которого – уменьшение искажений, возникающих в высокочастотной области в результате обработки сигнала.

Структура блока, выполняющего спектральное размытие, изображена на рис. 2.

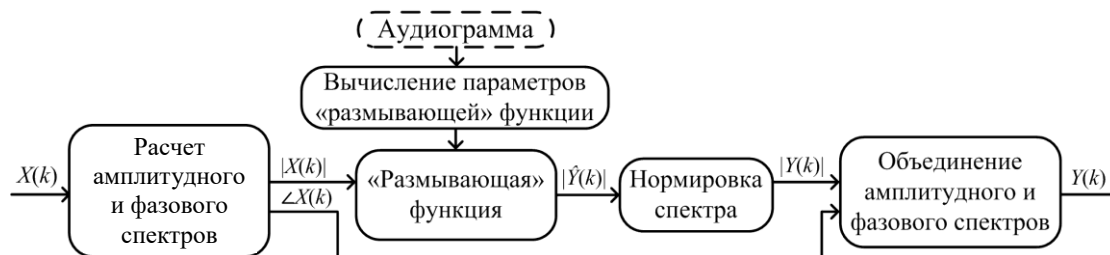


Рис. 2. Структура блока спектрального размытия

Fig. 2. Spectral smearing block structure

На вход блока поступает спектр текущего фрейма $X(k)$. Спектр раскладывается на амплитудную и фазовую составляющие. Амплитудный спектр обрабатывается «размывающей» функцией, а фазовый остается без изменений. Параметры «размывающей» функции извлекаются из аудиограммы тугоухого человека. Далее выполняется нормировка модифицированного амплитудного спектра по уровню мощности исходного спектра. Выход $Y(k)$ формируется путем объединения модифицированного амплитудного и исходного фазового спектров.

Спектральное размытие. У больных нейросенсорной тугоухостью, как правило, повреждены рецепторы улитки (наружные и внутренние волосковые клетки), преобразующие звуковые колебания в электрические импульсы. При этом поражение наружных волосковых клеток, обеспечивающих избирательную реакцию отделов улитки на звуки разных частот, приводит к нарушению анализа звуков по частоте и громкости. Эффект ухудшения частотного разрешения уха заключается в расширении частотной характеристики слуховых фильтров и, как следствие, размытии спектральных составляющих сигнала по оси частот. Снижение частотного раз-

решения уха сильно влияет на разборчивость речи в целом. Как следствие, человек не ощущает разницы между схожими по частоте звуками. Например, он не будет слышать разницы в произношении слов «стол» и «стул» [1].

При моделировании эффекта спектрального размытия применяется модель слуховых фильтров [7, 9], называемая гоех-фильтром (англ. ROUNDED-EXponential filter) [10]. Роех-фильтр позволяет моделировать слуховые фильтры как человека с нормальным слухом, так и слабослышащего [10]. Амплитудно-частотная характеристика (АЧХ) гоех-фильтра имеет симметричный вид относительно центральной частоты f_c и описывается выражением [10, 11]

$$A(f) = \left(1 + p \frac{|f - f_c|}{f_c} \right) \exp \left(-p \frac{|f - f_c|}{f_c} \right), \quad (1)$$

где f – частота; p – параметр, определяющий крутизну подъема и спада АЧХ фильтра.

Параметр p определяется формулой

$$p = \frac{4f_c}{ERB(f_c)}, \quad (2)$$

где ERB – эквивалентная прямоугольная полоса пропускания фильтра:

$$ERB(f_c) = 24,637\alpha(0,00437f_c + 1). \quad (3)$$

Здесь α – параметр, определяющий степень расширения полосы пропускания слухового фильтра.

При моделировании эффекта ухудшения частотного разрешения уха АЧХ гоех-фильтров используются для формирования «шаблонов» слуха нормально слышащего и тугоухого человека. Необходимость использования характеристик фильтров нормально слышащего человека обоснована ниже. Коэффициенты α в равенстве (3) выступают в качестве настроечных параметров для модели слуховых фильтров тугоухого человека, их значения определяются на основе анализа аудиограммы. Для нормально слышащего человека коэффициенты α принимаются равными единице [7]. Расчет параметров α описывается в следующем разделе.

Определим исходный (\mathbf{x}) и обработанный (\mathbf{y}) фреймы в блоке спектрального размытия в виде векторов

$$\mathbf{x} = [X(1), X(2), \dots, X(K)]^T, \quad \mathbf{y} = [Y(1), Y(2), \dots, Y(K)]^T, \quad (4)$$

где $X(k)$ и $Y(k)$ – отсчеты амплитудного спектра исходного и обработанного фреймов, K – количество точек на частотной оси:

$$K = (M / 2) + 1. \quad (5)$$

Здесь M – размер фрейма.

Для вычисления слухового образа, формирующегося у слабослышащего человека, необходимо умножить амплитудный спектр входного сигнала на АЧХ гребенки гоех-фильтров, моделирующих эффект ухудшения частотного разрешения уха:

$$\mathbf{y} = \mathbf{A}_1 \mathbf{x}, \quad (6)$$

где \mathbf{A}_1 – матрица, содержащая в строках АЧХ гоех-фильтров, моделирующих слуховые фильтры слабослышащего человека. Центральные частоты гоех-фильтров совпадают с частотной сеткой ДПФ-анализа, используемого в структуре на рис. 1.

В выражении (6) не учитывается тот факт, что обработанные спектральные составляющие сигнала y впоследствии будут восприниматься человеком с нормальным слухом, т. е. будут отображаться на нормальные слуховые фильтры. Чтобы компенсировать работу этих фильтров, в статье [7] предложено моделировать спектральное размытие при помощи выражения

$$y = A_N^{-1} A_1 x, \quad (7)$$

где A_N – матрица, аналогичная A_1 , но содержащая АЧХ гоех-фильтров, моделирующих слуховые фильтры человека с нормальным слухом.

Спектральные компоненты, сформированные по выражению (7) после прохождения через слуховые фильтры человека с нормальным слухом, в итоге дадут слуховой образ $A_1 x$, который соответствует тому, что слышит тугоухий человек.

Согласно выражению (3) при увеличении центральной частоты f_c увеличивается и ERB , что приводит к увеличению АЧХ гоех-фильтров (см. выражение (1)). Поскольку обработанный сигнал будет проходить через слуховые фильтры нормально слышащего человека, предлагается нормировать АЧХ фильтров путем их деления на каждой частоте f_c на ERB . Аналогичным образом для характеристик расширенных гоех-фильтров, соответствующих слуховой патологии, необходимо выполнить нормировку. Однако в данном случае нормировка выполняется путем деления АЧХ расширенных фильтров на каждой частоте f_c на коэффициент $ERB \cdot \alpha$. Различия в нормировочных коэффициентах связаны с тем, что ERB расширенных фильтров увеличиваются в α раз [7].

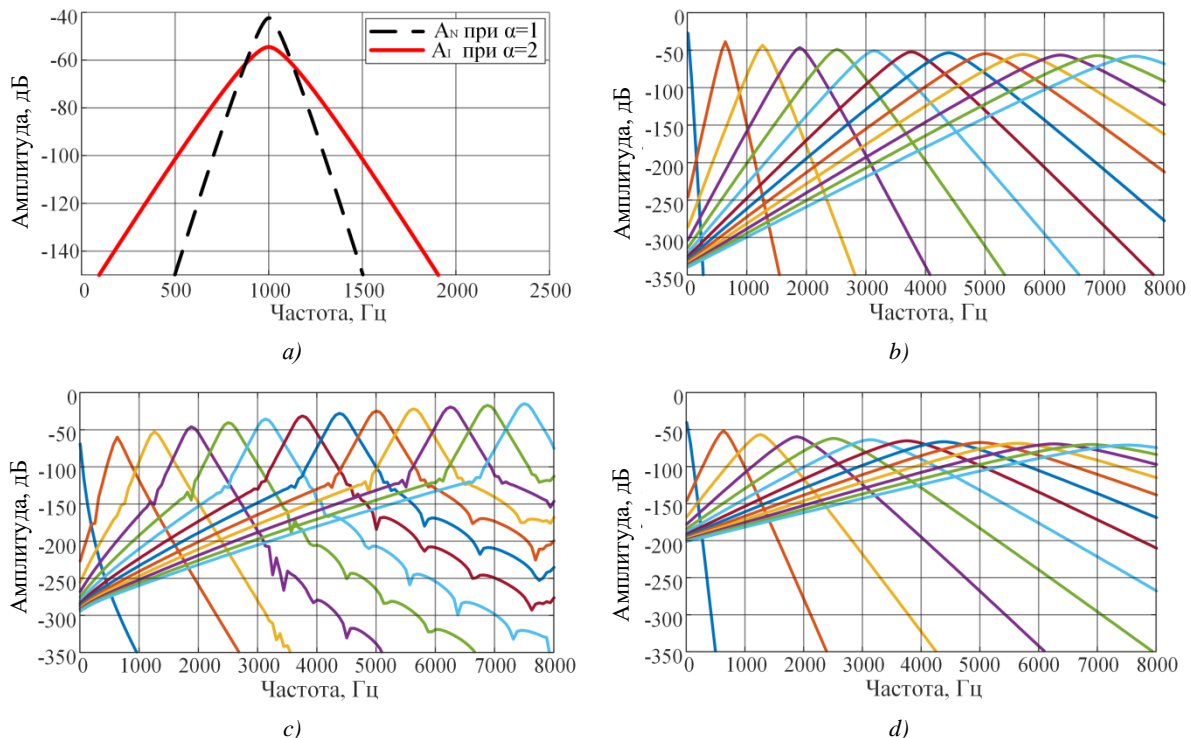


Рис. 3. Примеры АЧХ гоех-фильтров: а) АЧХ на центральной частоте 1000 Гц (нормальный слух и потеря слуха); б) фильтры, соответствующие матрице A_N ; в) фильтры, соответствующие матрице A_N^{-1} ; д) фильтры, соответствующие матрице A_1

Fig. 3. Roex-filters frequency response examples: a) frequency response at the central frequency 1 kHz (normal hearing and hearing loss); b) filters corresponding to matrix A_N ; c) filters corresponding to matrix A_N^{-1} ; d) filters corresponding to matrix A_1

На рис. 3, *a* показан эффект расширения полосы пропускания гоех-фильтра. АЧХ без расширения полосы пропускания соответствует нормальному слуху (в выражении (3) $\alpha = 1$). Слуховой патологии соответствует АЧХ с расширенной полосой пропускания (в выражении (3) $\alpha = 2$). На рис. 3, *b-d* приведены примеры АЧХ гоех-фильтров, входящих в матрицы \mathbf{A}_N , \mathbf{A}_N^{-1} и \mathbf{A}_I для $\alpha=2$ (показан каждый 10-й фильтр). На рис. 3, *b, d* видно, что при увеличении параметра α АЧХ гоех-фильтров расширяются. При этом с увеличением частоты пиковое значение амплитуды постепенно уменьшается. На рис. 3, *c* изображены фильтры, входящие в состав матрицы \mathbf{A}_N^{-1} . С увеличением частоты максимальное значение амплитуды этих фильтров увеличивается, что компенсирует уменьшение пикового значения АЧХ фильтров из матрицы \mathbf{A}_N .

Примеры обработки фрейма входного сигнала блоком спектрального размытия для гласного звука /а/ и согласного /ш/ приведены на рис. 4.

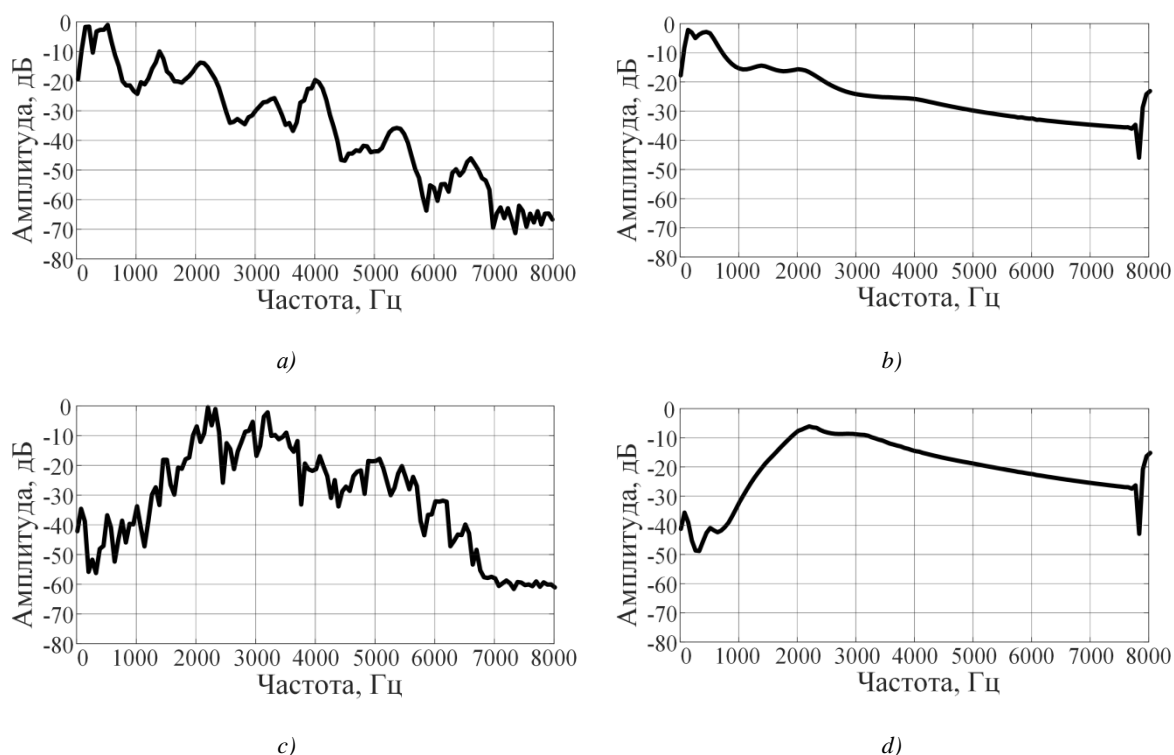


Рис. 4. Примеры обработки фрейма блоком спектрального размытия: *a*) исходный фрейм для звука /а/; *b*) обработанный фрейм для звука /а/; *c*) исходный фрейм для звука /ш/; *d*) обработанный фрейм для звука /ш/

Fig. 4. Examples of frame processing with spectral smearing block: a) original frame for syllable /a/; b) processed frame for syllable /a/; c) original frame for syllable /u/; d) processed frame for syllable /u/

На рис. 3 и 4 видно, что в результате обработки амплитудного спектра спектральные составляющие исходного сигнала «размываются» вдоль частот оси. Также в высокочастотной области появляются искажения (артефакты), которые после синтеза выходного сигнала устраняются посредством фильтрации (см. рис. 1).

Анализ аудиограммы. Важным моментом является настройка метода моделирования эффекта ухудшения частотного разрешения уха согласно характеристикам слуха тугоухого человека. Наиболее простым методом оценки слуха человека служит аудиометрия [12, 13]. Результатом аудиометрии является аудиограмма, отражающая пороги слышимости человека на заданном наборе частот [1]. Аудиометрия представляет собой наиболее распространенный метод оценки слуха. По этой причине настройку метода моделирования эффекта ухудшения частотного разрешения уха предлагается выполнять по аудиограмме слабослышащего человека.

В работах [13–17] описана связь между уровнем потери слуха и шириной полосы пропускания слуховых фильтров на заданной частоте. Так, в работах [13–15, 17] авторы предлагают выбирать значение параметра α в зависимости от степени потери слуха (табл. 1). Степень потери слуха определяется как среднее арифметическое между значениями порогов слышимости (по воздушной проводимости) на частотах 500, 1000, 2000 и 4000 Гц [1]. Существенным недостатком такого подхода является слишком неточный выбор параметра α , учитывающий только степень потери слуха, а не уровень потери слуха на заданных частотах.

Таблица 1. Зависимость параметра α от степени потери слуха

Table 1. Dependence of the parameter α on hearing loss degree

Степень потери слуха <i>Hearing loss degree</i>	Средний порог слышимости, дБ <i>Average hearing threshold, dB</i>	Значение параметра α в выражении (3) <i>Parameter α value in expression (3)</i>
Нет потери слуха	<26	1
Первая	26–40	3
Вторая	41–55	
Третья	56–70	
Четвертая	71–90	6

В исследовании [16] установлено, что параметр α , отвечающий за расширение полосы пропускания слухового фильтра, зависит как от центральной частоты f_c , так и от уровня потери слуха на данной частоте HL и описывается выражением

$$\alpha(f_c, HL) = \begin{cases} 1, & \text{если } HL \leq 22, \forall f_c, \\ 10^{0,01348(HL-22)/(1-0,3551g(f_c/1000))}, & \text{если } 22 < HL < 65 \text{ при } f_c \leq 1000, \\ 10^{0,01348(HL-22)}, & \text{если } 22 < HL < 65 \text{ при } f_c > 1000, \end{cases} \quad (8)$$

где f_c – центральная частота гоух-фильтра, HL – уровень потери слуха на частоте f_c .

Уровень потери слуха HL определяется путем интерполяции аудиограммы на частотную сетку центральных частот гоух-фильтров. Из выражения (8) следует, что при потере слуха менее 22 дБ расширения полосы слуховых фильтров не происходит ($\alpha = 1$). Различия в выражениях для вычисления параметра α при частотах выше и ниже 1000 Гц отражают факт менее выраженного расширения слуховых фильтров на низких частотах [16].

Недостатком выражения (8) является то, что оно позволяет определить параметр α только для уровня потери слуха менее 65 дБ, т. е. охватывает первую, вторую и частично третью степени тугоухости и полностью игнорирует четвертую степень (табл. 1).

Таким образом, в существующих исследованиях авторы по-разному описывают связь между уровнем потери слуха и шириной полосы пропускания слухового фильтра. В настоящей работе предлагается доопределить выражение (8), предложенное в [16], используя данные табл. 1 [13–15, 17].

Выражение (8) описывает зависимость параметра α от частоты при уровне потери слуха менее 65 дБ. При потере слуха от 65 до 90 дБ предлагается находить параметр α путем линейной интерполяции (рис. 5).

Согласно предлагаемому подходу при потере слуха 65 дБ значение α фиксируется на уровне, определяемом формулой (8), потере 90 дБ (что соответствует четвертой степени тугоухости) сопоставляется значение $\alpha = 6$ [13–15, 17]. Для промежуточных уровней потери слуха от 65 до 90 дБ значения α вычисляются по уравнению прямой, проходящей через указанные крайние точки:

$$\alpha^*(f_c, HL) = 0,24HL + \alpha(f_c, 65)(3,6 - 0,04HL) - 15,6. \quad (9)$$

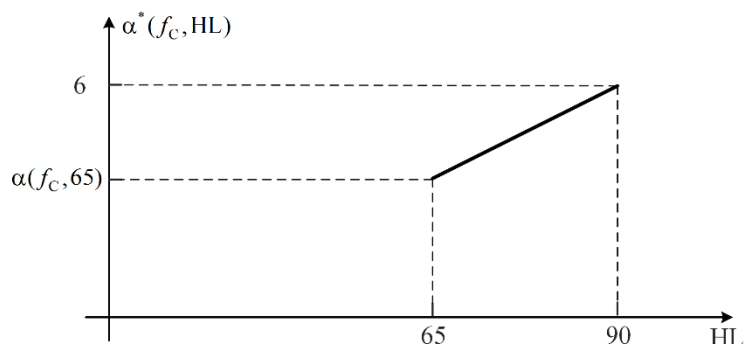


Рис. 5. Значения параметра α при уровне потери слуха от 65 до 90 дБ
 Fig. 5. Dependence of the parameter α at hearing loss level from 65 to 90 dB

Кривая зависимости α от частоты может иметь резкие перепады. С целью сглаживания графика зависимости $\alpha(f_c)$ и, как итог, внесения меньших искажений в сигнал используется процедура усреднения по семи точкам. На рис. 6 показан пример исходной и сглаженной зависимости $\alpha(f_c)$. Видно, что применение усреднения позволило сгладить резкие переходы в графике зависимости $\alpha(f_c)$.

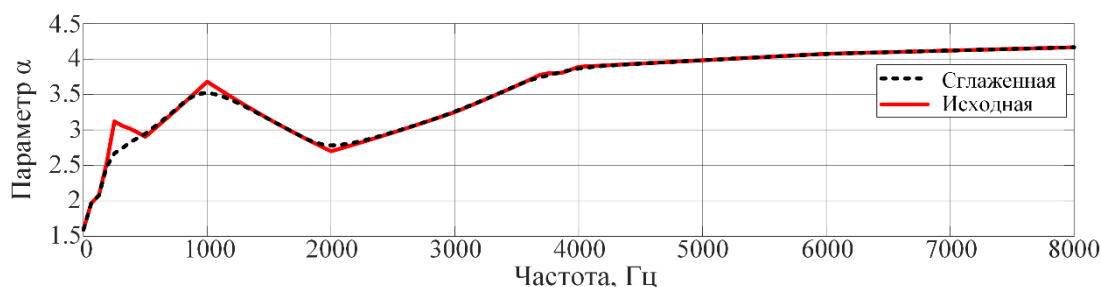


Рис. 6. Значения параметра α в зависимости от частоты
 Fig. 6. Frequency dependence of the parameter α

Таким образом, определение параметра α выполняется в три этапа (рис. 7). На первом этапе аудиограмма интерполируется на сетку центральных частот гоех-фильтров, которая совпадает с частотной сеткой ДПФ. Далее определяется значение уровня потери слуха HL на каждой частоте f_c . На втором этапе вычисляются значения параметра α на частотах f_c . Когда уровень потери слуха не превышает 65 дБ, используется выражение (8). При потере 65 дБ и более параметр α вычисляется по формуле (9). На третьем этапе выполняется усреднение значений параметра α .

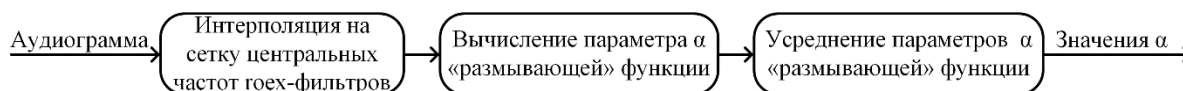


Рис. 7. Структура блока вычисления параметров «размывающей» функции
 Fig. 7. Structure of the block for calculating the parameters of the "smearing" function

Проверка метода моделирования эффекта ухудшения частотного разрешения уха. Предлагаемый метод реализован в среде MATLAB и протестирован на речевом сигнале. Целью эксперимента являлась качественная оценка того, как изменится речь при моделировании различных степеней тугоухости. Для этого тестовый сигнал обрабатывался предложенным мето-

дом, настроенным по аудиограммам, которые соответствуют третьей и четвертой степеням тугоухости (рис. 8, *d*). Спектрограммы исходного и обработанных сигналов показаны на рис. 8, *a–c*.

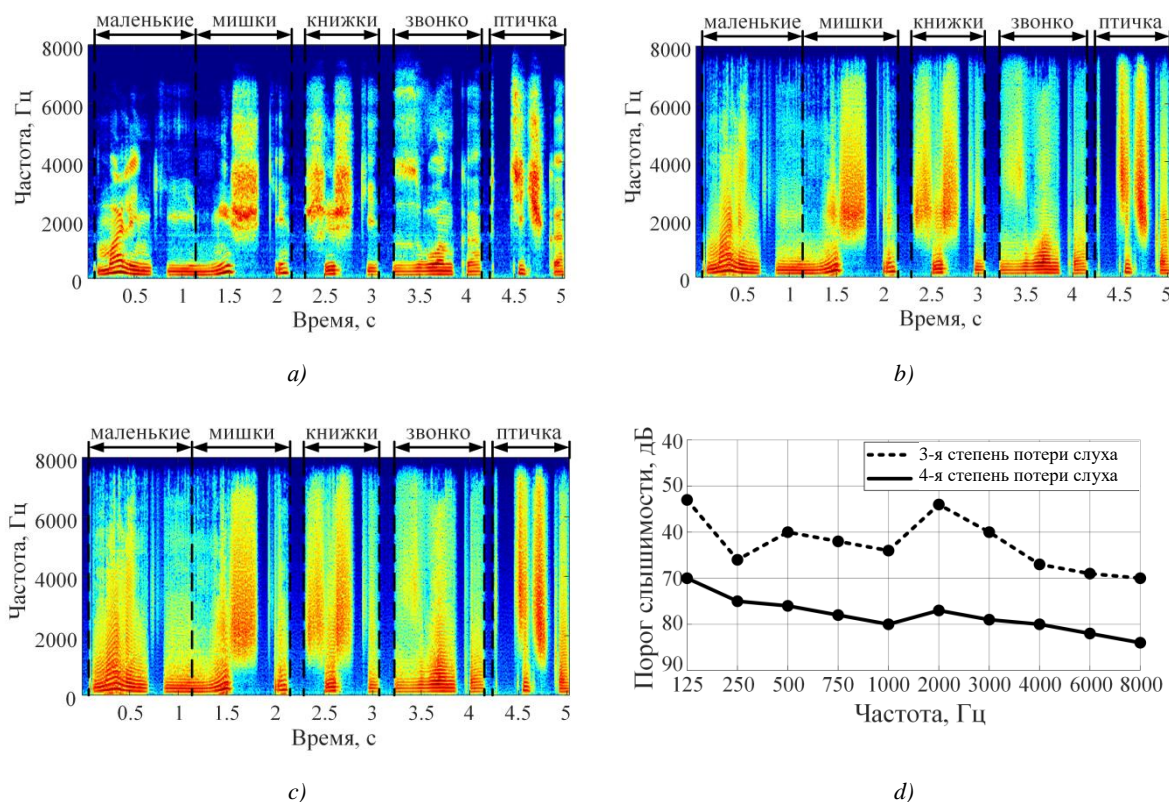


Рис. 8. Результаты экспериментов: *a*) спектрограмма исходного сигнала; *b*) спектрограмма обработанного сигнала для аудиограммы при третьей степени потери слуха; *c*) спектрограмма обработанного сигнала для аудиограммы при четвертой степени потери слуха; *d*) аудиограмма при третьей и четвертой степенях потери слуха

Fig. 8. Experimental results: *a*) spectrogram of original signal; *b*) spectrogram of processed signal for an audiogram at moderate hearing loss; *c*) spectrogram of processed signal for an audiogram at moderately severe hearing loss; *d*) an audiogram at moderate and moderately severe hearing loss

Из полученных спектрограмм видно, что при увеличении уровня потери слуха увеличивается и степень размытия спектральных компонентов звукового сигнала по частотной оси, что соответствует эффекту ухудшения частотного разрешения уха у больных нейросенсорной тугоухостью.

Экспериментальные исследования. Исследования проводились с целью определения влияния моделируемой степени потери слуха на разборчивость речевого сигнала (при наличии и отсутствии фонового шума). Для определения разборчивости речи использовался тест на основе артикуляционных таблиц слов из ГОСТ 16600 – 72. Согласно указанному стандарту для определения разборчивости речи участник эксперимента должен прослушать артикуляционную таблицу, состоящую из 50 слов. После каждого слова делается пауза, во время которой участник повторяет услышанное слово. На основании числа правильно повторенных слов определяется процент разборчивости речи [1].

Для проведения эксперимента были подготовлены 49 тестовых записей (см. рис. 9), каждая из которых содержала одну артикуляционную таблицу из ГОСТ 16600 – 72. Тестовые записи разделены на семь групп по семь записей каждая (табл. 2). Различные группы отличаются моделируемой степенью потери слуха и уровнем добавленного шума (использовалось отношение сигнал/шум (ОСШ) 0 и 5 дБ).

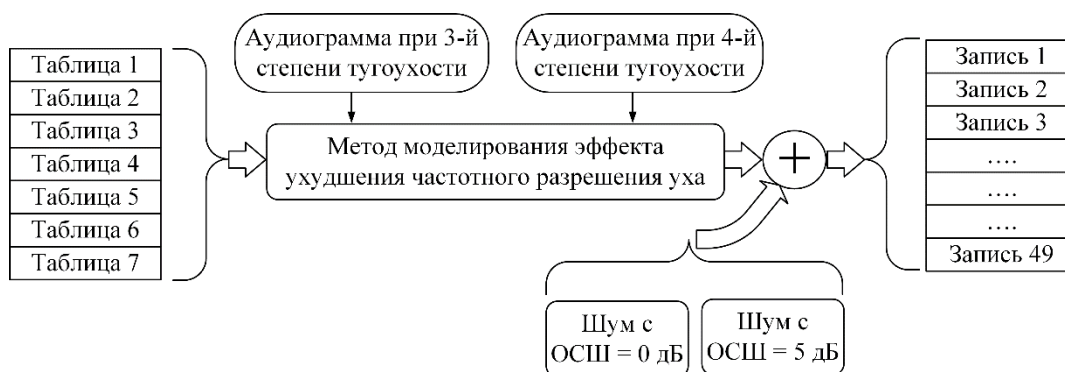


Рис. 9. Этапы подготовки тестовых записей
 Fig. 9. Test records preparation stages

Таблица 2. Группы тестовых записей
 Table 2. Test recordings groups

Группа Group	Параметры Settings	Уровень добавленного шума Added noise level	Условное обозначение Symbolic name
1	Без обработки	Без добавления шума	Norm
2	Аудиограмма при третьей степени потери слуха	Без добавления шума	HL3
3		ОСШ = 5 дБ	HL3 + 5dB
4		ОСШ = 0 дБ	HL3 + 0dB
5	Аудиограмма при четвертой степени потери слуха	Без добавления шума	HL4
6		ОСШ = 5 дБ	HL4 + 5dB
7		ОСШ = 0 дБ	HL4 + 0dB

В эксперименте участвовали 15 нормально слышащих людей в возрасте от 17 до 25 лет. Каждый участник эксперимента по очереди прослушивал семь тестовых записей – по одной из каждой группы. Эксперимент проводился таким образом, что каждый участник прослушивал уникальный набор тестовых записей. Полученные в результате данные о разборчивости речи для каждой группы тестовых записей представлены в виде диаграмм размаха на рис. 10. Диаграммы отражают изменение процента разборчивости речи в процентах в зависимости от моделируемой степени потери слуха и от уровня добавленного шума.

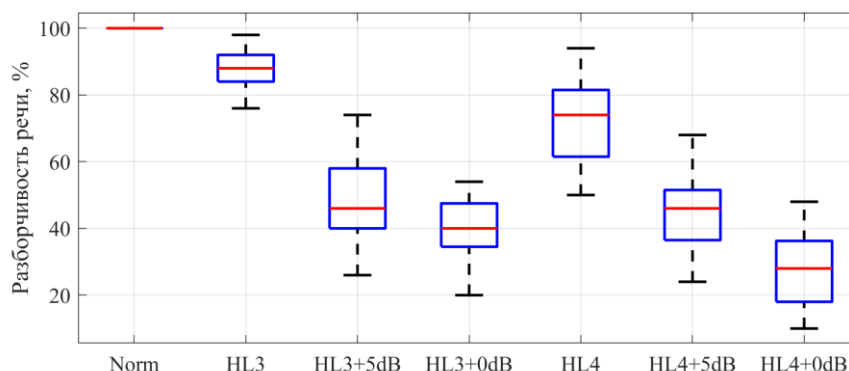


Рис. 10. Диаграммы размаха изменения разборчивости речи
 Fig. 10. Boxplots of speech intelligibility

На основании экспериментальных данных составлена табл. 3, где приведены средние, максимальные и минимальные значения разборчивости речи для каждой группы тестовых записей.

Таблица 3. Изменение разборчивости речи

Table 3. Speech intelligibility variation

Параметры обработки сигнала <i>Signal processing parameters</i>	Разборчивость речи, % <i>Speech intelligibility, %</i>		
	минимальная <i>minimum</i>	максимальная <i>maximum</i>	средняя <i>average</i>
Norm	100	100	100
HL3	76	98	88
HL3+0dB	20	54	39
HL3+5dB	26	74	48
HL4	50	94	72
HL4+0dB	10	48	28
HL4+5dB	24	68	44

Результаты эксперимента показывают, что при ухудшении частотного разрешения уха вследствие увеличения степени потери слуха снижается и разборчивость речи. При росте интенсивности уровня добавленного шума к тестовым сигналам разборчивость речи резко снижается.

В процессе тестирования фиксировалась не только правильность произнесенных слов, но и сами слова, что позволило провести анализ ошибок. На основании полученных результатов были составлены таблицы спутывания звуков (табл. 4–9), наиболее часто встречающихся в ответах участников при обработке звука предложенным методом с различными настроечными параметрами. Жирным шрифтом выделены самые часто спутываемые звуки.

Таблица 4. Спутывание звуков при обработке методом HL3

Table 4. Syllables mixing up after processing with HL3 method

		Должно быть произнесено <i>Must be pronounced</i>		
		/м/	/б/	/л/
Произнесено <i>Pronounced</i>	/н/	6	0	0
	/н'/	3	0	7
	/в/	0	4	0
	/й/	0	0	4

Таблица 5. Спутывание звуков при обработке методом HL3+5dB

Table 5. Syllables mixing up after processing with HL3+5dB method

		Должно быть произнесено <i>Must be pronounced</i>											
		/в/	/й/	/б/	/т/	/д/	/л/	/с/	/п/	/н/	/м/	/л'/	/п'/
Произнесено <i>Pronounced</i>	/в/	–	0	6	0	2	4	0	2	0	0	0	0
	/й/	0	–	0	0	0	0	0	0	0	0	0	7
	/б/	1	0	–	2	5	0	1	0	1	0	0	0
	/т/	0	0	1	–	1	8	6	6	0	0	0	0
	/т'/	0	0	0	4	0	0	0	0	0	0	0	0
	/д/	1	0	1	4	–	1	2	0	0	0	0	0
	/л/	6	0	0	1	1	–	18	14	0	0	0	0
	/с/	0	0	0	3	0	1	–	1	0	0	0	0
	/к/	0	0	0	7	1	3	0	0	1	0	1	0
	/п/	0	0	1	12	2	1	2	–	0	0	0	0
	/н'/	0	9	0	0	0	0	0	0	6	0	5	0
	/н/	0	0	0	0	0	0	0	0	–	6	1	0
/м/	0	0	0	0	0	0	0	0	6	–	1	0	

Таблица 6. Спутывание звуков при обработке методом HL3+0dB
 Table 6. Syllables mixing up after processing with HL3+0dB method

		Должно быть произнесено Must be pronounced												
		/_/_/	/с/	/м/	/н/	/п/	/т/	/л'/	/д'/	/в/	/ш/	/й'/	/з'/	/ф/
Произнесено Pronounced	/_/_/	–	15	0	1	6	3	0	5	8	1	0	0	0
	/к/	3	3	1	2	6	9	0	4	0	0	0	0	0
	/с/	2	–	0	0	0	11	0	0	0	4	0	0	0
	/м/	0	0	–	4	0	0	0	0	0	0	0	0	0
	/т/	0	4	0	0	0	1	0	0	0	1	0	4	0
	/н/	0	0	5	–	0	0	1	0	0	0	1	0	0
	/п/	0	3	0	0	–	10	0	0	0	2	0	0	0
	/т'/	8	10	0	0	9	–	0	0	0	0	0	0	1
	/л'/	0	0	0	0	0	0	–	0	0	0	7	0	0
	/д'/	1	2	0	0	0	4	0	0	1	0	0	0	6
/н'/	0	0	0	8	0	0	6	0	0	0	2	0	0	
/з/	0	2	0	0	0	0	0	8	1	1	0	0	0	

Таблица 7. Спутывание звуков при обработке методом HL4
 Table 7. Syllables mixing up after processing with HL4 method

		Должно быть произнесено Must be pronounced					
		/т/	/к/	/м/	/л'/	/д'/	/б/
Произнесено Pronounced	/п/	3	11	0	0	0	0
	/н/	0	0	8	0	0	0
	/н'/	0	0	3	7	0	0
	/д'/	5	0	0	0	1	2
	/т'/	–	4	0	0	6	1
	/в/	0	0	2	0	6	8

Таблица 8. Спутывание звуков при обработке методом HL4+5dB
 Table 8. Syllables mixing up after processing with HL4+5dB method

		Должно быть произнесено Must be pronounced												
		/_/_/	/н/	/т/	/д'/	/к/	/л'/	/с/	/б/	/п/	/з'/	/м/	/м'/	/й'/
Произнесено Pronounced	/_/_/	–	0	2	3	3	0	17	1	1	0	0	0	0
	/н/	0	–	0	0	0	1	0	0	0	0	13	0	0
	/т/	5	0	0	0	5	0	14	0	0	0	0	0	0
	/д'/	1	0	9	–	3	0	2	4	0	0	0	0	0
	/н'/	0	7	0	0	0	7	0	0	0	0	10	4	2
	/в/	2	0	0	4	0	0	1	10	0	0	2	0	0
	/к/	4	0	2	3	–	0	0	0	5	0	0	0	0
	/т'/	0	0	5	0	0	0	0	0	0	6	0	0	0
	/л'/	0	0	0	0	0	–	0	0	0	0	0	0	8
	/з/	0	0	0	6	0	0	0	0	0	0	0	0	0

Таблица 9. Спутывание звуков при обработке методом HL4+0dB
 Table 9. Syllables mixing up after processing with HL4+0dB method

		Должно быть произнесено Must be pronounced											
		/т/	/с/	/_/_/	/к/	/п/	/м/	/н/	/й'/	/б/	/в/	/з/	/л'/
Произнесено Pronounced	/в/	0	2	3	0	0	1	0	0	11	0	1	0
	/т/	–	15	7	0	5	0	0	0	0	0	0	0
	/с/	10	–	2	0	0	0	0	0	0	0	0	0
	/н'/	0	0	0	0	0	3	7	7	0	4	0	2
	/_/_/	8	11	–	9	1	0	0	0	0	0	0	0
	/к/	12	9	5	–	4	0	1	0	0	0	0	0
	/п/	2	4	0	1	–	0	2	0	1	0	0	0
	/м/	0	0	0	0	0	–	5	0	0	0	0	0
	/н/	0	0	0	0	0	6	–	0	0	0	0	3
	/д'/	2	0	0	1	0	0	0	0	1	0	8	0
/й'/	0	0	0	0	0	0	0	–	0	0	0	4	

Чаще всего при произношении услышанных в тестовых сигналах слов участники путали между собой звуки, близкие по частоте, например: ма/т'/ – ма/з'/, /п/оп – /т/оп, па/с/ – па/т'/, /т/ук – /с/ук, /т/ру/д/ – /п/ру/т/, бра/к/ – бра/т/, ча/т'/ – ча/с/. Важно отметить, что довольно часто участники эксперимента пропускали в начале и конце слов согласные звуки, как правило шипящие, например: / /кот – /с/кот, / /дай – /с/дай, / /стань – /в/стань, бор/т/ – бор/ /, / /ши/т'/ – /в/ши/т'/, сер/ / – сер/п/. Нередко участники путали между собой твердые и мягкие звуки, например: жес/т'/ – жес/т'/, бы/л/ – бы/л'/, ко/н/ – ко/н'/, а также мягкие согласные звуки и звук /й/ и наоборот, например: мо/л'/ – мо/й'/, но/й'/ – но/л'/, да/л'/ – да/й'/.

Таким образом, результаты эксперимента показывают, что чаще всего участники путали между собой схожие по частоте звуки и не слышали шипящие звуки в словах. Это, в свою очередь, соответствует природе эффекта ухудшения частотного разрешения уха [1].

Заключение. В работе предложен метод моделирования эффекта ухудшения частотного разрешения уха у больных нейросенсорной тугоухостью, основанный на применении частотного размытия с использованием слуховых фильтров. Особенностью модели является возможность ее настройки с использованием аудиограммы слабослышащего человека. Метод реализован в среде MATLAB. Проведены экспериментальные исследования изменения разборчивости речи при использовании предложенного метода. Результаты эксперимента показали, что при ухудшении частотного разрешения уха вследствие увеличения степени потери слуха разборчивость речи снижается, особенно при наличии в звуке фонового шума. На основании анализа экспериментальных данных выявлена зависимость между спутываемыми звуками, что указывает на корректность и работоспособность метода моделирования эффекта ухудшения частотного разрешения уха.

Вклад авторов. М. И. Порхун разработал метод моделирования эффекта ухудшения частотного разрешения уха у больных нейросенсорной тугоухостью, а также провел экспериментальные исследования; М. И. Вашкевич определил задачи, которые необходимо было решить в ходе проведения исследований, принял участие в разработке метода и интерпретации результатов эксперимента.

Список использованных источников

1. Королева, И. В. Введение в аудиологию и слухопротезирование / И. В. Королева. – СПб. : КАРО, 2012. – 400 с.
2. Erber, N. Applications of hearing-loss simulation in education of student clinicians / N. Erber // J. of the Academy of Rehabilitative Audiology. – 1995. – Vol. 28. – P. 37–50.
3. Simulation of hearing loss using compressive gammachirp auditory filters / H. Hu. [et al.] // IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP). – Prague, Czech Republic, 2011. – P. 5428–5431. <https://doi.org/10.1109/ICASSP.2011.5947586>
4. Graf, I. Simulation of the Effects of Sensorineural Hearing Loss / I. Graf. – Cambridge : Massachusetts Institute of Technology, 1997. – 112 p.
5. Mourgela, A. Perceptually motivated hearing loss simulation for audio mixing reference / A. Mourgela, T. Agus, J. D. Reiss // 147th Audio Engineering Society Intern. Convention. – N. Y., 2019.
6. Moore, B. Cochlear Hearing Loss: Physiological, Psychological and Technical Issues / B. Moore. – 2nd ed. – John Wiley & Sons, 2008. – 332 p. <https://doi.org/10.1002/9780470987889>
7. Baer, T. Effects of spectral smearing on the intelligibility of sentences in noise / T. Baer, B. Moore // The J. of the Acoustical Society of America. – 1993. – Vol. 94, no. 3. – P. 1229–1241. <https://doi.org/10.1121/1.408176>
8. Goodwin, M. The STFT, sinusoidal models, and speech modification / M. Goodwin // Handbook of Speech Processing. – Springer, 2008. – P. 229–258. https://doi.org/10.1007/978-3-540-49127-9_12
9. A hearing impairment simulation method using audiogram-based approximation of auditory characteristics / N. Jinbo [et al.] // 15th Annual Conf. of the Intern. Speech Communication Association (INTERSPEECH). – Singapore, 2014. – P. 490–494.
10. The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold / R. D. Patterson [et al.] // The J. of the Acoustical Society of America. – 1982. – Vol. 72, no. 6. – P. 1788–1803. <https://doi.org/10.1121/1.388652>

11. Effects of spectral smearing of stimuli on the performance of auditory steady-state response-based brain-computer interface / H. J. Hwang [et al.] // *Cogn Neurodyn.* – 2017. – Vol. 11, no. 6. – P. 515–527. <https://doi.org/10.1007/s11571-017-9448-y>
12. Фонлантен, А. Слуховые аппараты / А. Фонлантен, Х. Арндт. – Ростов н/Д. : Феникс, 2009. – 302 с.
13. Glasberg, B. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments / B. Glasberg, B. Moore // *The J. of the Acoustical Society of America.* – 1986. – Vol. 79, no. 4. – P. 1020–1033. <https://doi.org/10.1121/1.393374>
14. Baer, T. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech / T. Baer, B. Moore // *The J. of the Acoustical Society of America.* – 1994. – Vol. 95. – P. 2277–2280. <https://doi.org/10.1121/1.408640>
15. Moore, B. Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech / B. Moore, B. Glasberg // *The J. of the Acoustical Society of America.* – 1993. – Vol. 94, no. 4. – P. 2050–2062. <https://doi.org/10.1121/1.407478>
16. Nejime, Y. Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise / Y. Nejime, B. Moore // *The J. of the Acoustical Society of America.* – 1997. – Vol. 102, no. 1. – P. 603–615. <https://doi.org/10.1121/1.419733>
17. Effects of spectral smearing of stimuli on the performance of auditory steady-state response-based brain-computer interface / J. H. Hwang [et al.] // *Cognitive Neurodynamics.* – 2017. – Vol. 11, no. 6. – P. 515–527. <https://doi.org/10.1007/s11571-017-9448-y>

References

1. Koroleva I. V. Vvedenie v audiologiyu i slukhoprotezirovanie. *Introduction to Audiology and Hearing Care.* Saint Petersburg, KARO, 2012, 400 p. (In Russ.).
2. Erber N. Applications of hearing-loss simulation in education of student clinicians. *Journal of the Academy of Rehabilitative Audiology*, 1995, vol. 28, pp. 37–50.
3. Hu. H., Sang J., Lutman M. E., Bleeck S. Simulation of hearing loss using compressive gammachirp auditory filters. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Prague, Czech Republic, 2011, pp. 5428–5431. <https://doi.org/10.1109/ICASSP.2011.5947586>
4. Graf I. *Simulation of the Effects of Sensorineural Hearing Loss.* Cambridge, Massachusetts Institute of Technology, 1997, 112 p.
5. Mourgela A., Agus T., Reiss J. D. Perceptually motivated hearing loss simulation for audio mixing reference. *147th Audio Engineering Society International Convention.* New York, 2019.
6. Moore B. *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues.* 2nd ed., John Wiley & Sons, 2008, 332 p. <https://doi.org/10.1002/9780470987889>
7. Baer T., Moore B. Effects of spectral smearing on the intelligibility of sentences in noise. *The Journal of the Acoustical Society of America*, 1993, vol. 94, no. 3, pp. 1229–1241. <https://doi.org/10.1121/1.408176>
8. Goodwin M. The STFT, sinusoidal models, and speech modification. *Handbook of Speech Processing*, Springer, 2008, pp. 229–258. https://doi.org/10.1007/978-3-540-49127-9_12
9. Jinbo N., Takamichi S., Toda T., Neubig G., Sakti S., Nakamura S. A hearing impairment simulation method using audiogram-based approximation of auditory characteristics. *15th Annual Conference of the International Speech Communication Association (INTERSPEECH)*. Singapore, 2014, pp. 490–494.
10. Patterson R. D., Nimmo-Smith I., Weber D. L., Milroy R. The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *The Journal of the Acoustical Society of America*, 1982, vol. 72, no. 6, pp. 1788–1803. <https://doi.org/10.1121/1.388652>
11. Hwang J. H., Nam K. W., Jang D. P., Kim I. Y. Effects of spectral smearing of stimuli on the performance of auditory steady-state response-based brain-computer interface. *Cogn Neurodyn*, 2017, vol. 11, no. 6, pp. 515–527. <https://doi.org/10.1007/s11571-017-9448-y>
12. Vonlanthen A., Arndt H. *Hearing Instrument Technology for the Hearing Healthcare Professional.* Thomson Delmar Learning, 2007, 318 p.
13. Glasberg B., Moore B. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, 1986, vol. 79, no. 4, pp. 1020–1033. <https://doi.org/10.1121/1.393374>
14. Baer T., Moore B. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *The Journal of the Acoustical Society of America*, 1994, vol. 95, pp. 2277–2280. <https://doi.org/10.1121/1.408640>

15. Moore B., Glasberg B. Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech. *The Journal of the Acoustical Society of America*, 1993, vol. 94, no. 4, pp. 2050–2062. <https://doi.org/10.1121/1.407478>
16. Nejime Y., Moore B. Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*, 1997, vol. 102, no. 1, pp. 603–615. <https://doi.org/10.1121/1.419733>
17. Hwang J. H., Nam W. N., Jang P. D., Kim I. Y. Effects of spectral smearing of stimuli on the performance of auditory steady-state response-based brain-computer interface. *Cognitive Neurodynamics*, 2017, vol. 11, no. 6, pp. 515–527. <https://doi.org/10.1007/s11571-017-9448-y>

Информация об авторах

Порхун Максим Игоревич, магистр технических наук, ассистент кафедры электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: porhun@bsuir.by

Вашкевич Максим Иосифович, кандидат технических наук, доцент кафедры электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: vashkevich@bsuir.by

Information about the authors

Maksim I. Porhun, M. Sci. (Eng.), Assistant Lecturer of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: porhun@bsuir.by

Maksim I. Vashkevich, Cand. Sci. (Eng.), Associate Professor of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: vashkevich@bsuir.by



УДК 004.93
<https://doi.org/10.37661/1816-0301-2021-18-3-83-96>

Оригинальная статья
Original Paper

Нормализация данных в машинном обучении

В. В. Старовойтов[✉], Ю. И. Голуб

Объединенный институт проблем информатики

Национальной академии наук Беларуси,

ул. Сурганова, 6, Минск, 220012, Беларусь

[✉]E-mail: valerys@newman.bas-net.by

Аннотация. В задачах машинного обучения исходные данные часто заданы в разных единицах измерения и типах шкал. Такие данные следует преобразовывать в единое представление путем их нормализации или стандартизации. В работе показана разница между этими операциями. Систематизированы основные типы шкал, операции над данными, представленными в этих шкалах, и основные варианты нормализации функций. Предложена новая шкала частот и приведены примеры использования нормализации данных для их более корректного анализа.

На сегодняшний день универсального метода нормализации данных, превосходящего другие методы, не существует, но нормализация исходных данных позволяет повысить точность их классификации. Кластеризацию данных методами, использующими функции расстояния, лучше выполнять после преобразования всех признаков в единую шкалу.

Результаты классификации и кластеризации разными методами можно сравнивать различными оценочными функциями, которые зачастую имеют разные диапазоны значений. Для выбора наиболее точной функции можно выполнить нормализацию нескольких из них и сравнить оценки в единой шкале.

Правила разделения признаков древовидных классификаторов инвариантны к шкалам количественных признаков. Они используют только операцию сравнения. Возможно, благодаря этому свойству классификатор типа «случайный лес» в результате многочисленных экспериментов признан одним из лучших при анализе данных разной природы.

Ключевые слова: классификация объектов, кластеризация, нормализация данных, нормализация функций, сигмоида, гиперболический тангенс, случайный лес

Благодарности. Работа частично выполнена в рамках проектов БРФФИ Ф20РА–014 и Ф21ПАКГ–001.

Для цитирования. Старовойтов, В. В. Нормализация данных в машинном обучении / В. В. Старовойтов, Ю. И. Голуб // Информатика. – 2021. – Т. 18, № 3. – С. 83–96. <https://doi.org/10.37661/1816-0301-2021-18-3-83-96>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 02.07.2021
Подписана в печать | Accepted 11.08.2021
Опубликована | Published 29.09.2021

Data normalization in machine learning

Valery V. Starovoitov[✉], Yuliya I. Golub

*The United Institute of Informatics Problems
of the National Academy of Sciences of Belarus,
st. Surganova, 6, Minsk, 220012, Belarus
✉E-mail: valerys@newman.bas-net.by*

Abstract. In machine learning, the input data is often given in different dimensions. As a result of the scientific papers review, it is shown that the initial data described in different types of scales and units of measurement should be converted into a single representation by normalization or standardization. The difference between these operations is shown. The paper systematizes the basic operations presented in these scales, as well as the main variants of the function normalization. A new scale of parts is suggested and examples of the data normalization for correct analysis are given. Analysis of publications has shown that there is no universal method of data normalization, but normalization of the initial data makes it possible to increase the accuracy of their classification. It is better to perform data clustering by methods using distance functions after converting all features into a single scale. The results of classification and clustering by different methods can be compared with different scoring functions, which often have different ranges of values. To select the most accurate function, it is reasonable to normalize several functions and to compare their estimates on a single scale. The rules for separating features of tree-like classifiers are invariant to scales of quantitative features. Only comparison operation is used. Perhaps due to this property, the random forest classifier, as a result of numerous experiments, is recognized as one of the best classifiers in the analysis of data of different nature.

Keywords: object classification, clustering, data normalization, function normalization, sigmoid, hyperbolic tangent, random forest

Acknowledgements. This work was partially performed within the framework of the BRFFR projects F20RA–014 and F21PAKG–001.

For citation. Starovoitov V. V., Golub Y. I. Data normalization in machine learning. *Informatics*, 2021, vol. 18, no. 3, pp. 83–96 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-83-96>

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

Введение. Нормализация данных в машинном обучении – это метод предварительной обработки, при котором данные преобразуются, чтобы обеспечить равный вклад каждого показателя [1, 2]. Успех алгоритмов машинного обучения зависит от точности описания данных для получения обобщенной прогнозной модели проблемы классификации [3]. Важность нормализации данных для улучшения описания и повышения точности алгоритмов машинного обучения была отмечена многими исследователями [4]. Для применения алгоритмов, использующих некую метрику или сравнение данных разных типов, предварительно требуется представление данных в одной шкале измерения.

В литературе часто путают понятия нормализации и стандартизации данных (см., например, [1–4]). Цель настоящей публикации – уточнить эти понятия, показать разницу между ними и продемонстрировать, как их следует применять в области машинного обучения.

Шкалы описания данных. В машинном обучении данные представляют собой признаки, описывающие некоторые объекты, понятия или события. Признаки могут быть записаны в разных шкалах: категориальных (неметрических) и количественных (численных). Такая классификация шкал была предложена Стивенсом [5]. Она представлена первыми четырьмя типами шкал в табл. 1 и подвергалась множественной критике [6]. Общепризнанной классификации шкал данных не существует, предлагаются и другие варианты этой классификации [7]. Основная идея любой классификации шкал заключается в группировании однотипно описываемых данных и определении допустимых для каждой группы операций над данными одного типа.

Таблица 1. Дополненная классификация шкал Стивенса

Table 1. Augmented classification of Stevens scales

Шкала <i>Scale</i>	Свойство <i>Property</i>	Единицы измерения <i>Units measurements</i>	Матем. операции <i>Mathem. operations</i>	Доп. операции <i>Add. operations</i>	Центральная тенденция <i>Central trend</i>
Номинальная или наименований	Принадлеж- ность классу	Названия	$=, \neq$	Группировка	Мода
Порядковая или рангов	Сравнение, уровень	Порядковые величины	$-//-, >, <$	Сортировка	Медиана
Интервальная	Разница в интервалах	Относительные величины	$-//-, +, -$	Сравнение	Среднее арифмети- ческое
Отношений	Абсолютная величина	Неотрицательные величины	$-//-, *, /$	Отношение	Среднее геометриче- ское и гар- моническое
Частей	Часть от целого	Числа в диапазоне [0; 1]	$-//-$	Сравнение	Нет

Все шкалы делятся на категориальные и количественные, категориальные шкалы – на номинальные и порядковые (табл. 1). Чаще всего данные, представленные в категориальных шкалах, носят субъективный характер. Данные в шкалах наименований, например названия улиц, номера телефонов, могут быть представлены символами. Данные такого типа можно только сравнивать: равны они или нет. Данные в порядковых шкалах могут содержать оценки в виде чисел, позволяющих задать порядок и сравнить описания разных объектов. Например, ранжировать людей по росту не в сантиметрах, а выше или ниже, первый, второй или десятый. Школьные оценки описываются так же. Их можно сравнивать (больше или меньше, лучше или хуже), используя операции больше, меньше, равно. Вместе с тем описание данных в этих шкалах субъективно и применение к ним арифметических операций является некорректным. Например, нельзя утверждать, что ученик, получивший оценку 10, знает в два раза больше ученика, получившего оценку 5.

Данные в количественных шкалах представлены числами. Числа принимают значения из определенных шкал, которые можно разделить на три группы: интервальные, отношений и частей. Третья группа предлагается авторами настоящей статьи. Интервальные шкалы разбиты на равные интервалы, но не имеют строго определенного начала (нуля). В частности, время можно измерять в часах или годах, но начало исчисления условно. Такие данные можно складывать и вычитать, но их нельзя множить и делить, так как меняется цена интервала измерения. Шкалы отношений также разделены на равные интервалы, но они имеют строго определенный ноль, с которого начинается отсчет. Данные, представленные в таких шкалах, не могут быть отрицательными. К ним можно применять дополнительно операции умножения и деления, статистические операции. Например, двадцатилетний человек в два раза моложе сорокалетнего (данные представлены в шкале отношений), но родившийся в 2000 г. не будет в два раза младше родившегося в 1980 г. (данные представлены в шкале интервалов).

Классификацию Стивенса можно дополнить пятой группой – шкалами частей от целого (например, процентов в долях, КПД, частей угла, вероятности и т. п.). Такие шкалы начинаются с нуля и заканчиваются единицей, в них описываются безразмерные нормализованные данные.

Признаки одного и того же объекта могут описываться в разных шкалах. Например, ученик Иванов (по шкале наименований), ему 10 лет (по шкале отношений), он родился в 2011 г. (по шкале интервалов), он отличник (по шкале порядка), решил все задачи контрольной работы (единица по шкале частей), у него черные волосы (по шкале наименований).

В задачах классификации и кластеризации исходные данные должны быть описаны в числовом представлении и, желательно, преобразованы в единую шкалу измерений. Рассмотрим способы преобразования данных.

Нормализация и стандартизация данных. Данные, описанные в категориальной шкале, нормализовать невозможно.

Определение 1. Преобразование численных данных в диапазон с крайними значениями $[0; +1]$ будем называть нормализацией данных.

Следует отличать понятия нормализации и нормировки данных. Нормировка – это преобразование данных, измеренных в разных единицах (например, когда длины объектов описаны в сантиметрах и дюймах), в единую шкалу.

Чаще всего нормализованные данные имеют значения в диапазоне $[0; +1]$, реже – в диапазоне $[-1; +1]$, но данные, представленные в этих диапазонах, легко трансформируются из одного диапазона в другой. Если количественные данные описаны в одной шкале с разными интервалами, после нормализации их можно сравнивать и оценивать математически. Фактически нормализация данных – это их преобразование в шкалу частот.

Линейная нормализация набора произвольных данных x выполняется по формуле

$$y = (x - \min(x)) / (\max(x) - \min(x)), \quad (1)$$

где x – исходное множество данных, y – преобразованное множество данных, \min и \max – операции вычисления минимального и максимального значений.

Десятичное масштабирование является нелинейным методом нормализации данных в диапазон $[0; +1]$:

$$y = (x - \min(x)) / 10^j, \quad j = \log_{10} |\max(x) - \min(x)|. \quad (2)$$

Еще одним способом нормализации данных можно считать метод вычисления интервала значений, который используется при построении блочной диаграммы распределения данных, называемой ящиком с усами и предложенной в работе [8]. Если дано множество одномерных величин $X_n = \{x_1, x_2, \dots, x_n\}$, «ящик с усами» строится следующим образом. Вычисляются значения квартилей массива X_n $Q1, Q2, Q3, Q4$; межквартильный размах $IQR = Q3 - Q1$; границы диапазона основных значений ящика $L = Q1 - 1,5IQR$ и $R = Q3 + 1,5IQR$ (рис. 1). Диапазон основных значений данных будет равен $[L; R]$. Все точки вне этого диапазона классифицируются как выбросы. Для данных, имеющих нормальное распределение, всего около 0,7 % значений будут лежать за пределами вычисленного межквартильного интервала [9], а его легко преобразовать в диапазон значений $[0; +1]$ по формуле (1), где $L = \min(x)$, $R = \max(x)$.

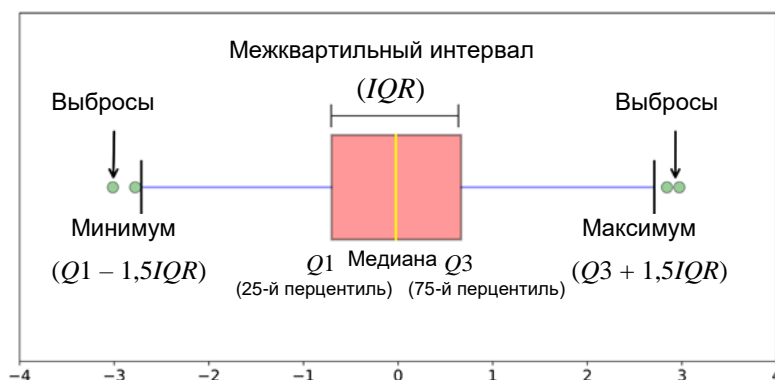


Рис. 1. Идея стандартного вычисления межквартильного интервала

Fig. 1. The idea of standard calculation of interquartile range

Для данных, не соответствующих нормальному распределению, в работе [10] скорректировали границы межквартильного интервала в сторону искажения данных, используя нижний и верхний полуквартильный диапазоны $A = Q2 - Q1$ и $B = Q3 - Q2$, которые определяют грани-

цы интервала как $[Q1 - 3A; Q3 + 3B]$. В статье [11] значения квартилей $Q1$ и $Q3$ заменены медианным значением второго квартиля $Q2$, а константа 1,5 заменена формулой, зависящей от размера диапазона. Такой подход увеличивает число вычислительных операций, поскольку применяется дополнительная сортировка, но, как правило, он незначительно уточняет границы искомого интервала.

В работе [12] для данных, соответствующих непрерывному унимодальному распределению, предложено вычислять отклонения от $Q1$ и $Q3$ в виде определенных функций. Однако для их построения требуется найти значение медианы попарных отклонений МС (medcouple), предложенной в статье [13] и вычисляемой от медианного значения исходного множества данных. Этот подход также требует дополнительных вычислений, поскольку сложность вычисления МС равна $O(n \log n)$ для n величин.

Обобщая описанные подходы, можно сделать вывод, что если множество данных X_n имеет ограниченный диапазон значений $[x_0; x_k]$, где $x_0 < x_k$, то его можно нормализовать, преобразовав в другое множество $X_n \rightarrow Y_n$ с фиксированным диапазоном значений $[0; +1]$ или $[-1; +1]$ либо с заданными свойствами. Например, преобразовав медианное значение множества, получим $y_{\text{med}}(x_{\text{med}}) = 0,5$ при $y(Q1) = y_1$, $y(Q3) = y_3$, $y(L) = 0$, $y(B) = 1$. Очевидно, что можно назначить и другие значения y для указанных значений x , например $y(x_{\text{min}}) = 0$, $y(x_{\text{max}}) = 1$.

Если данные имеют нормальное или примерно нормальное распределение, можно выполнить процедуру их стандартизации.

Определение 2. Преобразование данных в набор, имеющий определенные статистические характеристики, но неопределенные минимальные и максимальные значения, будем называть стандартизацией данных.

Наиболее распространенные методы стандартизации данных [2] приведены в табл. 2. Например, Z-преобразование меняет значения набора данных так, что он будет иметь нулевое среднее и единичную дисперсию.

Таблица 2. Методы стандартизации набора данных x

Table 2. Methods for dataset x standardization

Название Name	Формула Formula
Z-преобразование	$y = (x - \mu) / \sigma$, μ – среднее, σ^2 – дисперсия
Преобразование Парето	$y = (x - \mu) / \sqrt{\sigma^2}$
Масштабирование стабильности переменной (Variable Stability Scaling)	$y = \mu (x - \mu) / \sigma^2$
Степенное преобразование	$y = p - \mu^p$, $p = \sqrt{x - \min(x)}$
Med-MAD-преобразование	$y = (x - \text{med}(x)) / \text{MAD}$, $\text{MAD} = \text{med}(x - \text{med})$

Стандартизованные данные можно нормализовать, убрав выбросы и преобразовав значения в диапазон $[0; +1]$. Если данные не имеют нормального распределения, указанные выше методы просто приведут их к другому масштабу, но не преобразуют в шкалу с диапазоном $[0; +1]$. Это методы стандартизации, а не нормализации данных.

Нормализация функций. Не всегда известно полное множество данных, однако может быть известна функция, описывающая их природу или позволяющая оценить границы диапазона значений.

Определение 3. Под нормализацией функции будем понимать преобразование, нормализующее область ее значений в диапазон $[0; +1]$ или $[-1; +1]$.

Основные варианты нормализации функций приведены в табл. 3. Основное отличие нормализации данных от нормализации функций состоит в том, что набор данных всегда конечен и можно вычислить его минимальное и максимальное, среднее и медианное значения, среднеквадратическое отклонение и другие статистические характеристики, которые используются

для нормализации и стандартизации данных. При нормализации функций известно только уравнение функции. По нему можно определить предельные значения, а также значения функции в некоторых точках.

Таблица 3. Основные варианты нормализации функций [14]

Table 3. The main options of function normalization [14]

Название <i>Name</i>	Формула <i>Formula</i>	Стандартные параметры <i>Standard parameters</i>	Диапазон значений <i>Range of values</i>
Алгебраическая функция	$f(x) = x / (x + a), a > 0$	$a = 1$	$[-1; +1]$
Обобщение алгебраической функции	$f(x) = x^n / (x ^n + a), a > 0$	$n = 2, a = 1$	$[0; +1]$
Обобщенный вариант сигмоиды	$f(x) = (1 + \exp^{-bx})^{-a}, a, b > 0$	$a = b = 1$	$[0; +1]$
Арктангенс	$f(x) = \arctan(x/a), a > 0$	$a = 1$	$[-1; +1]$
Гиперболический тангенс	$f(x) = \tanh(x) = (e^{x/a} - e^{-x/a}) / (e^{x/a} + e^{-x/a}), a > 0$	$a = 1$	$[-1; +1]$
На базе гиперболического тангенса	$f(x) = 0,5 \cdot x [\tanh(0,01(x/a - \mu) / \sigma) + 1], a > 0$	$a = 1$	$[0; +1]$
Функция Гудермана	$f(x) = \text{gd}(x) = 4 \cdot \arctan(\tanh(x/a) / \pi), a > 0$	$a = 2$	$[-1; +1]$

Если область значений функции ограничена, к функции можно применить те же преобразования, которые применяют для нормализации данных. Например, если некоторые данные описываются функцией $y = \sin(x) + \cos(x)$, то она нормализуется в диапазон $[0; +1]$ следующим образом:

$$f(y) = (y + \sqrt{2}) / (2 \cdot \sqrt{2}).$$

Под нормализацией функции $a \leq y(x) < +\infty$ понимается такое ее преобразование $f(y)$, что $0 \leq f(y) \leq 1$ и $f(a) = 0, f(\infty) = 1$. Если функция монотонна, то из $x_1 < x_2$ следует $f(y_1) \leq f(y_2)$. Здесь x является переменной.

Функция с бесконечным диапазоном значений $-\infty < y(x) < +\infty$ обычно нормализуется в диапазон значений $[-1; +1]$, где $f(-\infty) = -1$ и $f(\infty) = 1$. Значения функции f можно преобразовать в диапазон $[0; +1]$ посредством дополнительного преобразования $g(f) = (f + 1) / 2$. Поэтому будем считать, что нормализованная функция вычисляет некоторую величину в диапазоне $[0; +1]$, где крайние значения соответствуют минимальному и максимальному значениям функции f .

Если количественный признак описывает объекты двух классов со средними значениями классов μ_1, μ_2 и дисперсией σ^2 , а данные в каждом классе имеют нормальное распределение с одинаковой дисперсией, то нормализованное значение признака дает логистическая функция (сигмоида) [15].

В работе [16] описаны свойства преобразования Вох-Сох (3), предложенного для «исправления» данных, не имеющих нормального распределения, и показано, что применение этого преобразования улучшает результаты классификации. Экспериментальные исследования показали, что если данные плохо описываются нормальным распределением, то требуется многократное применение этого преобразования и оптимизация оценки максимального правдоподобия относительно нормального распределения для каждого варианта. На рис. 2 слева изображена гистограмма множества данных, представленных в диапазоне $[3,944 \cdot 10^{-05}; 3,372]$ и имеющих распределение Вейбулла, а справа – результат их преобразования в новое представление в диапазоне $[-4,342; 1,376]$, имеющее нормальное распределение с уровнем значимости $p = 0,90$. Аппроксимация обоих наборов данных функцией нормального распределения показана кривыми красного цвета. Результат стандартизации данных получен посредством применения преобразования Вох-Сох к набору данных x с параметром $\lambda = 0,2$. Далее при необходимости можно нормализовать эти данные методами (1) или (2):

$$y = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{для } \lambda \neq 0, \\ \log(x) & \text{для } \lambda = 0, \end{cases} \quad (3)$$

где λ – константа.

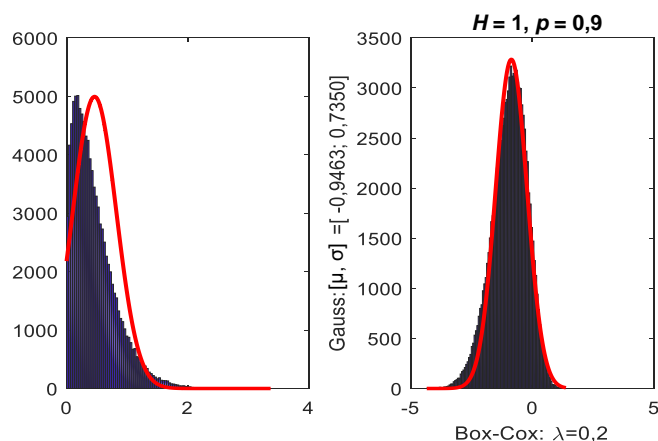


Рис. 2. Пример стандартизации данных, имеющих распределение Вейбулла с параметрами (0,5; 1,3), в нормально распределенное представление с параметрами (-0,946; 0,735)

Fig. 2. An example of standardization of data with Weibull distribution with parameters (0,5; 1,3) into a normally distributed representation with parameters (-0,946; 0,735)

Недостатком метода линейной нормализации (1) является наличие аномальных значений данных (выбросов), которые «увеличивают» диапазон. Это приводит к тому, что нормализованные значения концентрируются в узком диапазоне вблизи нуля. Чтобы избежать этого, следует определять диапазон с помощью не максимальных и минимальных значений, а среднего значения и дисперсии.

Дополнительно данным можно придать определенные свойства, например убрать выбросы, отцентрировать данные в середину диапазона.

Применение описанных функций. Рассмотрим примеры описания температуры воды в разных шкалах (табл. 4).

Таблица 4. Температура воды в единицах измерения разных шкал

Table 4. Water temperature in different scales measurement

Состояние воды Water condition	Кельвин Kelvin	Градус Цельсия Celsius	Градус Фаренгейта Fahrenheit	Градус Ньютона Newton	Градус Реомюра Reaumur
Замерзание	273,15	0	32	0	0
Кипение	373,15	100	212	33	80

Шкала Кельвина относится к шкалам отношений, поскольку в ней есть абсолютный ноль, а остальные шкалы – к интервальным. Все шкалы в табл. 4 связаны линейными зависимостями. Теоретический верхний предел температуры равен примерно 10^{32} К, т. е. можно считать, что диапазоны температурных шкал не ограничены справа. Самая низкая температура воздуха, зафиксированная на Земле, равна $-91,2^\circ\text{C}$ (181,95 К), а самая высокая $+56,7^\circ\text{C}$. Учитывая эти экстремальные значения, температуру воздуха, описанную в любой шкале, можно перевести в нормализованный вид в шкале частот с значениями от 0 до +1. Если нормализовать температуры, указанные в табл. 1, в шкале Кельвина по формуле $y = x / (x + 1)$, то для 0°C получим число 0,996 352, а для 100°C – число 0,997 327. Новые значения различаются только тысячными долями. Если температура воздуха нормализована линейно в указанном интервале [min; max], то значение 0°C будет преобразовано в число 0,616 63.

В зависимости от решаемой задачи нормализовать такой параметр, как значение температуры, можно разными способами. Пользуясь функциями принадлежности нечеткому множеству (например, функцией Гаусса), на базе показаний температуры можно также сформировать новый параметр, например степень комфортности температуры на улице или в помещении с диапазоном значений $[0; +1]$ и пиком, равным единице и соответствующим 22°C .

Формирование новых признаков на базе исходных данных. Один из самых известных алгебраических методов конструирования признаков называется методом главных компонент (principal component analysis, PCA). Главные компоненты – это новые признаки, сконструированные в виде линейных комбинаций исходных признаков [15].

Авторы работы [17] предложили применять PCA к численным признакам и нормализовать значение каждого признака, разделив на большее собственное число и умножив на собственное число, соответствующее этому признаку. На четырех наборах данных разных типов было продемонстрировано повышение точности классификации с помощью искусственных нейронных сетей (ИНС).

Функции принадлежности нечеткому множеству определяют степень принадлежности количественных величин определенному нечеткому множеству [18]. Диапазон их значений равен $[0; +1]$. Четыре основных типа функций определяются формой кривой функции принадлежности, типичными их представителями являются треугольная и трапециевидная функции, а также сигмоида и функция Гаусса (рис. 3). Следует отметить следующие особенности этих функций: не всегда максимальное значение функции соответствует крайним значениям диапазона области определения функции, одно и то же значение функции может соответствовать разным величинам диапазона области определения.

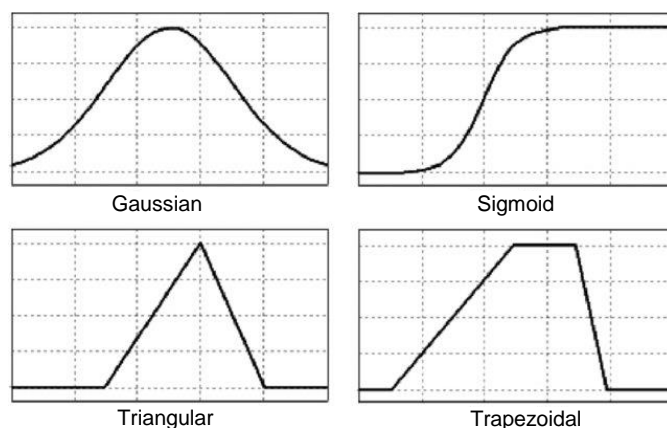


Рис. 3. Графики четырех основных типов функций нечеткой принадлежности

Fig. 3. Plots of four main types of fuzzy membership functions

В рамках настоящей статьи функции нечеткой принадлежности можно рассматривать как функции нормализации, которые преобразуют количественное описание данных в качественную шкалу значений, т. е. формируют другое представление этих данных. Например, рост человека, измеренный в сантиметрах или дюймах, можно качественно описать как низкий, средний и высокий. Рост двух и более человек можно сравнить как в количественной шкале, так и качественной.

Применение при кластеризации и классификации. Методы кластерного анализа и многие классификаторы рассчитывают расстояние между двумя точками в признаковом пространстве, используя евклидову метрику. Если один из признаков имеет более широкий диапазон значений, расстояние будет определяться этим доминирующим признаком. Диапазон всех признаков при классификации должен быть нормализован так, чтобы каждый вносил приблизительно пропорциональный вклад [19]. Еще одна причина, по которой применяется масштабирование значений признаков, заключается в том, что градиентный спуск сходится примерно в 14 раз

быстрее на нормализованных данных [20]. Авторы статьи [20] тестировали преобразование Парето на базе данных ImageNet. Для этой цели также применяют Z-преобразование [3]. Алгоритм логистической калибровки признака (z -оценки) описан в работе [15].

Сигмоида часто используется в ИНС в качестве функции активации, которая позволяет усиливать слабые сигналы и не усиливать сильные. Производная сигмоиды может быть легко выражена через саму функцию:

$$S'(x) = S(x) \times (1 - S(x)),$$

что позволяет сократить вычислительную сложность методов обучения ИНС. Вместо сигмоиды можно использовать гиперболический тангенс. Его производная также вычисляется через исходную функцию:

$$\tanh'(x) = 1 - \tanh^2(x).$$

Отметим, что максимальное значение производной сигмоиды равно 0,25, а максимальное значение производной тангенса – единице. Поэтому обучение ИНС с использованием гиперболического тангенса вместо сигмоиды происходит быстрее.

Если количественный признак описывает объекты двух классов со средними значениями классов μ_1 , μ_2 и дисперсией σ^2 , а в каждом классе данные имеют нормальное распределение с одинаковой дисперсией, то лучшее нормализованное значение признака дает сигмоида [15].

В статье [2] рассматривалось применение 14 методов нормализации и стандартизации данных и их влияние на производительность классификации с учетом полного набора функций, выбора функций и их взвешивания. Эксперименты были выполнены на 21 общедоступной базе реальных и синтезированных данных [2]. Было отмечено, что ни один метод не превосходит другие во всех экспериментах. Поэтому авторы отметили наборы лучших методов. Лучшими вариантами нормализации (точнее, стандартизации) были признаны Z-преобразование и преобразование Парето, а также гиперболический тангенс.

Преимущество древовидных классификаторов. Древовидные классификаторы используют деление признакового пространства по правилу сравнения значений одного признака с некоторой пороговой величиной, т. е. основой является операция сравнения. Если изменяется диапазон значений признаков, пороговые значения в правилах древовидного классификатора можно легко изменить, но сами правила не меняются. Древовидные модели нечувствительны к шкале количественного признака. Например, для обученного дерева не важно, измерена температура по шкале Цельсия или Фаренгейта. Несущественен и переход от линейной шкалы к логарифмической, т. е. порог разделения просто будет равен $\log(v)$, а не v [15]. Древовидные модели нечувствительны к монотонным преобразованиям шкалы признака – преобразованиям, не изменяющим относительного порядка его значений. Таким образом, древовидные модели не используют шкалу количественных признаков, а трактуют признаки как порядковые. То же самое справедливо для моделей на основе правил. Возможно, это является причиной того, что классификатор типа «случайный лес» в результате многочисленных экспериментов оказался лучшим среди 179 вариантов классификаторов 17 классов, протестированных на 121 множестве данных в статье [21]. В работе [22] показано, что классификатор «случайный лес» продемонстрировал более точную классификацию рака груди на два класса по 31 признаку по сравнению с байесовским классификатором.

Сравнение оценочных функций с разными диапазонами значений. На рис. 4 видно, что графики сигмоиды, гиперболического тангенса и функции Гудермана очень близки, графики арктангенса и алгебраической функции имеют несколько отличную форму, а последняя функция – наиболее плавную форму. При большом диапазоне нормализуемых значений лучше использовать арктангенс и алгебраическую функцию, задавая определенные значения функции нормализации $y = f(x)$ посредством выбора ее параметров. Например, можно получить $y = 0,5$ для исходного значения $x = 8$, применяя любую функцию нормализации. При этом диапазон

исходных значений $[0; 8]$ будет преобразован в диапазон $[0; 0,5]$, а оставшийся диапазон $[8; +\infty)$ – в диапазон $[0,5; +1]$. Разные функции нормализации по-разному преобразуют эти два поддиапазона.

В работе [23] сравнивались результаты бинарной классификации данных разной природы. Результаты оценивались по матрице ошибок размером 2×2 . Целью исследования было сравнение точности оценок двух функций: коэффициента корреляции Мэтьюса (МСС) с диапазоном значений $[-1; +1]$ и популярного в медицинских исследованиях диагностического отношения шансов (DOR) с диапазоном значений $[0; +\infty)$. Обе функции были нормализованы в диапазон значений $[0; +1]$: первая – линейно, вторая – с помощью алгебраической функции $y = x/(x + 1)$. Было показано, что функция DOR инвариантна к дисбалансу классов, но функция МСС является более информативной. При числе анализируемых объектов более 53 коэффициент корреляции Пирсона между значениями нормализованных функций в среднем был более 0,95 для четырех миллионов искусственно сгенерированных матриц ошибок.

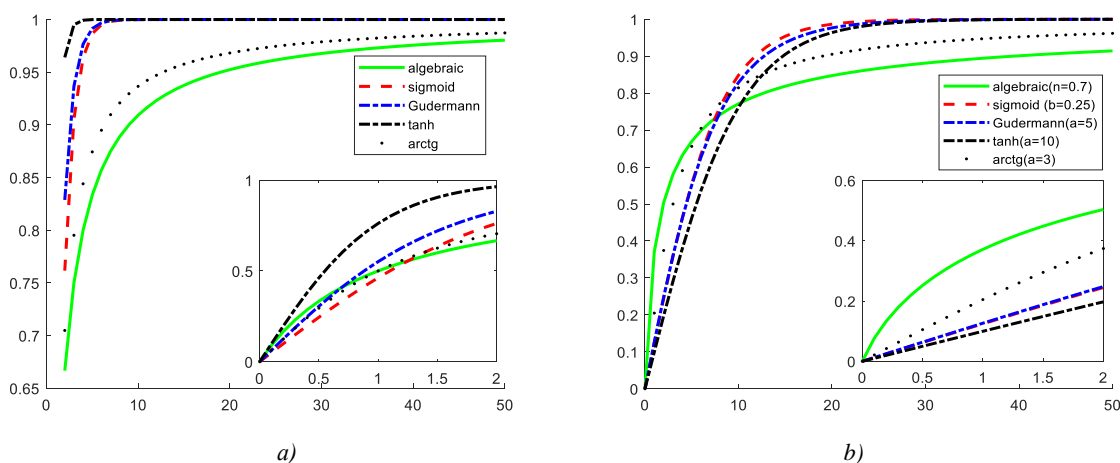


Рис. 4. Графики функций нормализации со стандартными параметрами (а) и измененными (б)
Fig. 4. Plots of normalization functions with the standard parameters (a) and modified (b)

Аналогично можно сравнить точность оценок, получаемых другими функциями с разными диапазонами значений, например коэффициентом корреляции Пирсона (диапазон $[0; +1]$) и среднеквадратическим отклонением (диапазон $[0; +\infty)$), нормализуя их значения.

Следует отметить: если данные двух типов преобразованы в одну шкалу, то из выражения $y_1 = f_1(x) > y_2 = f_2(x)$ не следует, что первое значение или признак лучше второго. Примером смогут служить коэффициенты корреляции Пирсона, Спирмена и Кендалла. Их значения обычно находятся в отношениях Пирсон > Спирмен > Кендалл. Следует оценивать динамику изменений этих функций на равных диапазонах из области их определения.

Проблема адекватности применения математических операций к данным, представленным в разных шкалах. Известно, что к данным рангового типа нельзя применять арифметические операции [24], однако это происходит повсеместно. Например, после защиты диссертации каждый член совета по защите выставляет оценку диссертации, затем вычисляется средний балл. Изначально выставляются оценки «отлично» (5), «хорошо» (4), «удовлетворительно» (3) и «неудовлетворительно» (2). В данном случае цифры – это символы, означающие только порядковые эквиваленты оценок. Пусть две трети совета поставила оценку «хорошо» и одна треть – «отлично». Сложить оценки «хорошо» и «отлично» нельзя, но их заменяют числами и тогда средний балл считается равным 4,3. Что это означает в данной шкале оценок? Оценка диссертации на 30 % выше оценки «хорошо»? Но такой оценки нет в применяемой шкале, по ней можно вычислить только медианное значение оценки, т. е. «хорошо».

Если те же оценки представить в виде чисел и перейти в шкалу частей, то можно вычислить, какую долю от максимально возможной суммы получила диссертация. Оценка будет равна 0,86.

Аналог такого подхода имеет место при вычислении оценок централизованного тестирования, если их разделить на максимальное число 100.

Нормализация данных в цифровой обработке изображений. В области классификации и кластеризации данных цифровые изображения часто используются как исходные данные. При линейном растяжении контраста или гамма-коррекции изображений они сначала подвергаются процедуре нормализации, т. е. преобразования в диапазон значений $[0; 1]$ по формуле (1). Насколько корректно такое преобразование? Напомним, что значения яркостей цифрового изображения являются безразмерными величинами и обозначают номер диапазона, в который попадает значение сигнала в результате его квантования при выполнении аналого-цифрового преобразования. Таким образом, значение яркости одного и того же сигнала может быть нулем или другим числом в зависимости от выбранного числа диапазонов при квантовании (два или более). Фактически яркость описывается в порядковой шкале с фиксированным числом значений, обычно кратным степени двойки. К таким данным нельзя применять арифметические операции, поэтому при обработке изображений значения яркости сначала преобразуют в другую шкалу (численную), а затем обратно в порядковую шкалу.

А. Чеддад в статье [25] предложил интересный метод нормализации цифровых изображений методом Вох-Сох, утверждая, что он повышает качество преобразованных изображений по сравнению с другими подходами. Авторы настоящей статьи проверили, действительно ли возможна нормализация исходных данных, представленных в виде изображений. На рис. 5 приведены примеры преобразования изображения методом, предложенным Чеддадом, с разными значениями параметра λ . На рис. 6 показаны гистограммы двух вариантов преобразования изображений методом Вох-Сох. При $\lambda < 1$ значения яркости нелинейно увеличиваются, при $\lambda > 1$ – нелинейно уменьшаются. Оказалось, что, как правило, яркости цифровых изображений имеют распределения, существенно отличные от нормального, и методом Вох-Сох их невозможно привести к нормальному распределению. В данном примере результатом является степенное преобразование значений яркости из исходного диапазона $[14; 248]$ в новые диапазоны $[3,3322; 5,5683]$, $[5,2013; 12,0929]$, $[10,6401; 45,4126]$, $[34,6116; 414,7514]$ с последующим преобразованием по формуле (1) в диапазон $[0; 255]$.



Рис. 5. Слева – оригинальное изображение, далее – изображения, преобразованные по формуле (3) с параметром λ , равным 0; 0,25; 0,6; 1,1
 Fig. 5. First left is the original image, then the images are transformed according to the formula (3) with the parameter λ equal to 0; 0,25; 0,6; 1,1

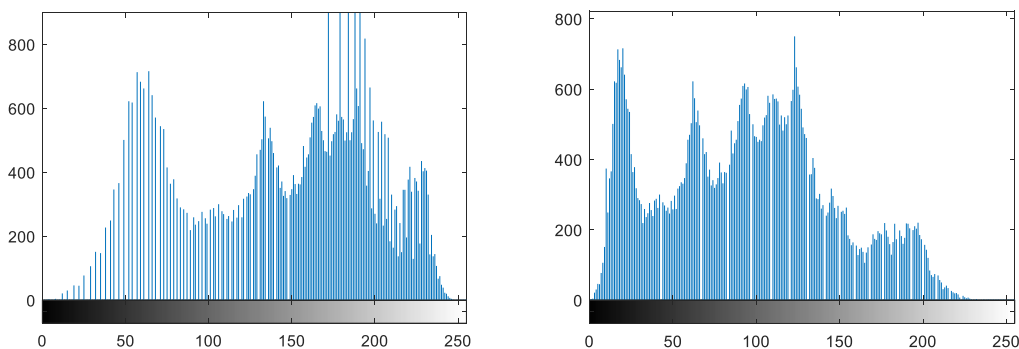


Рис. 6. Гистограммы изображений после преобразований с параметром λ , равным 0 и 1,1
 Fig. 6. Histograms of the images after transformations with parameters λ as 0 and 1,1

Заключение. Задачи классификации и кластеризации относятся к области машинного обучения. Данные в них часто представлены в разных шкалах. Больше информации и возможностей для анализа несут данные, представленные в количественных шкалах. Данные, представленные во всех шкалах, кроме номинальной, можно сравнивать, используя операции больше, меньше либо равно.

В статье описаны основные классы шкал представления данных разных типов. Приведены основные подходы к нормализации и стандартизации данных, а также нормализации функций с конечным и бесконечным диапазонами значений.

Многие экспериментальные исследования показывают, что универсального метода нормализации данных, превосходящего другие методы, не существует, но нормализация исходных данных позволяет повысить точность их классификации [4, 19, 20, 26, 27]. Следовательно, кластеризацию данных (особенно методами, использующими функции расстояния) лучше выполнять после преобразования всех признаков в единую шкалу.

Результаты классификации и кластеризации разными методами можно сравнивать различными оценочными функциями. Они имеют разные диапазоны значений. Для выбора наиболее подходящей функции можно выполнить нормализацию всех сравниваемых функций, тогда сравниваемые оценки будут в одной шкале – шкале частей. Байесовский классификатор использует понятие вероятности наступления некоторого события, а вероятность также можно считать описанием данных в шкале частей.

Древовидные модели классификации инвариантны к шкалам количественных признаков. При изменении шкал просто пересчитываются пороговые значения, но не меняются правила в древовидных моделях классификации. Возможно, благодаря этому свойству классификатор типа «случайный лес» является одним из лучших при анализе данных разной природы.

В следующей статье будет показано, как использовать нормализацию признаков и оценки классификации при анализе данных различной природы.

Вклад авторов. В. В. Старовойтов определил план статьи и задачи, которые необходимо было решить при проведении исследований, принял участие в интерпретации результатов; Ю. И. Голуб описала представления данных разных типов при их анализе, выполнила экспериментальные исследования различных вариантов нормализации и стандартизации данных, предложила рекомендации по нормализации и стандартизации входных данных для решения задач классификации. Шкала частей для представления данных была разработана совместно.

Список использованных источников

1. Aksoy, S. Feature normalization and likelihood-based similarity measures for image retrieval / S. Aksoy, R. M. Haralick // *Pattern Recognition Letters*. – 2001. – Vol. 22, no. 5. – P. 563–582.
2. Singh, B. Investigating the impact of data normalization on classification performance / B. Singh // *Applied Soft Computing J.* – 2020. – Vol. 97. – P. 105524.
3. Nayak, S. C. Impact of data normalization on stock index forecasting / S. C. Nayak, B. B. Misra, H. S. Behera // *Intern. J. of Computer Information Systems and Industrial Management Applications*. – 2014. – Vol. 6. – P. 257–269.
4. Naeini, A. A. Assessment of normalization techniques on the accuracy of hyperspectral data clustering / A. A. Naeini, M. Babadi, S. Homayouni // *Intern. Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*. – 2017. – Vol. 42. – P. 27–30.
5. Stevens, S. S. On the theory of scales of measurement / S. S. Stevens // *Science. New Series*. – 1946. – Vol. 103, no. 2684. – P. 677–680.
6. Орлов, А. И. Теория измерений как часть методов анализа данных / А. И. Орлов // *Социология: методология, методы, математическое моделирование*. – 2012. – № 35. – С. 155–174.
7. Velleman, P. F. Nominal, ordinal, interval, and ratio typologies are misleading / P. F. Velleman, L. Wilkinson // *The American Statistician*. – 1993. – Vol. 47, no. 1. – P. 65–72.
8. Tukey, J. W. *Exploratory Data Analysis* / J. W. Tukey. – Massachusetts : Addison-Wesley, 1977. – P. 39–49.
9. Bruffaerts, C. A generalized boxplot for skewed and heavy-tailed distributions / C. Bruffaerts, V. Verardi, C. Vermandele // *Statistics & Probability Letters*. – 2014. – Vol. 95. – P. 110–117.

10. Kimber, A. C. Exploratory data analysis for possibly censored data from skewed distributions / A. C. Kimber // *Applied Statistics*. – 1990. – Vol. 39. – P. 21–30.
11. Carling, K. Resistant outlier rules and the non-Gaussian case / K. Carling // *Computational Statistics & Data Analysis*. – 2000. – Vol. 33, no. 3. – P. 249–258.
12. Hubert, M. An adjusted boxplot for skewed distributions / M. Hubert, E. Vandervieren // *Computational Statistics & Data Analysis*. – 2008. – Vol. 52, no. 12. – P. 5186–5201.
13. Brys, G. A robust measure of skewness / G. Brys, M. Hubert, A. Struyf // *J. of Computational and Graphical Statistics*. – 2004. – Vol. 13. – P. 996–1017.
14. Kyurkchiev, N. Sigmoid Functions: Some Approximation and Modelling Aspects / N. Kyurkchiev, S. Markov. – Saarbrücken : LAP Lambert Academic Publishing, 2015. – 120 p.
15. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. – М. : ДМК Пресс, 2015. – 402 с.
16. Bicego, M. Properties of the Box-Cox transformation for pattern classification / M. Bicego, S. Baldo // *Neurocomputing*. – 2016. – Vol. 218. – P. 390–400.
17. Zhang, Q. Weighted data normalization based on eigenvalues for artificial neural network classification / Q. Zhang, S. Sun // *Proc. of Intern. Conf. Neural Information Processing*. – 2009. – Vol. 5863. – P. 349–356. https://doi.org/10.1007/978-3-642-10677-4_39
18. Zadeh, L. A. Fuzzy sets / L. A. Zadeh // *Information and Control*. – 1965. – Vol. 8, no. 3. – P. 338–353.
19. Więckowski, J. How the normalization of the decision matrix influences the results in the VIKOR method? / J. Więckowski, W. Sałabun // *Procedia Computer Science*. – 2020. – Vol. 176. – P. 2222–2231.
20. Ioffe, S. Batch normalization: accelerating deep network training by reducing internal covariate shift / S. Ioffe, C. Szegedy // *32nd Intern. Conf. on Machine Learning, Lille, France, 7–9 July 2015*. – Lille, 2015. – Vol. 37. – P. 448–456.
21. Do we need hundreds of classifiers to solve real world classification problems? / M. Fernández-Delgado [et. al.] // *The J. of Machine Learning Research*. – 2014. – Vol. 15, no. 1. – P. 3133–3181.
22. Lemons, K. Comparison between Naive Bayes and random forest to predict breast cancer / K. A. Lemons // *Intern. J. of Undergraduate Research & Creative Activities*. – 2020. – Vol. 12, art. 12. – P. 1–5. <http://doi.org/10.7710/2168-0620.0287>
23. Chicco, D. The benefits of the Matthews correlation coefficient (MCC) over the diagnostic odds ratio (DOR) in binary classification assessment / D. Chicco, V. Starovoitov, G. Jurman // *IEEE Access*. – 2021. – Vol. 9. – P. 47112–47124. <https://doi.org/10.1109/ACCESS.2021.3068614>
24. Новиков, Д. А. Статистические методы в педагогических исследованиях (типовые случаи) / Д. А. Новиков. – М. : МЗ-Пресс, 2004. – 67 с.
25. Cheddad, A. On box-cox transformation for image normality and pattern classification // *IEEE Access*. – 2020. – Vol. 8. – P. 154975–154983. <https://doi.org/10.1109/ACCESS.2020.3018874>
26. Han, J. The influence of the sigmoid function parameters on the speed of backpropagation learning / J. Han, C. Moraga // *Intern. Workshop on Artificial Neural Networks, Malaga-Torremolinos, Spain, 7–9 June 1995*. – Malaga-Torremolinos, 1995. – P. 195–201.
27. Jain, A. Score normalization in multimodal biometric systems / A. Jain, K. Nandakumar, A. Ross // *Pattern Recognition*. – 2005. – Vol. 38, no. 12. – P. 2270–2285.

References

1. Aksoy S., Haralick R. M. Feature normalization and likelihood-based similarity measures for image retrieval. *Pattern Recognition Letters*, 2001, vol. 22, no. 5, pp. 563–582.
2. Singh B. Investigating the impact of data normalization on classification performance. *Applied Soft Computing Journal*, 2020, vol. 97, p. 105524.
3. Nayak S. C., Misra B. B., Behera H. S. Impact of data normalization on stock index forecasting. *International Journal of Computer Information Systems and Industrial Management Applications*, 2014, vol. 6, pp. 257–269.
4. Naeini A. A., Babadi M., Homayouni S. Assessment of normalization techniques on the accuracy of hyperspectral data clustering. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2017, vol. 42, pp. 27–30.
5. Stevens S. S. On the theory of scales of measurement. *Science. New Series*, 1946, vol. 103, no. 2684, pp. 677–680.
6. Orlov A. I. *Measurement theory as part of data analysis methods*. Sotsiologiya: metodologiya, metody, matematicheskoe modelirovanie [Sociology: Methodology, Methods, Mathematical Modeling], 2012, no. 35, pp. 155–174 (In Russ.).

7. Velleman P. F., Wilkinson L. Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 1993, vol. 47, no. 1, pp. 65–72.
8. Tukey J. W. *Exploratory Data Analysis*. Massachusetts, Addison-Wesley, 1977, pp. 39–49.
9. Bruffaerts C., Verardi V., Vermandele C. A generalized boxplot for skewed and heavy-tailed distributions. *Statistics & Probability Letters*, 2014, vol. 95, pp. 110–117.
10. Kimber A. C. Exploratory data analysis for possibly censored data from skewed distributions. *Applied Statistics*, 1990, vol. 39, pp. 21–30.
11. Carling K. Resistant outlier rules and the non-Gaussian case. *Computational Statistics & Data Analysis*, 2000, vol. 33, no. 3, pp. 249–258.
12. Hubert M., Vandervieren E. An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis*, 2008, vol. 52, no. 12, pp. 5186–5201.
13. Brys G., Hubert M., Struyf A. A robust measure of skewness. *Journal of Computational and Graphical Statistics*, 2004, vol. 13, pp. 996–1017.
14. Kyurkchiev N., Markov S. *Sigmoid Functions: Some Approximation and Modelling Aspects*. Saarbrücken, LAP Lambert Academic Publishing, 2015, 120 p.
15. Flach P. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. 1st ed. Cambridge University Press, 2012, 409 p.
16. Bicego M., Baldo S. Properties of the Box-Cox transformation for pattern classification. *Neurocomputing*, 2016, vol. 218, pp. 390–400.
17. Zhang Q., Sun S. Weighted data normalization based on eigenvalues for artificial neural network classification. *Proceedings of International Conference Neural Information Processing*, 2009, vol. 5863, pp. 349–356. https://doi.org/10.1007/978-3-642-10677-4_39
18. Zadeh L. A. Fuzzy sets. *Information and Control*, 1965, vol. 8, no. 3, pp. 338–353.
19. Więckowski J., Sałabun W. How the normalization of the decision matrix influences the results in the VIKOR method? *Procedia Computer Science*, 2020, vol. 176, pp. 2222–2231.
20. Ioffe S., Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, Lille, France, 7–9 July 2015*. Lille, 2015, vol. 37, pp. 448–456.
21. Fernández-Delgado M., Cernadas E., Barro S., Amorim D. Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research*, 2014, vol. 15, no. 1, pp. 3133–3181.
22. Lemons K. A. Comparison between Naive Bayes and random forest to predict breast cancer. *International Journal of Undergraduate Research & Creative Activities*, 2020, vol. 12, art. 12, pp. 1–5. <http://doi.org/10.7710/2168-0620.0287>
23. Chicco D., Starovoitov V., Jurman G. The benefits of the Matthews correlation coefficient (MCC) over the diagnostic odds ratio (DOR) in binary classification assessment. *IEEE Access*, 2021, vol. 9, pp. 47112–47124. <http://doi.org/10.1109/ACCESS.2021.3068614>
24. Novikov D. A. Statisticheskie metody v pedagogicheskikh issledovaniyakh (tipovye sluchai). *Statistical Methods in Pedagogical Research (Typical Cases)*. Moscow, MZ-Press, 2004, 67 p. (In Russ.).
25. Cheddad A. On box-cox transformation for image normality and pattern classification. *IEEE Access*, 2020, vol. 8, pp. 154975–154983. <http://doi.org/10.1109/ACCESS.2020.3018874>
26. Han J., Moraga C. The influence of the sigmoid function parameters on the speed of backpropagation learning. *International Workshop on Artificial Neural Networks, Malaga-Torremolinos, Spain, 7–9 June 1995*. Malaga-Torremolinos, 1995, pp. 195–201.
27. Jain A., Nandakumar K., Ross A. Score normalization in multimodal biometric systems. *Pattern Recognition*, 2005, vol. 38, no. 12, pp. 2270–2285.

Информация об авторах

Старовойтов Валерий Васильевич, доктор технических наук, профессор, главный научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: valerys@newman.bas-net.by

Голуб Юлия Игоревна, кандидат технических наук, доцент, старший научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: 6423506@gmail.com

Information about the authors

Valery V. Starovoitov, Dr. Sci. (Eng.), Professor, Chief Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
E-mail: valerys@newman.bas-net.by

Yuliya I. Golub, Cand. Sci. (Eng.), Associate Professor, Senior Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
E-mail: 6423506@gmail.com

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ INFORMATION TECHNOLOGIES



УДК 004.832.32
<https://doi.org/10.37661/1816-0301-2021-18-3-97-105>

Оригинальная статья
Original Paper

Применение логики первого порядка для выявления организаторов и исполнителей противоправных действий в преступных группах

А. М. Соболев^{1✉}, Е. И. Козлова¹, Ю. А. Чернявский²

¹Белорусский государственный университет,
ул. Курчатова, 5, 220064, Минск, Беларусь
✉E-mail: sobolat@bsu.by

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

Аннотация. Известны три основных семейства алгоритмов вывода в логике первого порядка: прямой логический вывод и его применение к дедуктивным базам данных и продукционным системам; процедуры обратного логического вывода и системы логического программирования; системы доказательства теорем на основе метода резолюции. При решении конкретных проблем наиболее эффективными являются алгоритмы, позволяющие охватить все факты и аксиомы, которые должны быть учтены в процессе логического вывода. Рассматривается пример применения логики первого порядка с целью выявления виновных в противоправных действиях. На основании высказываний формируется база знаний из выражений, с помощью которых составляется выражение логики первого порядка. Приводятся доказательства рассуждений, полученных в прямом логическом выводе с помощью дерева доказательств. Прямой логический вывод предусматривает выполнение допустимых этапов логического вывода на основе всех известных фактов, поэтому также рассматривается метод на основе резолюции при реализации обратного логического вывода с учетом выражения, полученного при прямом логическом выводе. Данное выражение преобразуется в конъюнктивную нормальную формулу с помощью законов булевой алгебры и доказывается методом исключения событий с помощью операции конъюнкции.

Ключевые слова: логика первого порядка, прямой логический вывод, обратный логический вывод, формирование рассуждений, метод резолюции, конъюнктивная нормальная формула, дерево доказательства

Для цитирования. Соболев, А. М. Применение логики первого порядка для выявления организаторов и исполнителей противоправных действий в преступных группах / А. М. Соболев, Е. И. Козлова, Ю. А. Чернявский // Информатика. – 2021. – Т. 18, № 3. – С. 97–105. <https://doi.org/10.37661/1816-0301-2021-18-3-97-105>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 16.11.2020
Подписана в печать | Accepted 23.07.2021
Опубликована | Published 29.09.2021

Application of first-order logic to identify organizers and perpetrators of illegal actions in teams of a limited circle of people

Alexander M. Sobol¹✉, Elena I. Kozlova¹, Yuri A. Chernyavsky²

¹Belarusian State University,
st. Kurchatova, 5, 220064, Minsk, Belarus

✉E-mail: sobolam@bsu.by

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

Abstract. There are three main families of inference algorithms in first-order logic: direct inference and its application to deductive databases and production systems; backward inference procedures and logic programming systems; theorem proving systems based on the resolution method. When solving specific problems, the most effective algorithms are those that allow you to cover all the facts and axioms and must be taken into account in the process of inference. An example is considered in which it is necessary to prove the guilt of a person in murder. On the basis of statements, a knowledge base is formed from expressions, with the help of which an expression of first-order logic is compiled and proved using direct logical inference. The proof of the reasoning obtained in direct inference using the proof tree is given. However, direct inference provides for the implementation of all admissible stages of logical inference based on all known facts. The article also considers a method based on the resolution when implementing the reverse inference, taking into account the expression obtained in the direct inference. This expression is converted into a conjunctive normal formula using the laws of Boolean algebra and is proved by the elimination of events using the conjunction operation.

Keywords: first-order logic, direct inference, reverse inference, reasoning formation, resolution method, conjunctive normal formula, proof tree

For citation. Sobol A. M., Kozlova E. I., Chernyavsky Y. A. Application of first-order logic to identify organizers and perpetrators of illegal actions in teams of a limited circle of people. *Informatics*, 2021, vol. 18, no. 3, pp. 97–105 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-97-105>

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

Введение. Вопросы тактики выявления руководителей преступных групп и организации борьбы с ними до настоящего времени практически не изучались с использованием математических алгоритмов и математической логики. Большинство научных публикаций^{1–4} отражают результаты исследований с применением юридических, психологических, исторических, организационных и технических подходов [1–3]. В то же время использование математического аппарата логики при формировании выводов о степени участия подозреваемых в совершении уголовно наказуемых деяний может существенно облегчить процесс принятия следователями решений в ходе проводимых расследований.

Рассматриваемый в настоящей статье пример формирования рассуждений с помощью определенных выражений логики первого порядка с использованием прямого логического вывода и на основе резолюции при реализации обратного логического вывода призван продемонстрировать некоторые возможности применения математического аппарата при принятии решений специалистами в области юриспруденции, а также обучения специалистов в области информационных

¹Мазунин, Я. М. Тактика выявления и доказывания вины организаторов преступных групп : автореф. дис. ... канд. юр. наук / Я. М. Мазунин ; Юр. ин-т МВД России. – М., 1996. – 23 с.

²Корнакова, С. В. Логика для юристов : учеб. пособие / С. В. Корнакова. – Иркутск : Изд-во БГУЭП, 2015. – 122 с.

³Быков, В. М. Тактика выявления организаторов преступных групп : учеб. пособие / В. М. Быков, А. В. Куница. – Саратов : Изд-во Саратов. юр. ин-та МВД России, 2005. – 224 с.

⁴Математическая логика и теория алгоритмов для программистов : учеб. пособие / Д. В. Гринченков, С. И. Потоцкий. – М. : Кнорус, 2014. – 206 с.

технологий применению знаний и навыков построения алгоритмов и информационных (экспертных, рекомендательных) систем на примерах из областей, отличных от естественно-научных или технических.

Описание проблемы: в небольшом городе районного значения участились события криминального характера, связанные в основном с работниками городской клинической больницы. Цель расследования – доказать, что гражданин Ловелас, который представляется бизнесменом, является руководителем преступной группировки и лично совершал преступления.

Представление хода рассматриваемых событий на естественном языке. Во время возвращения из областного центра в район попал в аварию автобус междугороднего сообщения, в котором находилась и в результате аварии погибла жена заведующего хирургическим отделением городской больницы (гражданка Родионова). В этом же автобусе возвращалась на свою малую родину для устройства на работу в клинику доктор Стрелкова, хирург по специальности. В автомобильной аварии доктор Стрелкова не пострадала, а по прибытии в больницу была принята на постоянную работу на должность хирурга в отделение, заведующим которого являлся Родионов.

Событие 1. Когда доктор Стрелкова возвращалась домой, на нее напала группа хулиганов, но их быстро разогнал неожиданно оказавшийся рядом мужчина спортивного вида, который позже представился доктору бизнесменом Ловеласом. С этого момента и до определенного времени между доктором Стрелковой и гражданином Ловеласом сохранялись хорошие личные отношения. При этом, однако, она не могла не отметить определенные странности поведения Ловеласа: частые и неожиданные отъезды, объясняемые необходимостью решать неотложные вопросы с компаньонами по бизнесу, скрытность в сочетании с тревожностью.

Несколько позже о своем существовании и противоправной деятельности в окрестностях места жительства доктора Стрелковой напомнила группа хулиганов, которая на этот раз ограбила и жестоко избила молодую женщину. Доктор Стрелкова на служебном автомобиле медицинской помощи доставила пострадавшую в хирургическое отделение районной больницы, лично прооперировала, после чего распорядилась положить больную в палату хирургического отделения для пациентов, находящихся на стационарном лечении. После завершения операции пришедшая в себя пострадавшая сообщила доктору Стрелковой важную новость о том, что, когда ограбившие и избившие ее бандиты ушли, к ней подошел мужчина (некто), нанес ей ножевые ранения и быстро удалился.

При очередной встрече с Ловеласом доктор рассказала ему о случившемся. Ловелас заинтересовался состоянием здоровья потерпевшей после операции и получил обнадеживающий ответ об успешном проведении операции и скором выздоровлении пациентки.

Событие 2. Незнакомые сотрудникам больницы мужчины попытались проникнуть в палату хирургического отделения, в которой лежала прооперированная женщина. Они не смогли сразу реализовать свое намерение, поскольку у входной двери в палату был установлен пост полиции. Один мужчина при виде полицейского сразу покинул здание больницы. Некоторое время спустя возник пожар в одной из комнат больницы, не занятой людьми. Когда пожар начал распространяться по территории больницы, сработала пожарная сигнализация и полицейский оставил вверенный ему пост охраны. Некто вошел в палату потерпевшей, задушил ее подушкой и после этого попытался незаметно скрыться.

Для расследования преступления на территорию больницы прибыла оперативная группа полиции. Просмотр записей камер видеонаблюдения позволил выявить выходящего из здания человека, не сотрудника больницы, в котором полицейский, охранявший палату убитой, опознал одного из мужчин, приходивших к палате до пожара. На записи было видно, что он сел в припаркованную у здания машину и уехал. Кроме того, полицейский узнал в выходящем во время пожара из больницы мужчине второго визитера, который также не являлся сотрудником больницы. Примерно в то же самое время доктор Стрелкова увидела мирно сидящего на лавочке в садике больницы своего знакомого Ловеласа. Об этом она доложила заведующему хирургическим отделением Родионову, который также знал этого человека. Ранее подруга покойной жены Родионова показывала ему сохраненную ею совместную фотографию его жены и Ловеласа.

Событие 3. Родионов, обеспокоенный судьбой Стрелковой, прибыл в ее квартиру, где его встретил Ловелас с ножом в руках. В результате возникшего противоборства Родионов получил легкое ножевое ранение, а Ловелас оказался в полузадушенном состоянии. Прибывшие на место события полицейские и следователи составили протокол происшествия. В качестве позитива можно отметить состоявшуюся вскоре после изложенных событий свадьбу Родионова и Стрелковой.

Особенность применения алгоритма прямого логического вывода в рассматриваемом примере. Особенность заключается в том, что на любом этапе невозможны какие-либо новые логические выводы на основании высказываний, заключения по которым уже явно имеются в базе знаний. Необходимо доказать следующее: из описанных выше событий следует, что гражданин Ловелас является преступником. Вначале имеющиеся факты представим в виде определенных выражений в логике первого порядка [3].

Все события в виде высказываний можно представить следующим образом. Ловелас (*lov*) спас Стрелкову (*s*) (произошла встреча Ловеласа и Стрелковой). После этого банда (*n*) напала на девушку (*girl*) и некто (*x*) нанес ей ножевые ранения (событие 1).

Далее кто-то попытался пробраться в палату девушки, но она охранялась полицейскими (*pol*), для устранения охраны был организован пожар (*f*). В результате охрана отвлеклась и девушка была убита (событие 2).

При просмотре камер видеонаблюдения оперативная группа выявила выходящего во время пожара кого-то, в итоге произошло противоборство на квартире (*flat_s*) Стрелковой (событие 3).

Введем обозначения событий и высказываний (таблица).

Обозначения переменных и высказываний

Variables notation and description

Обозначение <i>Variable</i>	Пояснение <i>Description</i>
<i>lov</i>	Постоянная, обозначающая Ловеласа
<i>n</i>	Постоянная, обозначающая банду
<i>girl</i>	Постоянная, обозначающая девушку
<i>x</i>	Постоянная, обозначающая кого-то, он же «некто»
<i>pol</i>	Постоянная, обозначающая полицейских
<i>f</i>	Постоянная, обозначающая пожар
<i>r</i>	Постоянная, обозначающая Родионова
<i>s</i>	Постоянная, обозначающая Стрелкову
<i>flat_s</i>	Постоянная, обозначающая квартиру Стрелковой
<i>Disperse (lov, s)</i>	Встреча Ловеласа и Стрелковой
<i>Event₁ (x, n, girl)</i>	Высказывание, обозначающее событие 1
<i>Event₂ (f, x, pol)</i>	Высказывание, обозначающее событие 2
<i>Event₃ (lov, r, flat_s)</i>	Высказывание, обозначающее событие 3
<i>Criminal (x)</i>	Преступник (объект <i>x</i>)
<i>Attack (n, girl)</i>	Нападение объекта <i>n</i> на объект <i>girl</i>
<i>Save (lov, s)</i>	Спасение объектом <i>lov</i> объекта <i>s</i>
<i>Organizer (x, f)</i>	Организатор события <i>f</i> с участием объекта <i>x</i>
<i>Leave (pol)</i>	Уход объекта <i>pol</i>
<i>Murder (p, girl)</i>	Убийство объектом <i>p</i> объекта <i>girl</i>
<i>Meet (lov, r)</i>	Встреча объектов <i>lov</i> и <i>r</i>

Рассмотрим высказывание

$$\exists z \text{Attack}(n, s) \wedge \text{Save}(z, s) \Rightarrow \text{Disperse}(z, s). \quad (1)$$

Высказывание (1) с квантором существования означает, что существует некоторый объект (*z*), удовлетворяющий определенному условию: объект *z* спас Стрелкову от банды, в результате чего встретились Ловелас и Стрелкова. В реальной ситуации Ловелас (объект *lov*) спас Стрел-

кову от банды. Импликация в выражении (1) означает, что после двух событий состоялась встреча. В этом случае при атаке банды на девушку и спасении кем-то незнакомым ее жизни встретились два объекта: кто-то (Ловелас) и Стрелкова. Для устранения квантора существования введем в высказывание константу *lov*:

$$Attack(n, s) \wedge Save(lov, s) \Rightarrow Disperse(lov, s). \quad (2)$$

Далее состоялось другое нападение банды: кто-то после нападения подошел к обессиленной, лежащей на земле девушке и нанес ей тяжелые ножевые ранения:

$$Attack(n, girl) \wedge Attack(x, girl) \Rightarrow Event_1(x, n, girl). \quad (3)$$

Следует установить правомерность действия полицейского (*pol*), оставившего во время пожара (*f*) вверенный ему пост охраны у палаты раненой девушки:

$$Organizer(x, f) \wedge Leave(pol) \wedge Murder(x, girl) \Rightarrow Event_2(f, x, pol). \quad (4)$$

Также необходимо учесть события, которые произошли в квартире гражданки Стрелковой (*y*). Когда Родионов (*r*) пришел в квартиру, его встретил Ловелас с ножом в руках. Произошло противостояние, в результате которого Родионов получил легкое ножевое ранение, а Ловелас был обезврежен:

$$Meet(lov, r) \wedge Attack(lov, r) \Rightarrow Event_3(lov, r, flat_s). \quad (5)$$

С учетом введенных обозначений приведем высказывание (6), описывающее все события:

$$Disperse(lov, s) \wedge Event_1(x, n, girl) \wedge Event_2(f, x, pol) \wedge Event_3(lov, r, flat_s). \quad (6)$$

Используя выражение (6), требуется доказать, что Ловелас является преступником *Criminal(lov)*:

$$Disperse(lov, s) \wedge Event_1(x, n, girl) \wedge Event_2(f, x, pol) \wedge Event_3(lov, r, flat_s) \Rightarrow Criminal(lov). \quad (7)$$

Применение правила резолюции для доказательства того, что гражданин Ловелас является руководителем преступной группировки и лично совершал преступления. Правило резолюции [1] чаще всего используется при логических выводах для выражений, представленных в конъюнктивной нормальной форме (КНФ), и заключается в следующем. Пусть s_1, s_2 – дизъюнкты КНФ S , x – булева переменная. Если $x \in s_1, \bar{x} \in s_2$, то дизъюнкт $r = (S \setminus \{x\}) \cup (S \setminus \{\bar{x}\})$ является логическим следствием КНФ S . Дизъюнкт r называется резольвентой дизъюнктов s_1, s_2 , при этом КНФ S и $S \vee r$ эквивалентны. Доказательство на основе резолюции заключается в получении резольвент для заданной КНФ до тех пор, пока не будет получена пара однолитеральных дизъюнктов x_i, \bar{x}_i , что приведет к выводу пустого дизъюнкта и доказательству $S = 0$ – невыполнимости КНФ S .

Согласно правилу резолюции доказательство того, что из базы знаний (в рассматриваемом случае обозначаемой KB) следует высказывание α ($KB \models \alpha$), осуществляется путем доказательства невыполнимости выражения $KB \wedge \neg\alpha$, т. е. путем получения пустого выражения. Логические высказывания (1)–(7) представим в КНФ для доказательства с помощью метода резолюции. Для примера представим в КНФ выражение (7). По законам алгебры логики

$$A \Rightarrow B \Leftrightarrow \neg A \vee B, \quad \neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B, \quad (8)$$

поэтому выражение (7) примет вид

$$Disperse(lov, y) \vee \neg Event_1(x, n, girl) \vee \neg Event_2(f, x, pol) \vee \vee \neg Event_3(lov, r, flat_s) \vee Criminal(lov). \tag{9}$$

Остальные выражения будут выглядеть следующим образом:

$$\begin{aligned} &\neg Attack(n, s) \vee \neg Save(lov, s) \vee Disperse(lov, s); \\ &\neg Attack(n, girl) \vee \neg Attack(x, girl) \vee Event_1(x, n, girl); \\ &\neg Murder(x, girl) \vee Event_2(f, x, pol); \\ &\neg Attack(x, r) \vee Event_3(x, r, flat_s); \\ &Attack(n, s); \\ &Save(lov, s); \\ &Attack(n, girl); \\ &Attack(p, girl); \\ &Murder(p, girl); \\ &Attack(lov, r); \\ &Organizer(x, f); \\ &Leave(pol). \end{aligned} \tag{10}$$

Между тем множество высказываний (10) включает и отрицаемую цель $\neg Criminal(lov)$. Также в него входят $Organizer(x, f)$ и $Leave(pol)$, однако при использовании метода резолюции рассматривать их не будем. Представленная база знаний не содержит функциональных символов, в отсутствие которых в логике первого порядка логический вывод существенно упрощается.

Выражение (7) доказывается с помощью алгоритма прямого логического вывода, сформированного в виде дерева. (Алгоритм следует читать снизу вверх и слева направо.) Истинность факта $Criminal(lov)$ подтверждается доказательством четырех конъюнктов из выражения (7). Когда алгоритм достигает последнего конъюнкта, переменная x (некто) заменяется константой lov . Нижний уровень дерева соответствует первоначальным фактам, средний уровень – фактам, выведенным логическим путем в первой итерации. Выведенный в средней итерации факт соответствует верхнему уровню.

На основании импликаций (1)–(6) реализуем следующие итерации. В первой итерации выражение (7) имеет невыполненные предпосылки, а (3), (4) выполняются с подстановкой $\{lov/x\}$. Во второй итерации выражение (7) выполняется путем добавления высказывания $Criminal(lov)$.

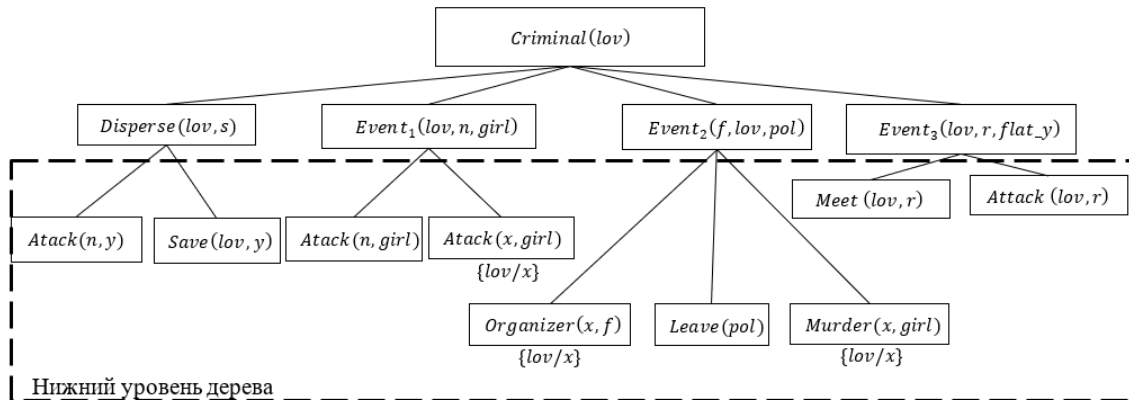


Рис. 1. Дерево доказательства, сформированного путем прямого логического вывода, что гражданин Ловелас является преступником
Fig. 1. The tree formed by direct inference evidence that the person Lovelace is a criminal

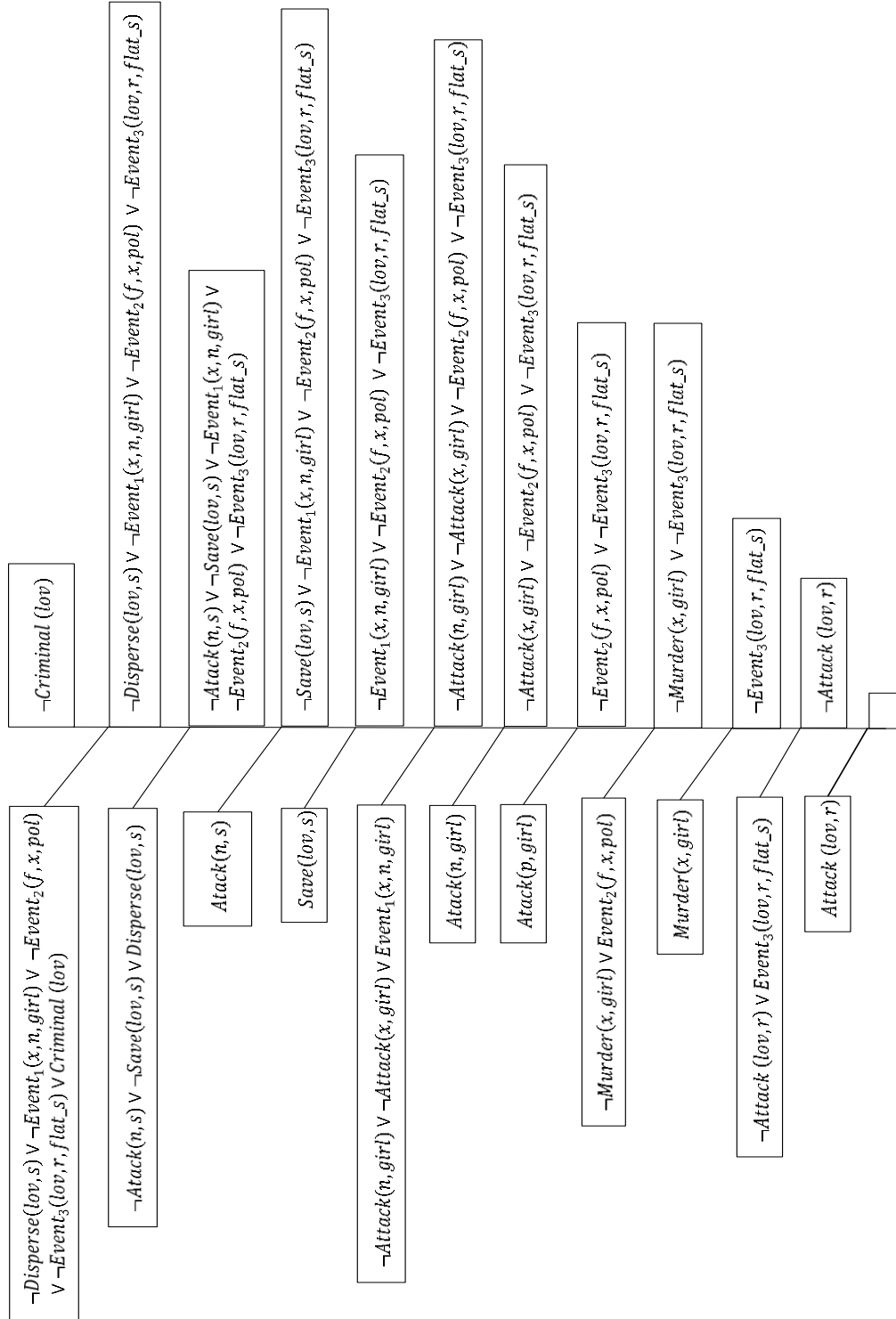


Рис. 2. Процедура доказательства с помощью резолюции того, что гражданин Ловелас организовал и лично совершил уголовно наказуемые преступления
 Fig 2. The procedure for proving, by the resolution, that the person Lovelace organized and personally committed criminal offenses

При работе с алгоритмом прямого логического вывода (см. рис. 1) возможны следующие источники осложнений:

1. «Внутренний цикл» алгоритма предполагает поиск всех возможных унификаторов, так как предпосылка некоторого правила унифицируется с подходящим множеством факторов в базу знаний. Такая вычислительная операция, называемая согласованием с шаблоном, может быть весьма ресурсозатратной.

2. Имеет место повторная проверка каждого правила во всех итерациях на предмет выполнения его предпосылки, даже если в каждой итерации в базе знаний вносится лишь очень немного дополнений.

3. Алгоритм может генерировать много фактов (резольвент), которые не имеют отношения к текущей цели.

Обратный логический вывод – частный случай резолюции, в котором применяется стратегия определения того, какая операция резолюции должна быть выполнена в следующую очередь.

В рассматриваемом случае структура доказательства по методу резолюции показана на рис. 2. Структура состоит из единственного «хребта», который начинается с целевого выражения. Выражения, расположенные вдоль «хребта», точно соответствуют последовательным значениям целевых переменных в алгоритме обратного логического вывода. Для метода резолюции всегда выбирается выражение, положительный литерал которого унифицируется с самым левым литералом текущего выражения в «хребте» (справа). Метод резолюции применяется к выражениям из базы знаний до тех пор, пока не образуется пустое выражение, которое обозначено квадратом на рис. 2. Приведем пример работы метода резолюции. Для начала возьмем выражение с левой стороны «хребта» и сделаем операцию конъюнкции с его правой частью:

$$\begin{aligned} & \neg Disperse(lov, s) \vee \neg Event_1(x, n, girl) \vee \neg Event_2(f, x, pol) \vee \\ & \vee \neg Event_3(lov, r, flat_s) \vee Criminal(lov) \wedge \neg Criminal(lov) \Leftrightarrow \\ \Leftrightarrow & \neg Disperse(lov, s) \vee \neg Event_1(x, n, girl) \vee \neg Event_2(f, x, pol) \vee \neg Event_3(lov, r, flat_s). \end{aligned} \quad (11)$$

Следует иметь в виду, что метод резолюции нельзя использовать для выработки всех логических следствий из множества высказываний. Он позволяет осуществлять поиск ответов на данный конкретный вопрос с помощью метода отрицания цели. Кроме того, алгоритм обратного логического вывода (в сравнении с алгоритмом прямого логического вывода) отличается наличием повторяющихся состояний и неполнотой. Используемая в рассмотренном примере весьма ограниченная по объему база знаний не позволила проиллюстрировать как относительные преимущества, так и недостатки алгоритма прямого логического вывода и алгоритма, основанного на применении правила резолюции. Отмеченные проблемы можно было бы решить путем существенного расширения участия органов охраны правопорядка в расследовании преступлений, совершенных не только руководителем криминальной группировки, но и ее членами, используя при этом соответствующие региональные базы знаний.

Заключение. В рассмотренном примере показано, как можно применять алгоритм прямого логического вывода к определенным выражениям в логике первого порядка и каким образом он может быть реализован эффективно. Использование выражений типа *Situation* \Rightarrow *Response* (ситуация – отклик) особенно полезно в тех случаях, когда логический вывод осуществляется в ответ на вновь поступающую информацию. Представленные в работе выводы получены без привлечения инструментальных средств, поскольку основной задачей статьи являлось доказательство того, что в случае возможности формирования базы знаний на основе только определенных выражений создание рассуждений с помощью прямого или обратного логического вывода может оказаться более предпочтительным и позволит избежать издержек, характерных для метода резолюции при использовании высказываний общего значения.

В дальнейшем предполагается построить описание предметной области, решить поставленную задачу с помощью языка логического программирования Prolog и проверить полученные в работе выводы.

Вклад авторов. А. М. Соболев представил ход событий с помощью логических выражений, доказал, что Ловелас является преступником, с помощью метода резолюции, редактировал текст статьи; Е. И. Козлова участвовала в создании графического материала и редактировании статьи, подобрала литературу, доказала, что Ловелас является преступником, с помощью прямого и обратного логических выводов; Ю. А. Чернявский представил ход рассматриваемых событий на естественном языке.

Список использованных источников

1. Самойлов, А. В. О применении основных логических законов и последствиях их нарушения в уголовно-процессуальной деятельности / А. В. Самойлов, Д. Е. Снегирева // Auditorium : электр. науч. журнал Курского гос. ун-та. – 2014. – № 4. – С. 96–102.
2. Владимиров, Л. Е. Учение об уголовных доказательствах / Л. Е. Владимиров. – Тула : Автограф, 2000. – 464 с.
3. Рассел, С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг. – 2-е изд. – М. : Вильямс, 2006. – 1408 с.

References

1. Samoilo A. V., Snegireva D. E. *On the application of basic logical laws and the consequences of their violation in criminal procedural activity*. Auditorium : jelektronnyj nauchnyj zhurnal Kurskogo gosudarstvennogo universiteta [Auditorium: Electronic Scientific Journal of Kursk State University], 2014, no. 4, pp. 96–102 (In Russ.).
2. Vladimirov L. E. *Uchenie ob ugovolnyh dokazatel'stvah. The Doctrine of Criminal Evidence*. Tula, Autograph, 2000, 464 p. (In Russ.).
3. Russell S., Norvig P. *Artificial Intelligence: A Modern Approach*. 2nd ed., Prentice Hall, 2002, 1408 p.

Информация об авторах

Соболев Александр Михайлович, магистр физико-математических наук, аспирант кафедры интеллектуальных систем, факультет радиоп физики и компьютерных технологий, Белорусский государственный университет.

E-mail: sobolam@bsu.by

Козлова Елена Ивановна, кандидат физико-математических наук, доцент, заведующий кафедрой интеллектуальных систем, факультет радиоп физики и компьютерных технологий, Белорусский государственный университет.

E-mail: kozlova@bsu.by

Чернявский Юрий Александрович, кандидат технических наук, доцент кафедры информатики, факультет компьютерных систем и сетей, Белорусский государственный университет информатики и радиоэлектроники.

E-mail: chernyavskiy@bsuir.by

Information about the authors

Alexander M. Sobol, M. Sci. (Phys.-Math.), Postgraduate Student, Department of Intelligent Systems, Faculty of Radiophysics and Computer Technologies, Belarusian State University.

E-mail: sobolam@bsu.by

Elena I. Kozlova, Cand. Sci. (Phys.-Math.), Head of the Department of Intelligent Systems, Faculty of Radiophysics and Computer Technologies, Belarusian State University.

E-mail: kozlova@bsu.by

Yuri A. Chernyavsky, Cand. Sci. (Eng.), Associate Professor of the Department of Informatics, Faculty of Computer Systems and Networks, Belarusian State University of Informatics and Radioelectronics.

E-mail: chernyavskiy@bsuir.by



УДК 528.85; 528.87; 528.88; 528.94
<https://doi.org/10.37661/1816-0301-2021-18-3-106-114>

Оригинальная статья
Original Paper

Совершенствование инструментальных средств обработки и анализа космической информации

С. А. Золотой¹, И. Б. Страшко¹, Д. С. Котов^{1✉}, И. М. Нестерович¹, В. В. Рубо¹, К. И. Костюк²

¹УП «Геоинформационные системы»,
ул. Сурганова, 6, Минск, 220012, Беларусь
✉E-mail: dkotau@gis.by

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

Аннотация. Показывается необходимость создания информационной системы, имеющей специализированные сервисы, которые позволяют ученым и специалистам выполнять тематическую обработку данных дистанционного зондирования Земли, изменяя определенным образом параметры обработки данных, и самостоятельно анализировать полученную информацию.

Рассматривается разработка специализированной информационной системы для оперативного предоставления информации на примере создания программного комплекса распространения оперативной космической информации и работающего во взаимодействии с ним программного комплекса прогноза урожайности зерновых культур.

Представляется структура программного комплекса распространения оперативной космической информации. Приводятся технические и функциональные характеристики его отдельных элементов. Показываются связи подсистем комплекса при приеме, обработке, анализе данных и предоставлении информации заинтересованным специалистам и исследователям. Обращается внимание на возможности широкого применения комплекса для решения научных и прикладных задач в различных областях знаний.

Описывается практическое применение вегетационных индексов при анализе оперативной космической информации. Приводятся данные о практических результатах прогнозирования развития зерновых сельскохозяйственных культур и количественных показателях сеансов с космических аппаратов. Показывается возможность дальнейшего совершенствования и развития созданной информационной системы.

Ключевые слова: дистанционное зондирование Земли, тематическая обработка данных, космическая информация, спутниковая метеорология, данные о парниковых газах, изучение облачного покрова, содержание углекислого газа, мониторинг чрезвычайных ситуаций, прогнозирование урожайности, вегетационный индекс

Для цитирования. Совершенствование инструментальных средств обработки и анализа космической информации / С. А. Золотой [и др.] // Информатика. – 2021. – Т. 18, № 3. – С. 106–114. <https://doi.org/10.37661/1816-0301-2021-18-3-106-114>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 15.05.2021
Подписана в печать | Accepted 18.06.2021
Опубликована | Published 29.09.2021

The improvement of the instrumental means for processing and analyzing of space information

Sergey A. Zolotoy¹, Igor B. Strashko¹, Dzmitry S. Kotau^{1✉}, Iryna M. Nestsiarovich¹, Vitali V. Rouba¹, Karolina I. Kostyuk²

¹UE "Geoinformation Systems",
st. Surganova, 6, Minsk, 220012, Belarus
✉E-mail: dkotau@gis.by

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

Abstract. The need of the creation of information system with specialized services which allow scientists and specialists to perform thematic processing of Earth remote sensing data, changing the data processing parameters in a certain way, and independently analyze the information received is shown.

To achieve the goal of rapid provision of information, the development of specialized information system is considered on the example of the creation of the software package for the dissemination of operational space information, as well as a software package for predicting the yield of grain crops.

The structure of the software package for the dissemination of operational space information is shown. The technical and functional characteristics of the separate components of the package are given. The relations of the subsystems of the package are shown during data receiving, processing, analyzing and providing information to interested specialists and researchers. Attention is drawn to the possibilities of wide application of the software package for the dissemination of operational space information for solving scientific and applied problems in various fields of knowledge.

The practical application of vegetation indexes in the analysis of operational space information is shown. The data of the practical results of forecasting of the grain crops development, as well as the quantitative indicators of satellites sessions are presented. The possibility of further improvement and development of the created information system is shown.

Keywords: remote sensing of the Earth, thematic data processing, space information, satellite meteorology, greenhouse gas data, cloud cover studying, carbon dioxide content, emergency monitoring, yield forecasting, vegetation index

For citation. Zolotoy S. A., Strashko I. B., Kotau D. S., Nestsiarovich I. M., Rouba V. V., Kostyuk K. I. The improvement of the instrumental means for processing and analyzing of space information. *Informatics*, 2021, vol. 18, no. 3, pp. 106–114 (In Russ.). <https://doi.org/10.37661/1816-0301-2021-18-3-106-114>

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

Введение. Для всех организаций, осуществляющих дистанционное зондирование Земли (ДЗЗ) из космоса, актуальной задачей является предоставление полученных данных ученым и специалистам, ведущим исследования различных явлений и процессов, происходящих в окружающей среде. Данная задача решалась разработчиками УП «Геоинформационные системы» – Национального оператора Белорусской космической системы дистанционного зондирования Земли (БКСДЗ) – в рамках создания специализированной информационной системы, которая не только предоставляет ученым и специалистам данные дистанционного зондирования, но и обеспечивает возможности поиска и выборки необходимых пользователям данных, а также работы с различными инструментами для проведения удаленного анализа полученной информации. При создании подобной информационной системы важными факторами, определяющими ее архитектуру и принципы функционирования, являются разработка соответствующих сервисов и обеспечение их дальнейшего совершенствования. Необходимость в создании такой системы связана с тем, что формирование структуры приема спутниковых данных, систем хранения и тематической обработки больших объемов информации требует достаточно серьезных финансовых вложений.

Для обеспечения специалистов Республики Беларусь, ведущих исследования в спутниковой метеорологии, возможностью доступа к оперативным данным ДЗЗ и с целью формирования инструментов их тематической обработки и анализа в рамках государственной программы «Научно-технологические и технические» на 2016–2020 гг. выполнена опытно-конструкторская работа «Создание распределенной системы приема, обработки и распространения оперативной космической информации с космических аппаратов (КА) (AQUA, SUOMI NPP, NOAA 20, MetOp, Feng-Yun 3D)». Результатом ее выполнения стало создание программного комплекса распространения оперативной космической информации (ПК РОКИ) и программного комплекса обнаружения тепловых аномалий (ПК ОТА).

В настоящей статье представлен опыт создания ПК РОКИ, а также работающего во взаимодействии с ним программного комплекса прогноза урожайности зерновых культур (ПК «Прогноз ЗК»).

Назначение и архитектура ПК РОКИ. Комплекс предоставляет данные о европейской территории, получаемые с КА AQUA, TERRA, SUOMI NPP, NOAA 20, MetOp A/B/C, NOAA 18, NOAA 19 и Feng-Yun 3D, на станцию приема Национального оператора БКСДЗ в режиме времени, близкого к реальному. Распространяемые данные хранятся и доступны в течение не менее шести месяцев с момента приема. Подробная информация о спутниках и датчиках, используемых для получения информации для ПК РОКИ, приведена на сайте Всемирной метеорологической организации¹.

В рамках опытно-конструкторской работы в УП «Геоинформационные системы» вместе с антенной системой создана и установлена новая приемная станция (ПРС) X/L, которая предназначена для приема данных, передаваемых с метеорологических КА ДЗЗ по радиоканалам X/L-диапазонов частот.

Характеристики ПРС X/L:

- расчет расписания пролета КА ДЗЗ по начальным условиям NORAD;
- одновременный прием радиосигналов по двум каналам круговой поляризации правого направления вращения в L-диапазоне частот от 1,69 до 1,71 ГГц и X-диапазоне частот от 7,75 до 8,4 ГГц;
- диаметр параболического зеркала ПРС X/L 3,5 м;
- синхронизация времени по сигналам GPS;
- время подготовки станции к очередному сеансу связи с КА ДЗЗ после завершения предыдущего сеанса связи не более 3 мин;
- масса станции 1200 кг.

Параболическое зеркало диаметром 3,5 м установлено на опорно-поворотном устройстве, которое имеет две оси вращения. Зеркальная система ПРС X/L состоит из рефлектора и подвеса облучателя.

За счет двойного преобразования частот в L-диапазоне обеспечивается гарантированный прием сигналов в диапазоне работы станции на частотах 1,690–1,710 ГГц.

Спутниковые данные, поступающие с приемной антенны, проходят автоматизированную обработку и размещаются на серверах в виде базовых продуктов обработки (в основном с уровнем обработки L1B). ПК РОКИ ведет мониторинг положения метеоспутников на орбите, их полосы обзора и зоны радиовидимости антенны, что дает возможность планирования и предикативного анализа данных. Также комплекс осуществляет настройку уведомлений по электронной почте для зарегистрированных пользователей о поступлении в каталог интересующих их данных ДЗЗ.

ПК РОКИ сокращает время на поиск и получение информации специалистам и исследователям, занимающимся следующими научными и практическими разработками:

- получением данных о температуре и влажности воздуха, парниковых газах в разных слоях атмосферы;
- изучением облачного покрова и интенсивности осадков, в том числе в мощных кучево-дождевых облаках;

¹WMO Oscar | List of all instruments [Electronic resource]. – Mode of access: <https://www.wmo-sat.info/oscar/instruments>. – Date of access: 19.04.2021.

- анализом изменений гидрометеопоказателей для повышения достоверности краткосрочных и долгосрочных прогнозов погоды;
- измерением параметров атмосферных аэрозолей, созданием моделей содержания углекислого газа, оценкой качества воздуха;
- мониторингом отдельных видов чрезвычайных ситуаций, моделированием распространения природных пожаров и зон затопления;
- наблюдениями за движением и температурой льдов, картографированием и прогнозированием их динамики;

– мониторингом состояния сельскохозяйственных культур и прогнозированием урожайности.

Доступ к информации программного комплекса можно получить после прохождения простой процедуры регистрации в клиентском приложении².

Архитектура ПК РОКИ и схема информационного взаимодействия с пользователями и другими компонентами распределенной системы приема, обработки и распространения оперативной космической информации показаны на рис. 1.

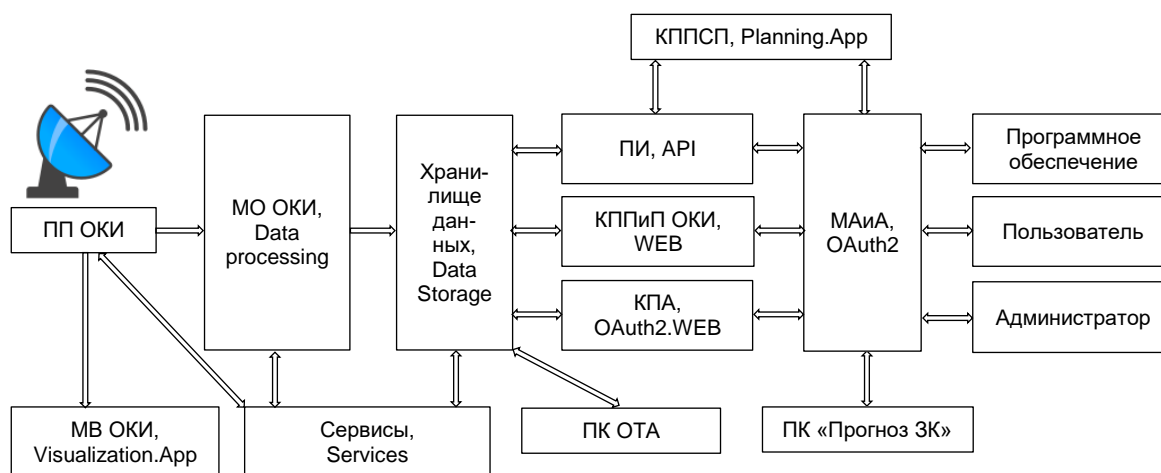


Рис. 1. Архитектура ПК РОКИ

Fig. 1. Architecture of the operational space information dissemination software package

ПК РОКИ состоит из программного интерфейса, модуля обработки оперативной космической информации, сервисов для поддержки работы системы, клиентского приложения администрирования – интерфейса модуля аутентификации и авторизации (МАиА, OAuth2), клиентского приложения просмотра и поиска оперативной космической информации, модуля визуализации оперативной космической информации в процессе приема данных, клиентского приложения планирования сеансов приема.

Подсистема приема оперативной космической информации (ПП ОКИ), ПК «Прогноз ЗК», ПК ОТА и хранилище данных не входят в состав ПК РОКИ, но являются компонентами распределенной системы приема, обработки и распространения оперативной космической информации.

Программный интерфейс (ПИ, API) для обеспечения взаимодействия с внешними программными комплексами реализован на основе технологии Representational State Transfer (REST). Формат структур входных и выходных данных – XML или JSON. Описание используемых методов можно посмотреть на сайте³.

²MeteoEye – сервис доступа к оперативным данным [Электронный ресурс]. – Режим доступа: <https://meteoeye.gis.by>. – Дата доступа: 19.04.2021.

³Подсистема обработки и распространения оперативной космической информации – программный интерфейс [Электронный ресурс]. – Режим доступа: <https://meteoeye.gis.by/api/index.html>. – Дата доступа: 19.04.2021.

Модуль обработки оперативной космической информации (МО ОКИ, Data processing) предназначен для автоматизации процессов первичной обработки принятой оперативной космической информации с целью формирования информационных ресурсов, соответствующих различным требованиям.

Программные средства модуля обеспечивают:

- контроль поступления данных от ПП ОКИ;
- сбор вспомогательных данных (начальных баллистических условий, калибровочных коэффициентов, прогностической информации и др.) для обработки;
- обработку получаемых данных и формирование базовых информационных ресурсов, реализующих информационную совместимость с составными частями модуля;
- формирование описательных данных, обзорных изображений к базовым информационным ресурсам, доступным пользователю через приложение просмотра и поиска оперативной космической информации;
- формирование информационных ресурсов для работы ПК «Прогноз ЗК» и ПК ОТА.

Информационные ресурсы, получаемые с помощью модуля обработки оперативной космической информации, совместимы с программными пакетами обработки метеорологических данных Direct Readout Laboratory⁴, Community Satellite Processing Package⁵ и AAPP⁶.

Сервисы для поддержки работы системы (Сервисы, Services) обеспечивают движение оперативной космической информации по компонентам системы, управление временем жизни базовых информационных ресурсов, организацию информационных рассылок посредством протокола SMTP и ведение журнала основных операций.

Клиентское приложение администрирования (КПА, OAuth2.WEB) является интерфейсом модуля аутентификации и авторизации. В качестве протокола проверки подлинности пользователей и программных компонентов системы, а также процедуры предоставления прав на выполнение определенных действий и подтверждения данных прав при попытке выполнения этих действий с помощью стороннего программного обеспечения вышеуказанным модулем используется протокол OpenID Connect, в качестве авторизации – протокол OAuth 2.0. Модуль аутентификации и авторизации поддерживает следующие варианты взаимодействия с приложением:

- с участием пользователя (Implicit) на основе логина и сеансового ключа;
- без участия пользователя (Client Credentials) на основе client_id и client_secret.

Клиентское приложение просмотра и поиска оперативной космической информации (КППИ ОКИ, WEB) реализовано как веб-приложение, работает в режиме гостевого доступа для неавторизованного пользователя или авторизованного доступа (рис. 2).

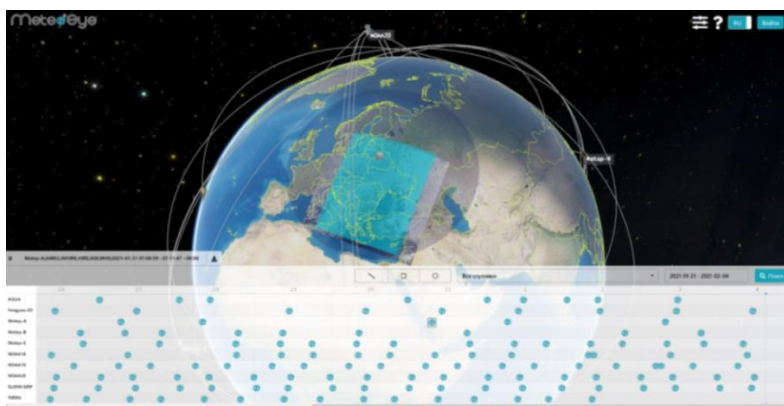


Рис. 2. Главная страница приложения просмотра и поиска оперативной космической информации

Fig. 2. Home page of the application for viewing and searching operational space information

⁴NASA – Direct Readout Laboratory [Electronic resource]. – Mode of access: <https://directreadout.sci.gsfc.nasa.gov/>. – Date of access: 19.04.2021.

⁵CSPP – CIMSS [Electronic resource]. – Mode of access: <https://cimss.ssec.wisc.edu/cspp/>. – Date of access: 19.04.2021.

⁶AAPP | NWP SAF [Electronic resource]. – Mode of access: <https://www.nwpsaf.eu/site/software/aapp/>. – Date of access: 19.04.2021.

Приложение имеет интерфейс на русском и английском языках, позволяет осуществлять поиск и просмотр полученной оперативной космической информации, мониторинг положения метеоспутников на орбите, их полосу обзора и зону радиовидимости антенны БКСДЗ. Для авторизованного пользователя также доступна загрузка базовых информационных ресурсов и просмотр статистики по загруженным ресурсам.

Модуль визуализации оперативной космической информации в процессе приема данных (МВ ОКИ, Visualization.App) реализован как кроссплатформенное приложение с визуальной оболочкой, т. е. может запускаться на любой платформе (кроме мобильных платформ) с графическим интерфейсом пользователя (graphical user interface, GUI). Модуль предназначен для визуализации текущих сеансов приема в режиме реального времени, а также для возможности отображения уже выполненных сеансов приема, доступных на сервере в тот момент (рис. 3).

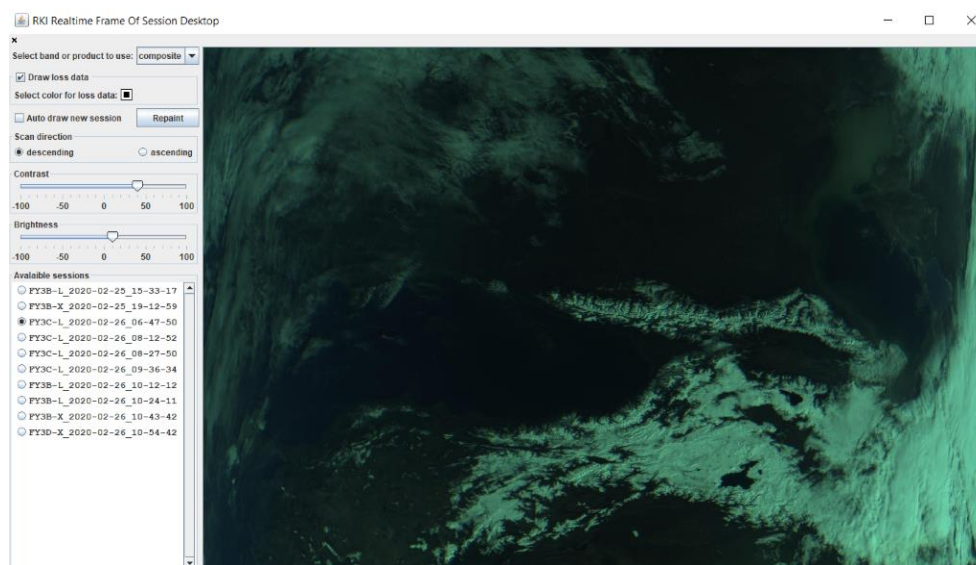


Рис. 3. Главная страница приложения визуализации оперативной космической информации в процессе приема данных

Fig. 3. Home page of the application for visualizing operational space information in the process of data receiving

Клиентское приложение планирования сеансов приема (КППСП, Planning.App) реализовано как веб-приложение и представляет собой интерфейс для управления сеансами приема.

Применение данных ПК «Прогноз ЗК». Признаком наличия растительности и ее состояния является спектральная отражательная способность, которой присущи большие различия при разных длинах волн. Знания о связи характеристик растительности с ее отражательной способностью позволяют использовать космические снимки для идентификации как типов растительности, так и ее состояния [1]. Вегетационный индекс – это показатель, рассчитываемый в результате математических операций с разными спектральными диапазонами (каналами) данных дистанционного зондирования и имеющий отношение к параметрам растительности в рассматриваемом пикселе изображения [2].

В настоящее время существует около 160 вариантов вегетационных индексов. Они подбираются экспериментально исходя из известных особенностей кривых спектральной отражательной способности растительности и почв [3]. Расчет большей части вегетационных индексов базируется на двух наиболее стабильных (не зависящих от прочих факторов) участках кривой спектральной отражательной способности растений. На красную зону спектра (0,62–0,75 мкм) приходится максимум поглощения солнечной радиации хлорофиллом, а на ближнюю инфракрасную зону (0,75–1,3 мкм) – максимум отражения энергии кле-

точной структурой листа, т. е. высокая фотосинтетическая активность, связанная, как правило, с большой фитомассой растительности, ведет к более низким значениям коэффициентов отражения в красной зоне спектра и большим значениям в ближней инфракрасной зоне. Хорошо известно, что отношение этих показателей друг к другу позволяет четко отделять растительность от прочих природных объектов [4]. В качестве вегетационного индекса выбран один из самых известных – нормализованный относительный индекс растительности (Normalized Difference Vegetation Index, NDVI), который вычисляется по формуле [5]

$$NDVI = \frac{NIR - RED}{NIR + RED},$$

где NIR – отражение в ближней инфракрасной области спектра, RED – отражение в красной области спектра.

На основе NDVI можно получить полноценный спектральный анализ и выявить участки, которые нуждаются в пересеве, внесении средств защиты растений или удобрений. Индекс умеренно чувствителен к изменениям почвенного и атмосферного фона, кроме случаев с бедной растительностью, и может перенасыщаться в условиях густой растительности⁷.

Данные ДЗЗ для работы ПК «Прогноз ЗК» предоставляет ПК РОКИ в виде файла с расширением GeoTIFF (Float 32) (рис. 4).

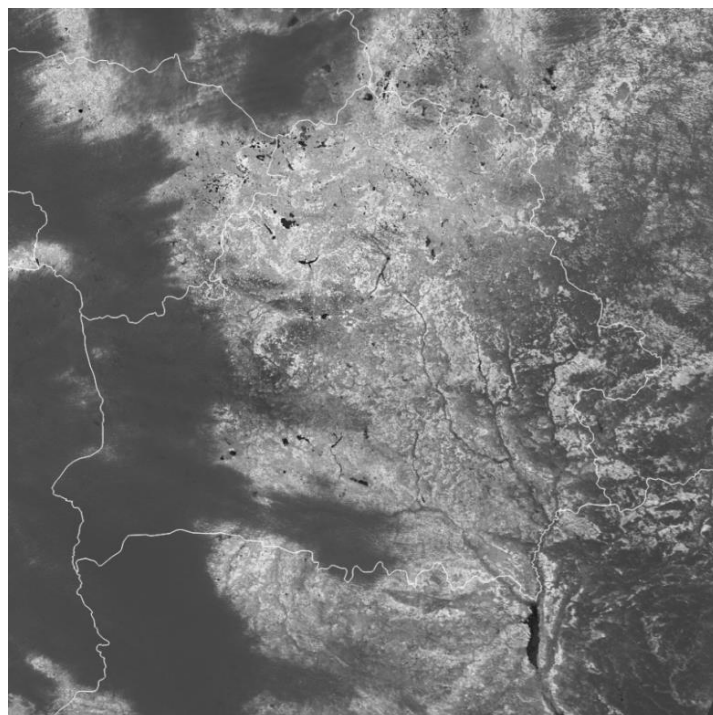


Рис. 4. Производный информационный ресурс NDVI по результатам обработки данных сеанса TERRA за 3 марта 2021 г.

Fig. 4. Derived information resource NDVI based on the results of data processing from the TERRA session on 3 March 2021

Взаимодействие ПК РОКИ и ПК «Прогноз ЗК» происходит следующим образом: ПК РОКИ осуществляет подготовку данных ДЗЗ с КА AQUA, SUOMI NPP, NOAA 20, MetOp, Fengyun-3, т. е. происходит дешифровка снимка и формирование результирующего изображения, содер-

⁷Вегетационные индексы NDVI, EVI, GNDVI, CVI, True Color [Электронный ресурс]. – Режим доступа: <https://www.soft.farm/ru/blog/vegetacionnye-indeksy-ndvi-evi-gndvi-cvi-true-color-140>. – Дата доступа: 19.04.2021.

жащего нормализованный разностный вегетационный индекс NDVI. ПК «Прогноз ЗК» выполняет загрузку результирующего файла NDVI с данными по территории Республики Беларусь и формирование прогнозной информации на основе оперативного анализа вегетационного развития зерновых сельскохозяйственных культур в пределах административных районов Республики Беларусь. На рис. 5 показана средняя оправдываемость такого метода прогнозирования урожайности зерновых культур, рассчитанная в ПК «Прогноз ЗК» за 2015–2019 гг.

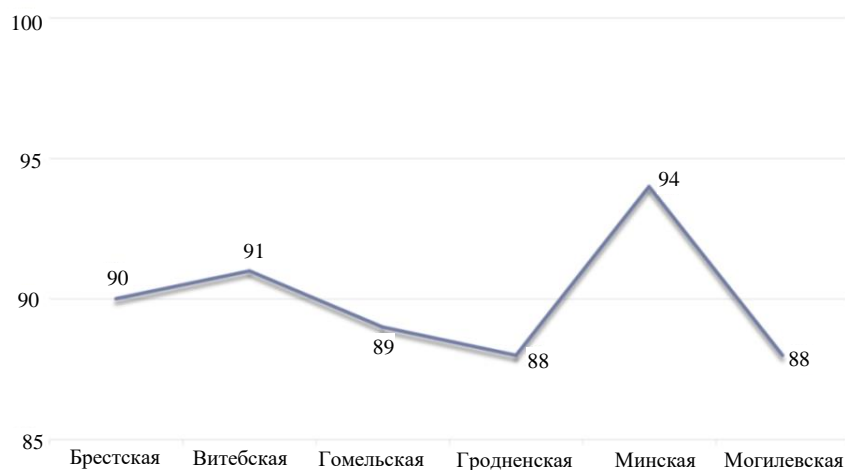


Рис. 5. Средняя оправдываемость прогноза урожайности зерновых культур, %
 Fig. 5. Average justifiability of grain yield forecast, %

Заключение. На современном этапе развития науки, техники и технологий требуется не столько получение и обработка данных ДЗЗ Национальным оператором БКСДЗ в соответствии с заранее заданными потребителями параметрами, сколько создание специализированных клиентских приложений и сервисов, позволяющих заинтересованным ученым и специалистам выполнять тематическую обработку данных ДЗЗ, изменяя определенным образом параметры обработки данных, и самостоятельно анализировать полученную информацию. Для решения этой актуальной задачи в рамках государственной программы «Научеёмкие технологии и техника» на 2016–2020 гг. был создан ПК РОКИ.

ПК РОКИ уже сегодня применяется для решения научных и прикладных задач в различных областях знаний. Он обеспечивает возможность поиска и выборки необходимых пользователям спутниковых данных и предоставляет различные способы, позволяющие проводить удаленный анализ спутниковой информации.

С момента тестового запуска 18 июля 2019 г. и в последующий период эксплуатации ПК РОКИ, который продлится до конца февраля 2021 г., уже принято более 7500 сеансов с КА AQUA, TERRA, SUOMI NPP, NOAA 20, MetOp A/B/C, NOAA 18, NOAA 19 и Feng-Yun 3D. ПК РОКИ является источником данных ДЗЗ для работы ПК «Прогноз ЗК».

При создании распределенной системы приема, обработки и распространения оперативной космической информации с КА изначально был заложен принцип дальнейшего совершенствования и развития, что позволяет с определенной периодичностью проводить улучшения ПК РОКИ в соответствии с новыми требованиями, выдвигаемыми современными исследованиями и практикой применения как отдельных модулей, так и информационной системы в целом.

Вклад авторов. С. А. Золотой – обоснование концепции и актуальности статьи; И. Б. Страшко – консультирование по вопросам создания ПК РОКИ; Д. С. Котов – обзор и аналитический анализ статьи, формулирование выводов; И. М. Нестерович – сбор, анализ и обобщение данных; В. В. Рубо – сбор и систематизация данных; К. И. Костюк – сбор и систематизация данных, интерпретация результатов применения данных ПК «Прогноз ЗК».

Список использованных источников

1. Черепанов, А. С. Вегетационные индексы / А. С. Черепанов // Геоматика. – 2011. – № 2. – С. 98–102.
2. Воронина, П. В. Классификация тематических задач мониторинга сельского хозяйства с использованием данных дистанционного зондирования MODIS / П. В. Воронина, Е. А. Мамаш // Вычислительные технологии. – 2014. – Т. 19, № 3. – С. 76–102.
3. Чаюк, А. И. Оценка динамики состояния растительности с использованием вегетационных индексов [Электронный ресурс] / А. И. Чаюк, Н. В. Клебанович. – Режим доступа: <https://elib.bsu.by/bitstream/123456789/148510/1/edvuvi.pdf>. – Дата доступа: 19.04.2021.
4. Черепанов, А. С. Спектральные свойства растительности и вегетационные индексы / А. С. Черепанов, Е. Г. Дружинина // Геоматика. – 2009. – № 3. – С. 28–32.
5. Боровлев, А. Ю. Вегетационный индекс как инструмент количественной оценки территорий, подвергнутых сплошным рубкам леса [Электронный ресурс] / А. Ю. Боровлев, Н. В. Шуктомов, Ю. А. Паутов. – Режим доступа: http://spbftu.ru/site/upload/201612201741_Borovlev_1.pdf. – Дата доступа: 19.04.2021.

References

1. Cherepanov A. S. *Vegetation indices data*. Geomatika [Geomatics], 2011, no. 2, pp. 98–102 (In Russ.).
2. Voronina P. V., Mamash E. A. *Classification of thematic monitoring for agriculture problems using remote sensing MODIS data*. Vychislitel'nye tehnologii [Computational Technologies], 2014, vol. 19, no. 3, pp. 76–102 (In Russ.).
3. Chayuk A. I., Klebanovich N. V. *Otsenka dinamiki sostoyaniya rastitel'nosti s ispol'zovaniyem vegetatsionnykh indeksov*. Assessment of the dynamics of the state of vegetation using vegetation indices. Available at: <https://elib.bsu.by/bitstream/123456789/148510/1/edvuvi.pdf> (accessed 19.04.2021).
4. Cherepanov A. S., Druzhinina E. G. *Spectral properties of vegetation and vegetation indices data*. Geomatika [Geomatics], 2009, no. 3, pp. 28–32 (In Russ.).
5. Borovlev A. Ju., Shuktomov N. V., Pautov Ju. A. *Vegetacionnyj indeks kak instrument kolichestvennoj ocenki territorij, podvergnutyh sploshnym rubkam lesa*. Vegetation Index as a Tool for Quantitative Assessment of Territories Subjected to Continuous Logging. Available at: http://spbftu.ru/site/upload/201612201741_Borovlev_1.pdf (accessed 19.04.2021).

Информация об авторах

Золотой Сергей Анатольевич, кандидат технических наук, директор, УП «Геоинформационные системы».

E-mail: gis@gis.by

Страшко Игорь Болеславович, заместитель директора по производству, УП «Геоинформационные системы».

E-mail: IStrashko@gis.by

Котов Дмитрий Сергеевич, кандидат технических наук, ученый секретарь, УП «Геоинформационные системы».

E-mail: dkotau@gis.by

Нестерович Ирина Максимовна, младший научный сотрудник отдела инновационных технологий, УП «Геоинформационные системы».

E-mail: nesterovichirina20@gmail.com

Рoubо Виталий Вильгельмович, начальник отдела инновационных технологий, УП «Геоинформационные системы».

E-mail: vroubo@gis.by

Костюк Каролина Ивановна, заведующий сектором лаборатории биоинформатики, Объединенный институт проблем информатики Национальной академии наук Беларуси.

E-mail: kostyukkarolina@gmail.com

Information about the authors

Sergey A. Zolotoy, Cand. Sci. (Eng.), Director, UE "Geoinformation Systems".

E-mail: gis@gis.by

Igor B. Strashko, Deputy Director for Production Affairs, UE "Geoinformation Systems".

E-mail: IStrashko@gis.by

Dzmitry S. Kotau, Cand. Sci. (Eng.), Scientific Secretary, UE "Geoinformation Systems".

E-mail: dkotau@gis.by

Iryna M. Nestsiarovich, Junior Research Scientist of the Department of Innovative Technologies, UE "Geoinformation Systems".

E-mail: nesterovichirina20@gmail.com

Vitali V. Rouba, Head of the Department of Innovative Technologies, UE "Geoinformation Systems".

E-mail: vroubo@gis.by

Karolina I. Kostyuk, Head of the Bioinformatics Laboratory Sector, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.

E-mail: kostyukkarolina@gmail.com

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