

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

ИНФОРМАТИКА

ТОМ 17

4

ОКТЯБРЬ-ДЕКАБРЬ
2020

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.

Основные разделы журнала: математическое моделирование; обработка сигналов, изображений и речи; прикладные информационные технологии; интеллектуальные системы; космические информационные технологии и геоинформатика; параллельные вычисления; системы, приборы и устройства; распознавание образов; информационная безопасность; автоматизация проектирования.

Журнал «Информатика» включен Высшей аттестационной комиссией Республики Беларусь в список научных изданий для опубликования результатов диссертационных исследований. В декабре 2017 г. журнал 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.

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

Адрес редакции:

ул. Сурганова, 6, к. 305, г. Минск, Беларусь
Тел. +375 (017) 351 26 22
E-mail: rio@newman.bas-net.by

Сайт журнала: inf.grid.by

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 academies in the given field. The journal focuses on the most significant and modern papers such as research projects results and PhD/DSc papers in computer sciences, IT and at the boundaries.

The journal covers the following topics: mathematical modeling; processing and recognition of signals, images and speech; applied information technology; intelligent systems; space information technology and GIS technologies; parallel computing; systems, devices and equipment; image recognition; information security; computer aided design.

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 the free access to reviewed electronic scientific paper, improving scientific information traffic and also raising quotation of works of the authors who are published in this journal (please use <https://elibrary.ru> or section for authors https://elibrary.ru_author_tools).

The journal is edited for a wide range of specialists in IT and computer sciences.

For further information:

Phone +375 (017) 351 26 22
E-mail: rio@newman.bas-net.by
Office 305, Surganova 6, 220012
Minsk, Belarus

Learn more at: <https://inf.grid.by/jour>

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

ИНФОРМАТИКА

Том 17, № 4, октябрь-декабрь 2020

Ежеквартальный научный журнал

Издается с января 2004 г.

Учредитель и издатель – Объединенный институт проблем информатики
Национальной академии наук Беларуси

Г л а в н ы й р е д а к т о р

Тузиков Александр Васильевич, д-р физ.-мат. наук, профессор, член-корреспондент
Национальной академии наук Беларуси, генеральный директор Объединенного института
проблем информатики Национальной академии наук 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, профессор, член-корреспондент
Национальной академии наук Беларуси, Объединенный институт проблем информатики
Национальной академии наук 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, профессор, академик Национальной академии наук
Беларуси, Белорусский государственный университет, Минск, Беларусь

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

Бибило Петр Николаевич, д-р техн. наук, профессор, Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь

Бобов Михаил Никитич, д-р техн. наук, профессор, ОАО «АГАТ – системы управления» – управляющая
компания холдинга «Геоинформационные системы управления», Минск, Беларусь

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

Дудин Александр Николаевич, д-р физ.-мат. наук, профессор, Белорусский государственный университет,
Минск, Беларусь

Карпов Алексей Анатольевич, д-р техн. наук, доцент, Санкт-Петербургский институт информатики
и автоматизации Российской академии наук, Санкт-Петербург, Россия

Килин Сергей Яковлевич, д-р физ.-мат. наук, профессор, академик Национальной академии наук
Беларуси, Президиум Национальной академии наук Беларуси, Минск, Беларусь

Краснопрошин Виктор Владимирович, д-р техн. наук, профессор, Белорусский государственный
университет, Минск, Беларусь

Крот Александр Михайлович, д-р техн. наук, профессор, Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь

Кругликов Сергей Владимирович, д-р воен. наук, канд. техн. наук, доцент, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь

Кундас Семен Петрович, д-р техн. наук, профессор, Белорусский национальный технический университет, Минск, Беларусь

Лиходед Николай Александрович, д-р физ.-мат. наук, профессор, Белорусский государственный университет, Минск, Беларусь

Матус Петр Павлович, д-р физ.-мат. наук, профессор, Институт математики Национальной академии наук Беларуси, Минск, Беларусь

Скляр Валерий Анатольевич, д-р техн. наук, профессор, Университет Авейру, Португалия

Сотсков Юрий Назарович, д-р физ.-мат. наук, профессор, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь

Стемпковский Александр Леонидович, д-р техн. наук, профессор, академик Российской академии наук, Институт проблем проектирования в микроэлектронике Российской академии наук, Москва, Россия

Харин Юрий Семенович, д-р физ.-мат. наук, профессор, член-корреспондент Национальной академии наук Беларуси, Научно-исследовательский институт прикладных проблем математики и информатики Белорусского государственного университета, Минск, Беларусь

Чернявский Александр Федорович, д-р техн. наук, профессор, академик Национальной академии наук Беларуси, Институт прикладных физических проблем им. А. Н. Севченко Белорусского государственного университета, Минск, Беларусь

Ярмолик Вячеслав Николаевич, д-р техн. наук, профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

ИНФОРМАТИКА

Том 17, № 4, октябрь-декабрь 2020

Ответственный за выпуск *Мойсейчик Светлана Сергеевна*

Редактор *Гончаренко Галина Борисовна*

Корректор *Михайлова Анна Антоновна*

Компьютерная верстка *Бутевич Ольга Борисовна*

Сдано в набор 17.11.2020. Подписано в печать 16.12.2020. Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс. Ризография. Усл. печ. л. 12,8. Уч.-изд. л. 12,5. Тираж 50 экз. Заказ 8.

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. 17, no. 4, October-December 2020

Published quarterly

Issued since January 2004

Founder and publisher – the United Institute of Informatics Problems
of the National Academy of Sciences of Belarus

Editor-in-Chief

Alexander V. Tuzikov, Dr. Sci. (Phys.-Math.), Professor, Corresponding Member of the National Academy of Sciences of Belarus, General Director of the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

Deputy Editor-in-Chief

Mikhail Y. Kovalyov, Dr. Sci. (Phys.-Math.), Professor, Corresponding Member of the National Academy of Sciences of Belarus, the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

Editorial Board

Sergey V. Ablameyko, Dr. Sci. (Eng.), Professor, Academician of the National Academy of Sciences of Belarus, Belarusian State University, Minsk, Belarus

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

Petr N. Bibilo, Dr. Sci. (Eng.), Professor, the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

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

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

Alexander N. Dudin, Dr. Sci. (Phys.-Math.), Professor, Belarusian State University, Minsk, Belarus

Alexey A. Karpov, Dr. Sci. (Eng.), Associate Professor, St. Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences, Saint Petersburg, Russia

Sergey Ya. Kilin, Dr. Sci. (Phys.-Math.), Professor, Academician of the National Academy of Sciences of Belarus, Presidium of the National Academy of Sciences of Belarus, Minsk, Belarus

Viktor V. Krasnoproshin, Dr. Sci. (Eng.), Professor, Belarusian State University, Minsk, Belarus

Alexander M. Krot, Dr. Sci. (Eng.), Professor, the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

Sergey V. Kruglikov, Dr. Sci. (Milit.), Cand. Sci. (Eng.), Associate Professor, the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

Semen P. Kundas, Dr. Sci. (Eng.), Professor, Belarusian National Technical University, Minsk, Belarus

Nikolai A. Likhoded, Dr. Sci. (Phys.-Math.), Professor, Belarusian State University, Minsk, Belarus

Petr P. Matus, Dr. Sci. (Phys.-Math.), Professor, Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus

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

Yuri N. Sotskov, Dr. Sci. (Phys.-Math.), Professor, the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus

Alexander L. Stempkovsky, Dr. Sci. (Eng.), Professor, Academician of the Russian Federation Academy of Sciences, the Institute for Design Problems in Microelectronics of the Russian Federation Academy of Sciences, Moscow, Russia

Yuriy S. Kharin, Dr. Sci. (Phys.-Math.), Professor, Corresponding Member of the National Academy of Sciences of Belarus, Research Institute for Applied Problems of Mathematics and Informatics of the Belarusian State University, Minsk, Belarus

Alexander F. Cherniavsky, Dr. Sci. (Eng.), Professor, Academician of the National Academy of Sciences of Belarus, A. N. Sevchenko Institute of Applied Physical Problems of the Belarusian State University, Minsk, Belarus

Vyacheslav N. Yarmolik, Dr. Sci. (Eng.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

INFORMATICS

Vol. 17, no. 4, October-December 2020

Issue Head *Sviatlana S. Maiseichyk*

Editor *Halina B. Hancharenka*

Corrector *Hanna A. Mikhailava*

Computer Imposition *Volha B. Butsevich*

Sent for press 17.11.2020. Output 16.12.2020. Format 60×84 1/8. Offset paper. Headset Times. Riesography. Printed sheets 12,8. Publisher's signatures 12,5. Circulation 50 copies. Order 8.

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

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Рубанов И. В., Ковалев М. Я.** Вычислительные методы решения задачи комбинирования секторов воздушного пространства 7
- Романчак В. М.** Вейвлет-преобразование на конечном интервале 22
- Дрыбин Е. А., Садов С. В., Садов В. С.** Цифровая модель генератора псевдослучайных чисел на основе непрерывной хаотической системы 36

ОБРАБОТКА ИЗОБРАЖЕНИЙ, РЕЧИ И ТЕКСТА

- Ковалев В. А., Войнов Д. М., Малышев В. Д., Лапо Е. Д.** Компьютеризированная диагностика рака простаты на основе полнослайдовых гистологических изображений и методов глубокого обучения 48
- Качков Д. И.** Моделирование языка и двунаправленные представления кодировщиков: обзор ключевых технологий 61
- Липницкий С. Ф.** Интернет-поиск и лексико-семантическая обработка аналогов принятых решений в различных предметных областях 73

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

- Костюк К. И., Браницкий А. В., Роубо В. В., Нестерович И. М.** Международный стандарт HL7 FHIR как основа создания единого информационного пространства здравоохранения Республики Беларусь 83

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

- Nataraj S. K., Paulraj M. P., Yaacob S. Bin, Adom A. H.** Intelligent robot chair with communication aid using TEP responses and higher order spectra band features 92

УЧЕНЫЕ БЕЛАРУСИ

- Крот Александр Михайлович** (к 60-летию со дня рождения) 104
- Дудкин Александр Арсентьевич** (к 70-летию со дня рождения) 107

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

CONTENTS

MATHEMATICAL MODELING

- Rubanov I. V., Kovalyov M. Y.** Computational methods for airspace sectorisation..... 7
- Romanchak V. M.** Wavelet transformation on a finite interval 22
- Drybin Y. A., Sadau S. V., Sadau V. S.** Digital model of a pseudo-random number generator based on a continuous chaotic system..... 36

IMAGE, SPEECH AND TEXT PROCESSING

- Kovalev V. A., Voynov D. M., Malyshau V. D., Lapo E. D.** Computerized diagnosis of prostate cancer based on whole slide histology images and deep learning methods..... 48
- Kachkou D. I.** Language modeling and bidirectional coders representations: an overview of key technologies..... 61
- Lipnitsky S. F.** Internet search and lexical-semantic processing of analogs when making decisions in various subject areas..... 73

INFORMATION TECHNOLOGY

- Kostyuk K. I., Branitski A. V., Roubo V. V., Nesterovich I. M.** International standard HL7 FHIR as a working framework for development of common healthcare information space of the Republic of Belarus..... 83

INTELLIGENT SYSTEMS

- Nataraj S. K., Paulraj M. P., Yaacob S. Bin, Adom A. H.** Intelligent robot chair with communication aid using TEP responses and higher order spectra band features..... 92

SCIENTISTS OF BELARUS

- Krot Alexander Mikhailovich** (on his 60th birthday) 104
- Dudkin Alexander Arsentievich** (on his 70th birthday) 107

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

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

УДК 519.1 + 656.021
<https://doi.org/10.37661/1816-0301-2020-17-4-7-21>

Поступила в редакцию 09.07.2019
Received 09.07.2019

Принята к публикации 20.10.2020
Accepted 20.10.2020

Вычислительные методы решения задачи комбинирования секторов воздушного пространства

И. В. Рубанов^{1✉}, М. Я. Ковалев²

¹Белорусская государственная академия авиации, Минск, Беларусь
✉E-mail: irubanov@inbox.ru

²Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь

Аннотация. Рассматривается задача комбинирования элементарных секторов региона воздушного пространства, в которой должно быть получено минимальное количество комбинированных секторов при ограничениях на их нагрузку и допустимость комбинаций, таких как требования связности в пространстве или принадлежности заданному множеству разрешенных комбинаций. Предлагаются и тестируются вычислительные методы, которые могут быть применены для решения этих и более общих задач секторизации регионов воздушного пространства. В частности, предлагаются комбинаторные алгоритмы двух типов для построения разбиений конечного множества с заданными весами элементов и ретико-графовыми связями между ними. Разбиения строятся методом ветвей и границ с минимизацией количества подмножеств в итоговом разбиении при ограничении на суммарный вес элементов в подмножестве. В алгоритме первого типа в каждом узле дерева ветвей и границ формируются готовые компоненты итогового разбиения, оставшаяся часть исходного множества подвергается дальнейшему разбиению в нижестоящих узлах. В алгоритме второго типа в каждом узле формируется целиком текущее разбиение, компоненты которого дополняются в нижестоящих узлах. При сравнении показателей производительности алгоритмов рассматриваемые задачи делятся на две группы, в одной из которых на решение накладывается требование связности графовых структур, сопоставленных подмножествам итогового разбиения, а в другой такое требование отсутствует. Также предлагаются постановки задачи в терминах задачи целочисленного программирования.

Устанавливается вычислительная сложность двух вариантов задачи: типа задачи об упаковке в контейнеры, учитывающей ограничения на допустимость комбинаций, и типа задачи о покрытии.

Ключевые слова: секторизация воздушного пространства, безопасность полетов, разбиение конечного множества, целочисленное линейное программирование, метод ветвей и границ

Для цитирования. Рубанов, И. В. Вычислительные методы решения задачи комбинирования секторов воздушного пространства / И. В. Рубанов, М. Я. Ковалев // Информатика. – 2020. – Т. 17, № 4. – С. 7–21. <https://doi.org/10.37661/1816-0301-2020-17-4-7-21>

Computational methods for airspace sectorisation

Igor V. Rubanov^{1✉}, Mikhail Y. Kovalyov²

¹*Belarusian State Academy of Aviation, Minsk, Belarus*

✉E-mail: irubanov@inbox.ru

²*The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus*

Abstract. A problem of combining elementary sectors of an airspace region is considered, in which a minimum number of combined sectors must be obtained with restrictions on their load and feasibility of combinations such as the requirement of the space connectivity or the membership of a given set of permissible combinations. Computational methods are proposed and tested to be used for solution of general problems of airspace sectorization. In particular, two types of combinatorial algorithms are proposed for constructing partitions of a finite set with specified element weights and graph-theoretical relationships between the elements. Partitions are constructed by use of a branch and bound method to minimize the number of subsets in the final partition, while limiting the total weight of elements in the subset. In the first type algorithm, ready-made components of the final partition are formed in each node of the branch and bound tree. The remaining part of the original set is further divided at the lower nodes. In the second type algorithm, the entire current partition is formed in each node, the components of which are supplemented at the lower nodes. When comparing algorithms performance, the problems are divided into two groups, one of which contains a connectivity requirement, and the other does not. Several integer programming formulations are also presented.

Computational complexity of two problem variants is established: a bin packing type problem with restrictions on feasible combinations, and covering type problem.

Keywords: airspace sectorization, flight safety, finite set partitioning, integer linear programming, branch and bound method

For citation. Rubanov I. V., Kovalyov M. Y. Computational methods for airspace sectorisation. *Informatics*, 2020, vol. 17, no. 4, pp. 7–21 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-7-21>

Введение. Одной из актуальных задач организации воздушного движения является повышение пропускной способности воздушного пространства (ВП) заданного региона. Среди факторов, оказывающих влияние на пропускную способность, большое значение имеет возможность контроля событий, происходящих в областях (секторах) ВП региона. В частности, оператор сектора ВП может отслеживать и поддерживать связь лишь с ограниченным количеством воздушных судов (ВС). Для поддержки контроля ВП в условиях изменяющейся ситуации и при ограниченных ресурсах может производиться динамическая секторизация ВП.

В самом недавнем из найденных по теме настоящей работы источнике [1], представляющем собой обзор предшествующих публикаций, можно найти ссылки на наиболее цитируемые, по мнению авторов статьи, работы. В обзоре представлена классификация исследований по нескольким критериям. Одним из них является подход к формированию секторов ВП, заключающийся либо в кратковременном комбинировании сравнительно небольших областей ВП правильной формы (назовем такой принцип моделью гибкой секторизации), либо в комбинировании уже существующих секторов с определенными для них ресурсами управления (назовем этот принцип моделью объединения или комбинирования секторов) [2].

В статьях [2, 3] отмечено, что второй принцип на сегодняшний день выглядит более реалистичным. В отличие от первого он является достаточно легко реализуемым, предлагая баланс между экономической эффективностью и безопасностью. Кратковременное комбинирование секторов, существующих в достаточно долговременно сформированных границах, уже практикуется на основе экспертных оценок нагрузки и ресурсов управления [4].

В обзоре [1] проводится различие между понятиями «секторизация» и «конфигурация» ВП. Конфигурация «обеспечивает расписание для объединения и разделения элементарных секторов на секторы управления, которые подходят для заданного числа доступных контроллеров и ожидаемой структуры трафика» [1, с. 1]. Конфигурация является более комплексным подходом, принимающим во внимание большее количество таких факторов, как, например, определение временных интервалов, в течение которых осуществляется решение

задачи секторизации; прогнозирование с различным упреждением значений нагрузки; учет маршрутов движения ВС и т. д. Более общая задача конфигурации может включать ответы на важные вопросы определения метрик нагрузки и эффективности, при котором применение математических моделей секторизации имеет достаточный смысл с точки зрения повышения пропускной способности ВП либо уменьшения стоимости ресурсов.

В обзор [1] не включены описания работ по проблемам конфигурации, однако на них даются соответствующие ссылки. Оговаривается, что множества вопросов, затрагиваемых при решении задач секторизации и конфигурации, взаимопересекаются. С исследованиями, касающимися, например, вопросов определения метрик и тестирования эффективности моделей на их основе, можно ознакомиться в статьях [2–9]. Обзор [1] не включает источники русскоязычных авторов, среди которых можно упомянуть серию комплексных работ [5–7].

Задача конфигурации остается за пределами настоящего исследования. Предполагается, что на вход задачи подаются числовые данные, возможно, прошедшие какую-либо предобработку по более общим моделям. В работе предлагаются математические методы решения элементарной задачи комбинирования областей ВП (КОВП) при заданных в числовом виде ограничениях на нагрузку в итоговых секторах (будем называть их объединениями), максимизирующие эффективность, которая задается как их минимальное итоговое количество. Предполагается, что сформированная комбинация секторов неизменна в течение горизонта планирования.

В статье предлагается решение задачи при помощи комбинаторного алгоритма разбиения конечного множества методом ветвей и границ, подобного предлагаемому в работе [4], однако использующего несколько иной принцип разбиения. Дается пошаговое описание некоторых алгоритмов и блок-схема для одного из них. Приводятся результаты вычислительных тестов по сравнению производительности алгоритмов, основанных на обоих принципах разбиения, при выбранном критерии эффективности. Предлагается также формулировка задачи в нескольких вариантах как задачи целочисленного линейного программирования (ЦЛП).

Производительность работы алгоритмов и решателей можно ожидать достаточной для ВП относительно небольших регионов, включающих до 20 исходных областей.

Постановка задачи. Сформулируем элементарную задачу комбинирования секторов. Пусть имеется ВП, разделенное на несколько секторов. В каждом из них находится изменяющееся во времени количество ВС, которое будем называть нагрузкой сектора. Секторы могут быть разделены или объединены с соблюдением условия неперевышения предела нагрузки, заданного для каждого из них. Предположим, что каждому сектору сопоставлено некоторое количество требуемых ресурсов для обслуживания и итоговая секторизация должна быть произведена с минимально возможным общим количеством требуемых ресурсов.

Математическая постановка задачи может быть следующей:

1. Пусть требуемый регион ВП разделен на области, которые назовем, как в статье [1], *элементарными секторами*. К примеру, элементарными секторами могут быть части, образованные делением криволинейного в общем случае цилиндра, определяющего ВП, плоскостями или полуплоскостями, параллельными образующей цилиндра, и поверхностями, параллельными его основаниям (рис. 1). Обозначим множество элементарных секторов как S , а его элементы – как $s_i, i = \overline{1, n}$.

2. Будем рассматривать интервалы времени t , в течение которых нагрузку w_i в каждом из элементарных секторов и в их объединениях будем считать постоянной.

3. По каждому интервалу t решаем задачу такого разбиения множества S элементарных секторов s_i на подмножества S_j , при котором в каждом S_j нагрузка w_j не превысит допустимого значения k_j и при этом количество m подмножеств S_j будет минимально. Если для какого-либо s_i оказывается, что $w_i > k_i$, предварительно обновляем S , исключая этот элементарный сектор из обработки, и затем добавляем его к итоговому разбиению. Такой показатель эффективности постановки предполагает, что количество требуемых ресурсов одинаково для всех S_j .

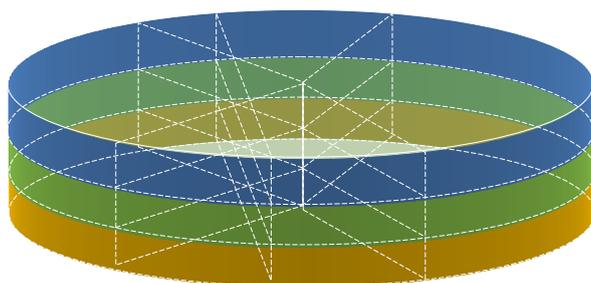


Рис. 1. Деление ВП на элементарные секторы

4. Имеются два варианта дополнительного условия допустимости объединения секторов:

- могут объединяться смежные элементарные секторы;
- задается конечное множество P допустимых объединений, включающее одноэлементные объединения.

5. Допустимая нагрузка k_j может определяться в следующих вариантах:

- как $\max_i (k_i)_j$ в вариантах 4 дополнительного условия допустимости объединения;
- задаться явно для каждого S_j в варианте 4 б);
- быть постоянной k для всех s_i, S_j .

6. Фактическая нагрузка w_j в итоговом объединении определяется как сумма фактических нагрузок w_i его элементов.

Задача в формулировке 1–4 не отвечает на вопросы определения временных интервалов t , возможного усреднения параметров нагрузки и т. д. Эти вопросы авторы относят к комплексной задаче конфигурации, которая, как уже говорилось во введении, находится за пределами настоящего исследования.

Вычислительная сложность модели. Предлагаемая модель с критерием допустимой нагрузки 5 в) постановки задачи без учета дополнительных условий из п. 4 может быть представлена как известная NP-трудная задача об упаковке в контейнеры [10]. Если принять дополнительное условие допустимости 4 а), то модель можно представить как обобщение задачи об упаковке в контейнеры. Условие смежности секторов можно задать в виде графа, тогда обычная задача о контейнерах будет задана на полном графе и из ее NP-трудности следует NP-трудность задачи, заданной на любом графе.

Далее, задачу с критерием допустимой нагрузки 5 а) можно рассмотреть как обобщение предыдущего варианта. Тогда из NP-трудности задачи, в которой допустимая нагрузка определяется по формуле из п. 5 а) с одинаковыми k_i , следует NP-трудность задачи с разными k_i .

Если принять дополнительное условие допустимости 4 б) и если множество P задано изначально, то можно вообще не задавать критерии допустимой нагрузки, а подмножества с превышением сразу отбрасывать, проводя предобработку P . Таким образом, модель с данными условиями может быть представлена как задача о точном покрытии минимальным количеством подмножеств. Если количество возможных связных объединений в S не слишком велико, то дополнительное условие допустимости 4 а) может быть заменено на 4 б). Вариант задачи с условием 4 б) встречается в приложениях в предметной области; в частности, именно он рассматривается в работе [4]. Доказательство NP-трудности варианта 4 б) задачи без критерия допустимой нагрузки будет дано ниже.

Введем следующие обозначения для сформулированных вариантов рассматриваемой задачи: КОВП₁ – задача с критерием допустимой нагрузки 5 а) или 5 в) и дополнительным условием 4 а) или без него, КОВП₂ – задача без критерия допустимой нагрузки и с условием 4 б).

Из комбинаторного анализа известно, что количество вариантов полного перебора всех разбиений множества из n элементов определяется известным числом Белла [11]:

$$B_n = \sum_{q=0}^n S(n, q). \quad (1)$$

Здесь $S(n, q)$ – числа Стирлинга второго рода,

$$S(n, m) = \frac{1}{m!} \sum_{q=0}^m (-1)^{m+q} C_m^q q^n, \quad (2)$$

где C_m^q – биномиальный коэффициент.

Один из комбинаторных алгоритмов, описанных в следующем разделе, основан на рекуррентной формуле для числа Белла

$$B_{n+1} = \sum_{q=0}^n C_n^q B_q, \quad B_0 = 1. \quad (3)$$

В выбранном вычислительном инструменте алгоритм выполняется за время не более нескольких минут для задач КОВП₁ с $n \leq 24$ в варианте 5 в) определения допустимой нагрузки.

Также сформулируем и докажем утверждение об NP-трудности задачи КОВП₂.

Теорема. *Задача КОВП₂ является NP-трудной в сильном смысле.*

Доказательство. Покажем, что КОВП₂ NP-трудна в сильном смысле путем псевдополиномиального сведения к ней NP-полной в сильном смысле задачи «Точное покрытие 3-множествами (ТП-3)» [12].

Задача ТП-3: для заданного семейства $C = \{C_1, \dots, C_r\}$ трехэлементных подмножеств множества $H = \{1, \dots, 3h\}$ следует определить, содержит ли C точное покрытие H , т. е. подсемейство X из C , такое, что каждый элемент $j \in H$ принадлежит точно одному трехэлементному множеству в X .

Для любого примера ТП-3 построим следующий пример КОВП₂: $S = H$, $P = P_3 \cup P_1$, где $P_1 = H$, $P_3 = C$ и требуется определить, существует ли точное покрытие S не более чем h подмножествами из P . Очевидно, что приведенное сведение является псевдополиномиальным. Покажем, что решение примера ТП-3 существует тогда и только тогда, когда существует решение соответствующего примера КОВП₂. Предположим, что подсемейство X из C является решением примера ТП-3. Поскольку $S = H$, $P_3 = C$ и $|X| = h$, множество X является решением соответствующего примера КОВП₂. Теперь предположим, что множество Y подмножеств из P является решением примера КОВП₂. Отметим, что если в Y есть хотя бы одно одноэлементное множество из P_1 , то множество S из $3h$ элементов невозможно точно покрыть h и менее подмножествами, каждое из которых состоит из одного либо трех элементов. Поэтому Y является подсемейством семейства $C = P_3$. Кроме того, если $|Y| \leq h - 1$, то общее количество элементов в Y не превосходит $3(h - 1)$ и оно не может покрывать S . Значит, $|Y| = h$ и Y являются решением примера ТП-3, что завершает доказательство.

Комбинаторные алгоритмы. Рассмотрим два типа комбинаторных алгоритмов: основанные на разделении (будем далее обозначать их алгоритмами типа А) и основанные на добавлении (будем обозначать их алгоритмами типа В). Применим алгоритмы к задаче КОВП₁ с критерием 5 в). Алгоритмы основаны на работе рекурсивной функции, которую далее будем обозначать как $f(arg)$.

Комбинаторные алгоритмы, основанные на разделении. В алгоритмах этого типа в данном узле на каком-либо уровне рекурсии получаем очередной класс J разбиения множества I номеров i элементов s_i исходного множества S , дополняя какое-либо подмножество arg_{pq} множества $arg \setminus i_0$ до самого аргумента arg функции $f(arg)$ для данного узла. Таким

образом, $J = arg \setminus arg_{pq}, arg_{pq} \in arg \setminus i_0$, где $p = |arg_{pq}|$, q – условный номер подмножества среди тех, мощность которых равна p , а i_0 – какой-либо фиксированный номер элемента в аргументе arg (рис. 2). Далее добавляем класс J к формируемому подобным образом по ветке дерева разбиению \underline{J} и, если J является допустимым, а ветка не приводит к количеству классов в \underline{J} , превышающему рекорд, вызываем $f(arg_{pq})$ и переходим на следующий уровень. Здесь и далее подчеркивание символа означает множество обозначаемых этим символом элементов. На рис. 2 показан шаг ветвления рекурсии для алгоритмов типа А при $i_0 = n$ и некоторые итоговые разбиения, получаемые на концах ветвей.

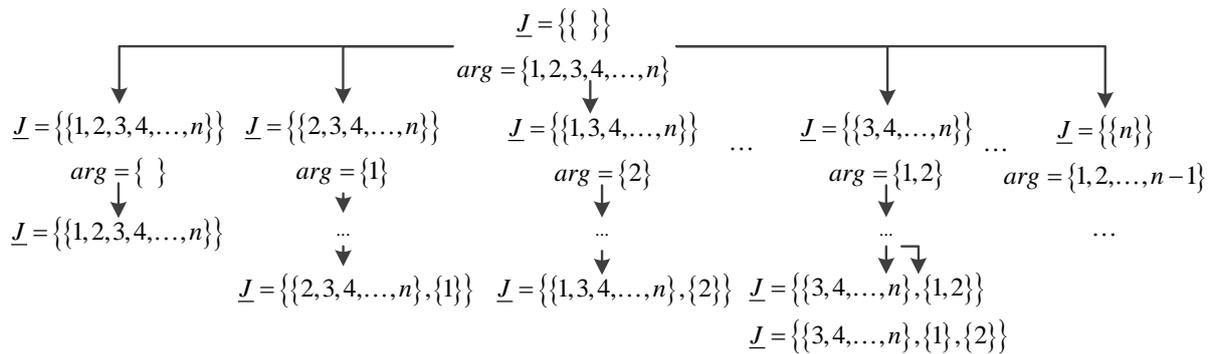


Рис. 2. Ветвление рекурсии алгоритмов типа А

Опишем базовый алгоритм, обозначим его А1.

Создадим пустое множество \underline{J} пустых классов J номеров i элементов s_i множества S , $i \in I, I = \{1, \dots, n\}$. Определим также пустое множество $\underline{\underline{J}}$ получаемых разбиений. Объявим рекорд $r = n$.

Определим рекурсивную функцию $f(arg)$ следующим образом:

а) проверим условие

$$|\underline{J}| < r - 1. \quad (4)$$

Если условие (4) выполняется, переходим к п. б), в противном случае – к п. е);

б) зафиксируем номер $i_0 \in arg$ какого-либо s_i и будем перечислять все подмножества arg_{pq} множества $arg \setminus \{i_0\}$ мощностью $p = 0, |arg \setminus \{i_0\}| - 1$. Как только берется очередное arg_{pq} , определяется множество $J' = arg \setminus arg_{pq}$ и проверяется условие

$$\sum_{i \in J'} w_i \leq k. \quad (5)$$

Если условие (5) выполняется, переходим к п. в); если нет, – к п. е). Закончив перечисление, перейдем к п. е);

в) создадим класс $J = J'$, добавим его к \underline{J} и перейдем к п. г);

г) если $arg_{pq} = \emptyset$, перейдем к п. д); если нет, вызовем $f(arg_{pq})$;

д) добавим \underline{J} к $\underline{\underline{J}}$, объявим новый рекорд $r = |\underline{J}|$, удалим последний элемент из \underline{J} и перейдем к п. е);

е) удалим последний элемент из \underline{J} и выйдем из функции $f(arg)$.

Далее, если $\sum_{i \in I} w_i = 0$, строим единственное разбиение, состоящее из одного класса и включающее все индексы. В противном случае для запуска алгоритма построения разбиений вызываем $f(I)$.

Искомое оптимальное разбиение определяется наборами индексов в одном из $\underline{J} \in \underline{J}$, $|\underline{J}| = r$. Если требуется найти все такие разбиения, в условии (4) следует поставить знак нестрогого неравенства, тогда в \underline{J} может оказаться несколько оптимальных разбиений.

Производительность алгоритма A1 можно улучшить, если учесть минимальное количество классов в допустимом разбиении очередного arg . В самом деле, если учитывать, что каждое arg может быть разбито не менее чем на

$$m_{arg} = m_0 = \left\lceil \sum_{i \in arg} w_i / k \right\rceil \quad (6)$$

частей, то при перечислении подмножеств в $arg \setminus i_0$ можно брать только те из них, в которых количество p элементов не меньше чем $m_{arg} - 1$. Существуют также улучшенные оценки [10], однако далее для простоты будем пользоваться оценкой (6). Условие (4) может быть определено неравенством

$$|\underline{J}| < r - m_{arg}, \quad (7)$$

т. е. ветви, приводящие к разбиениям, мощность которых больше рекорда, можно отсекал раньше, чем при выявлении их по исходной формуле (4).

Опишем улучшенный алгоритм, обозначим его A2.

Рекурсивную функцию определим следующим образом:

а) проверим условие (7). Если оно выполняется, переходим к п. б), в противном случае – к п. д);

б) зафиксируем номер $i_0 \in arg$ какого-либо s_i и будем перечислять все подмножества arg_{pq} множества $arg \setminus \{i_0\}$ мощностью $p = \overline{m_{arg} - 1, |arg \setminus \{i_0\}| - 1}$. Как только берется очередное arg_{pq} , определяется множество $J' = arg \setminus arg_{pq}$. Если $|J'| > 1$, проверяется условие (5). Если оно выполняется, переходим к п. в); если нет, – к п. д). Закончив перечисление, переходим к п. д);

в) создадим класс $J = J'$ и добавим его к \underline{J} . Если $arg_{pq} = \emptyset$, переходим к п. г); если нет, вызываем $f(arg_{pq})$;

г) объявим новый рекорд $r = |\underline{J}| + 1$, создадим класс $J = arg$ и добавим его к классу \underline{J} , который, в свою очередь, добавим к \underline{J} . Удалим последний элемент из \underline{J} и перейдем к п. д);

д) удалим последний элемент из \underline{J} и выйдем из функции $f(arg)$.

Если требуется проверить условие 4 а) постановки задачи, в описанных вариантах это легко сделать, добавив проверку множества S'_j , определяемого множеством J' , на соответствие условию связности к проверке условия (5) в п. г) обоих алгоритмов.

Например, если требуется, чтобы в каждом объединении S_j каждый элементарный сектор s_i был смежен с каким-либо другим по граням или, возможно, ребрам или углам, исходному множеству S можно сопоставить граф G , каждой вершине v_i которого сопоставлен сектор s_i , а ребра соответствуют смежности между s_i по нужному критерию. На рис. 3 показан граф смежности по граням элементарных секторов, который соответствует структуре, изображенной на рис. 1.

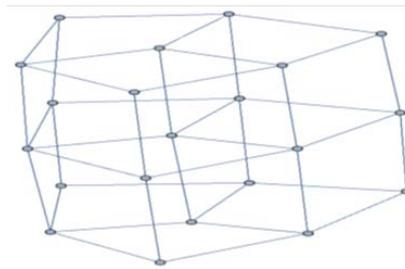


Рис. 3. Граф смежности элементарных секторов

Далее в качестве условия связности объединений можно взять связность подграфа $G_j \in G$, соответствующего объединению S_j .

На рис. 4 изображена структурная схема алгоритма A2 с проверками условия связности.

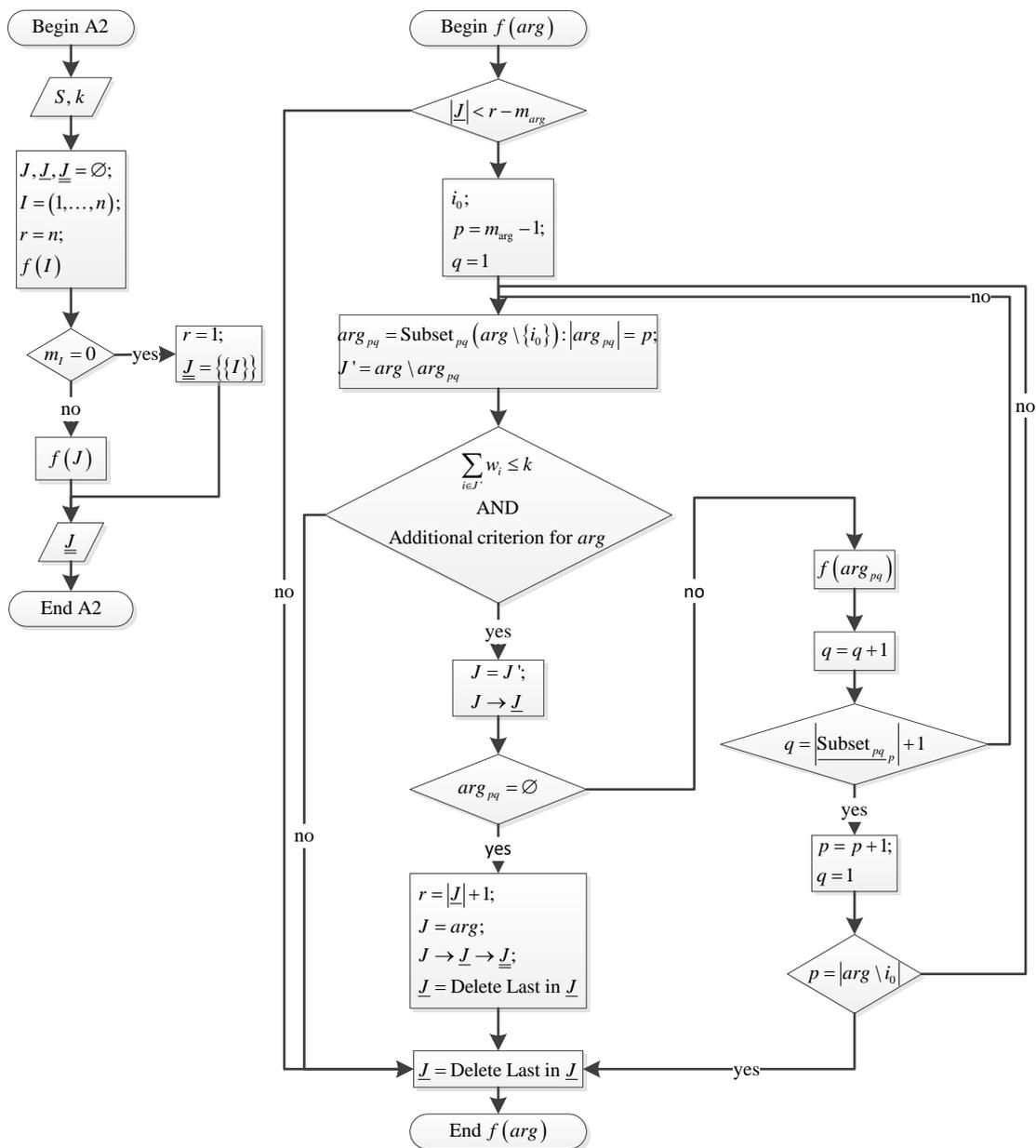


Рис. 4. Алгоритм A2

Комбинаторные алгоритмы, основанные на добавлении. Можно строить разбиения, начиная с множества $\underline{J} = \{\{1\}\}$ и получая каждый узел рекурсии на уровне i при помощи добавления элемента i к каждому из подмножеств верхнего узла и элемента $\{i\}$ ко всему множеству верхнего узла (рис. 5). При подобном принципе разбиения каждый узел представляет одно из разбиений множества элементов с номерами $1, \dots, i$ и на уровне i рекурсии можно получить все возможные такие разбиения. Итоговый набор будет получен на нижнем уровне. Если в каком-либо разбиении количество классов превышает рекорд, по этой ветке далее не идем.

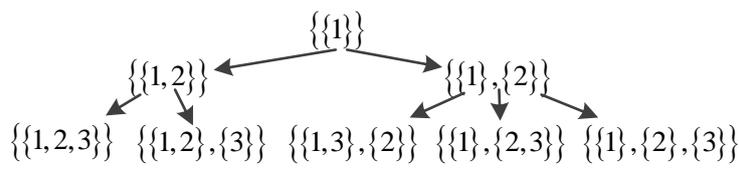


Рис. 5. Схема построения разбиения при помощи добавления элементов

Алгоритм с принципом построения разбиения, основанным на добавлении, предложен в работе [4] для задачи, сходной с обозначенной в настоящей работе как КОВП₂. При выбранных входных данных алгоритмы, построенные по схеме на рис. 5, показывают лучшее время выполнения для построения всего множества разбиений и для некоторых примеров задачи без условия связности. Однако, как будет предположено далее и как показывают результаты тестов, их производительность уступает алгоритму А2 с условием 4 а). Следует отметить, что улучшенная производительность А2 основана на использовании простого показателя эффективности, задаваемого как общее количество результирующих объединений. В работе [4] предлагаются, как можно понять, более комплексные критерии эффективности.

Опишем один из вариантов базового алгоритма, обозначим его В1.

Создадим пустое множество \underline{J} и множество $\underline{J} = \{\{1\}\}$. Объявим рекорд $r = n$, счетчик $i = 1$.

Определим рекурсивную функцию $f(arg)$ следующим образом:

а) проверим условие

$$|arg| < r. \quad (8)$$

Если оно выполняется, перейдем к п. б); если нет, – к п. д);

б) проверим условие

$$i = n. \quad (9)$$

Если оно выполняется, добавим arg к \underline{J} и объявим рекорд $r = |arg|$. В противном случае перейдем к п. в);

в) перечислим подмножества J аргумента \underline{J} , в каждой итерации увеличим счетчик на единицу, проверяя условие (5) для $i \in J \cup \{i\}$. Если оно выполняется, определим $J = J \cup \{i\}$ и вызовем $f(\underline{J})$; если нет, уменьшим счетчик на единицу и перейдем к следующей итерации. Закончив перечисление, перейдем к п. г);

г) увеличим счетчик на единицу, определим $\underline{J} = \underline{J} \cup \{\{i\}\}$ и вызовем $f(\underline{J})$;

д) уменьшим счетчик на единицу и выйдем из $f(arg)$.

Для запуска алгоритма вызовем $f(\underline{J})$.

Если требуется найти все решения, в условии (8) следует разместить знак нестрогого неравенства. Несколько ускоряет работу алгоритма В1 выполнение п. г) только по условию (8) со строгим неравенством. Также можно ожидать несколько более быстрое выполнение алгоритма В1 для задач, в которых оптимум совпадает с какой-либо оценкой m_l . Например, $m_l = m_0$ (см. (6)), если изначально определить $r = m_l$. Тогда для построения множества всех оптимальных разбиений можно пропустить п. а), а п. г) выполнять по условию (8). Для построения единственного разбиения в качестве условия (8) можно взять проверку условия $\underline{J} = \emptyset$, а п. г) также выполнять по условию (8). Для задач, в которых количество классов оказывается большим m_l , можно предусмотреть итеративное выполнение алгоритма с увеличением r на каждом шаге на единицу. Назовем такие алгоритмы В2.

Итоги тестов показали, что индивидуальных задач, для которых выполняется $r = m_0$ (т. е. оптимальное значение соответствует не самой лучшей оценке), оказывается больше примерно на порядок даже при небольших $k = 1, 15$. На рис. 6 показано распределение количества таких задач (обозначены зелеными точками) по сравнению с теми, в которых $r - m_0 = 1$ (обозначены желтыми точками), и теми, в которых $r - m_0 \geq 2$ (обозначены красными точками), по результатам четырех серий тестов. Для каждой из этих серий взято 135 индивидуальных задач с параметрами $k = 1, 15$, $n = 2, 10$, $w_i = R(\{0, \dots, k\})$, R – функция случайного выбора значений из аргумента. Таким образом, для большого количества задач можно ожидать меньшее среднее время выполнения алгоритма В2 по сравнению с В1.

Что касается проверки связности подграфа $G_j \in G$, то она менее удобна, чем в алгоритмах типа А. Не видно простой возможности осуществить такую проверку, пока не сформировано полное разбиение, поскольку на каждом шаге рекурсии любой из классов может быть дополнен. С другой стороны, операция взятия подмножеств заданной длины более вычислительно затратна, чем простое дополнение множества одним элементом. Как уже отмечалось ранее, тесты показывают более медленную работу по сравнению с алгоритмом А2 для одной и той же индивидуальной задачи КОВП₁ с критерием допустимой нагрузки 5 в). Для задачи КОВП₂ в каждом узле рекурсии можно проводить проверку принадлежности формируемых подмножеств разбиения подмножествам из P подобно способу, предложенному в работе [4].

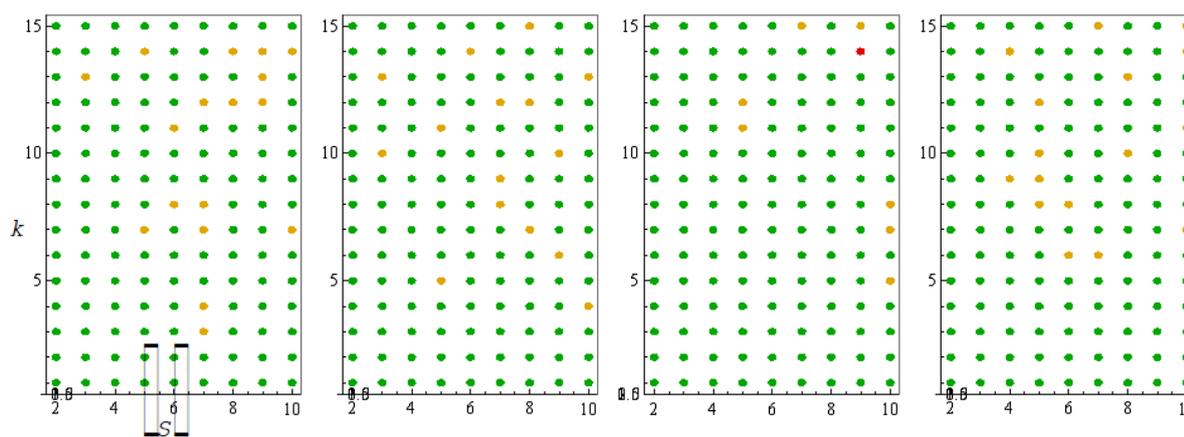


Рис. 6. Распределение задач $r = m_0$ и $r > m_0$

Вычислительные тесты комбинаторных алгоритмов. Рассмотрим сравнительные результаты вычислительных тестов для алгоритмов А2, В1, В2 в пакете Wolfram Mathematica 10.0.2.0 на компьютере с процессором Intel Celeron 2,7 ГГц и оперативной памятью 4 Гб.

В табл. 1–3 приведены результаты поиска одного решения при помощи алгоритмов A2, B1 и B2 для задач КОВП₁ с условием 4 а) (группа I) и без него (группа II). В качестве условия связности взята связность подграфов одной из структур на рис. 7. Задачи для обеих групп тестов для всех алгоритмов сгенерированы по одному и тому же набору из 75 массивов W весов $w_i = R(\{0, \dots, k\})$ при $n \in \{8, 9, 12, 15, 16\}$ и $k \in \{10, 15, 20\}$. Для каждой пары параметров взято пять задач, по которым определяется среднее время в секундах. В крайнем справа столбце приводится среднее время выполнения для каждого $|S|$ по всем k отдельно для групп I и II. Для сравнения алгоритмов можно сопоставлять значения в соответствующих ячейках трех таблиц.

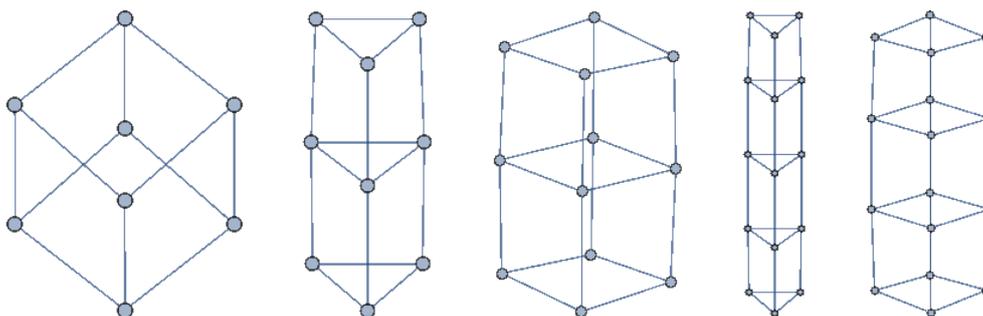


Рис. 7. Графовые структуры для тестов задач с условием связности подграфов

Таблица 1

Результаты тестов алгоритма A2

Группа	$n \backslash k$	10	15	20	Среднее
I	8	0	0	0	0
II		0,001 040 01	0,001 040 01	0,002 080 01	0,001 386 68
I	9	0,001 040 01	0,003 120 02	0,001 040 01	0,001 733 34
II		0,002 080 01	0,002 080 01	0,001 040 01	0,001 733 34
I	12	0,008 32	0,012 48	0,014 56	0,011 787
II		0,006 240 04	0,009 36	0,009 36	0,0214 935
I	15	0,100 881	0,082 161	0,127 921	0,103 654
II		0,227 761	0,126 881	0,060 32	0,138 321
I	16	0,316 162	0,238 162	0,190 321	0,248 215
II		2,538 656	0,138 321	0,442 003	1,039 66

Таблица 2

Результаты тестов алгоритма B1

Группа	$n \backslash k$	10	15	20	Среднее
I	8	0,002 080 01	0,002 080 01	0,001 040 01	0,001 733 34
II		0	0,001 040 01	0	0,000 346 67
I	9	0,003 120 02	0,005 200 03	0,002 080 01	0,003 466 69
II		0,003 120 02	0	0,001 040 01	0,001 386 68
I	12	0,112 321	0,0572	0,266 242	0,145 254
II		0,049 920 3	0,034 320 2	0,056 16	0,046 800 3
I	15	3,098 18	0,997 366	4,822 511	2,972 686
II		0,310 962	0,552 244	0,240 242	0,367 816
I	16	62,298 479	24,185 355	19,164 203	35,216 012
II		9,388 14	2,442 976	0,958 886	4,263 334

Таблица 3

Результаты тестов алгоритма В2

Группа	$n \backslash k$				
		10	15	20	Среднее
I	8	0,002 080 01	0,001 040 01	0,002 080 01	0,001 733 34
II		0,001 040 01	0	0	0,000 346 67
I	9	0,004 160 03	0,003 120 02	0,003 120 02	0,003 466 69
II		0,003 120 02	0,001 040 01	0,003 120 02	0,002 426 68
I	12	0,027 04	0,031 200 2	0,158 081	0,072 107 1
II		0	0,003 120 02	0,050 960 3	0,018 026 8
I	15	2,702 977	0,546 003	4,306 668	2,518 549
II		0,042 640 3	0,089 440 6	0,062 400 4	0,064 827 1
I	16	45,717 653	24,205 115	12,229 438	27,384 069
II		8,654 94	1,419 61	0,745 685	3,606 74

Как и ожидалось, алгоритм В2 работает эффективнее, чем В1, и является самым быстрым для задач, относящихся к группе II и $n < 16$, а алгоритм А2 работает значительно быстрее В1 и В2 для задач группы I. Единичные тесты алгоритма А2 для поиска одного решения задач, где $n \in \{18, 20, 24\}$ и задано условие связности, показывают результаты, которые можно считать приемлемыми для использования в практических целях конфигурирования секторов ВП небольшого региона. Время работы для $n = 24$ принимает значение порядка 2–3 мин.

Модель целочисленного линейного программирования. Задачу КОВП₂ с критерием допустимой нагрузки 5 в) без условия связности можно сформулировать в варианте постановки задачи ЦЛП для задачи упаковки в контейнеры:

$$\begin{cases}
 \sum_{j=1}^n y_j \rightarrow \min; \\
 \sum_{i=1}^n w_i x_{ij} \leq k y_j, \forall j \in \{1, \dots, n\}; \\
 \sum_{j=1}^n x_{ij} = 1, \forall i \in \{1, \dots, n\}; \\
 x_{ij}, y_j \in \{0, 1\},
 \end{cases} \quad (10)$$

где переменные x_{ij} обозначают размещение элемента s_i в объединение S_j , переменные y_j – наличие хотя бы одного элемента в объединении S_j .

Модель показывает хорошую производительность при использовании специализированных решателей. Например, единичный тест решения задач без условия связности при $n \in \{10, 20, 30, 40\}$, $k = 15$, $w_i = R(\{0, \dots, k\})$ с применением надстройки для MS Excel решателя IBM ILOG CPLEX показывает следующие значения времени работы решателя до получения оптимального решения при начальном решении $x_{ij}, y_j = 0$: 3 с для $n = 10$; 10 с для $n = 20$; 20 с для $n = 30$; 30 с для $n = 40$. Заметим, что только количество переменных модели (10) составляет $n^2 + n$.

Можно пробовать учесть условие связности, решая задачу (10) повторно с добавлением исключающих ограничений для недопустимых объединений. Например, пусть в результате решения получено несколько таких объединений. Обозначим множество индексов элементов в каком-либо недопустимом объединении J'' и множество индексов таких объединений \underline{J}'' . Тогда к задаче (10) могут быть добавлены ограничения

$$\sum_{i \in J''} x_{ij} \leq |J''| - 1, \forall J'' \in \underline{J}'' \quad (11)$$

затем произведена попытка повторного решения и т. д.

Для задачи с критерием допустимой нагрузки 5 б) и дополнительным условием 4 б) можно предложить следующую постановку:

$$\begin{cases} \sum_{j \in \underline{J}''_1} y_j \rightarrow \min; \\ y_j \sum_{i=1}^n w_i a_{ij} \leq k_j, \forall j \in \underline{J}''_1; \\ \sum_{j \in \underline{J}''_1} a_{ij} y_j = 1, \forall i \in \{1, \dots, n\}; \\ y_j \in \{0, 1\}, \end{cases} \quad (12)$$

где \underline{J}''_1 – множество индексов объединений из P ; $a_{ij} \in \{0, 1\}$ – коэффициент, определяющий возможность вхождения элемента s_i в объединение S_j . Если проводится предобработка P с отбрасыванием подмножеств с превышением допустимой нагрузки, т. е. решается задача КОВП₂, то в постановке (12) остается только вторая группа ограничений и получается вариант постановки задачи о покрытии. Преимуществом (12) может являться небольшое количество переменных.

В работе [9] предложена оригинальная постановка для учета условия связности с применением модели потоков, однако количество только переменных потоков там составляет n^3 . Тесты производительности модели ЦПП с ограничениями для условия связности и сравнение их результатов с постановкой из [9] пока не проводились.

Также, как следует из определения круга решаемых в настоящем исследовании задач, авторы не претендуют на сравнение предложенной формулировки с теми, которые предлагаются в более комплексных моделях.

Следует отметить, что решение задачи методом математического программирования с использованием решателей обладает тем недостатком, что известные решатели не имеют настроек для получения множества оптимальных решений в случае их существования. Предложенные в предыдущем разделе алгоритмы, использующие метод ветвей и границ, могут быть легко адаптированы для получения всех оптимальных решений, которых в случае рассматриваемой задачи разбиения конечного множества часто оказывается несколько и может оказаться довольно много для заданной индивидуальной задачи. Получение множества решений может дать возможность учесть какие-либо дополнительные неформальные критерии допустимости или оптимальности при рассмотрении практической задачи комбинирования секторов ВП.

Заключение. В статье представлено описание и результаты тестов нескольких вычислительных методов решения элементарной задачи комбинирования областей ВП с критерием оптимальности и ограничениями, определенными в разделе «Постановка задачи». Предложенные методы могут найти применение при моделировании комплексных задач конфигурации ВП, разделенного на исходные элементарные секторы.

По результатам тестов на производительность предлагаются следующие способы решения задач, обозначенных как КОВП₁ и КОВП₂. Задачу КОВП₁ с условием связности и количеством исходных секторов в пределах 24 предлагается решать при помощи комбинаторного алгоритма, обозначенного как А2, время решения прогнозируется в пределах нескольких минут. Алгоритм А2 может быть применен также для решения всех сформулированных вариантов задачи, если требуется получение нескольких решений. Для задачи КОВП₁ с условием связности и количеством исходных секторов более 24 авторы считают целесообразным провести дополнительные исследования производительности нескольких моделей, построенных по методу ЦПП. Задачу КОВП₁ без условия связности и задачу КОВП₂ более рационально решать мето-

дом ЦПП с применением специализированных решателей. Применение такого метода дает в результате производительность менее минуты для количества элементарных секторов в пределах 40.

Список использованных источников

1. Flener, P. Automatic Airspace Sectorisation: A Survey [Electronic resource] / P. Flener, J. Pearson. – Department of Information Technology, Uppsala University, Sweden, 2018. – Mode of access: <https://arxiv.org/abs/1311.0653>. – Date of access: 03.06.2020.
2. Applying graph theory to problems in air traffic management / A. H. Farrahi [et al.] // 17th AIAA AVIATION Technology, Integration, and Operations Conference, Denver, Colorado, 5–9 June 2017 / AIAA AVIATION Forum. – Denver, Colorado, 2017. – 20 p.
3. Bloom, M. Combining airspace sectors for the efficient use of air traffic control resources / M. Bloom, P. Kopardekar // Navigation, and Control (GNC) Conference and Exhibit, Honolulu, HI, 18–21 Aug. 2008. – American Institute of Aeronautics and Astronautics, 2008. – 15 p.
4. Gianazza, D. Forecasting workload and airspace configuration with neural networks and tree search methods / D. Gianazza // Artificial Intelligence. – 2010. – № 174(7–8). – P. 530–549.
5. Дегтярев, О. В. Решение задач секторизации района управления воздушным движением. I. Основные принципы и проблемы секторизации воздушного пространства и ее формализация как оптимизационной задачи / О. В. Дегтярев, В. Н. Минаенко, М. О. Орехов // Известия РАН. Теория и системы управления. – 2009. – № 3. – С. 56–72.
6. Дегтярев, О. В. Решение задач секторизации района управления воздушным движением. II. Разработка алгоритмов секторизации / О. В. Дегтярев, В. Н. Минаенко, М. О. Орехов // Известия РАН. Теория и системы управления. – 2010. – № 4. – С. 117–135.
7. Вересов, К. А. Решение задач секторизации района управления воздушным движением. III. Разработка алгоритмов определения конфигураций и временного графика работы секторов управления / К. А. Вересов, О. В. Дегтярев, В. Н. Минаенко // Известия РАН. Теория и системы управления. – 2013. – № 2. – С. 133–153.
8. Bloom, M. Algorithms for combining airspace sectors / M. Bloem, P. Gupta, P. Kopardekar // Air Traffic Control Quarterly. – Sept. 2009. – Vol. 17, no. 3.
9. Drew, M. C. A method of optimally combining sectors / M. C. Drew // 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO), Hilton Head, South Carolina, 21–23 Sept. 2009. – American Institute of Aeronautics and Astronautics, 2009. – P. 7057.
10. Martello, S. An Knapsack Problems. Algorithms and Computer Implementations / S. Martello, P. Toth. – John Wiley & Sons, 1990. – 296 p.
11. Стенли, Р. Перечислительная комбинаторика : в 2 т. / Р. Стенли. – М. : Мир, 1990. – Т. 1. – 440 с.
12. Гэри, М. Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон. – М. : Мир, 1982. – 416 с.

References

1. Flener P., Pearson J. *Automatic Airspace Sectorisation: A Survey*. Department of Information Technology, Uppsala University, Sweden, 2018. Available at: <https://arxiv.org/abs/1311.0653> (accessed 03.06.2020).
2. Farrahi A. H., Goldberg A. T., Bagasol L. N., Jaewoo J. Applying graph theory to problems in air traffic management. *17th AIAA AVIATION Technology, Integration, and Operations Conference, Denver, Colorado, 5–9 June 2017, AIAA AVIATION Forum*. Denver, Colorado, 2017, 20 p. <https://doi.org/10.2514/6.2017-3775>
3. Bloom M., Kopardekar P. Combining airspace sectors for the efficient use of air traffic control resources. *Navigation, and Control (GNC) Conference and Exhibit, Honolulu, HI, 18–21 August 2008*. American Institute of Aeronautics and Astronautics, 2008, 15 p.
4. Gianazza D. Forecasting workload and airspace configuration with neural networks and tree search methods. *Artificial Intelligence*, 2010, no. 174(7–8), pp. 530–549.
5. Degtyarev O. V., Minaenko V. N., Orekhov M. O. Reshenie zadach sektorizacii rajona upravlenija vozdushnym dvizheniem [Reservation of regional sector regulations]. I. Osnovnye principy i problemy sektorizacii vozdushnogo prostranstva i ee formalizacija kak optimizacionnoj zadachi [Basic principles and problems of airspace sectorization and its formalization as an optimization problem]. *Izvestija Rossijskoj akademii nauk. Teorija i sistemy upravlenija [Bulletin of the Russian Academy of Sciences. Control Theory and Systems]*, 2009, no. 3, pp. 56–72.

6. Degtyarev O. V., Minaenko V. N., Orekhov M. O. Reshenie zadach sektorizacii rajona upravlenija vozdušnym dvizheniem [Solving the Problems of Sectorization of the Air Traffic Control Area]. II. Razrabotka algoritmov sektorizacii [Development of sectorization algorithms]. *Izvestija Rossijskoj akademii nauk. Teorija i sistemy upravlenija [Bulletin of the Russian Academy of Sciences. Control Theory and Systems]*, 2010, no. 4, pp. 117–135.
7. Veresov K. A., Degtyarev O. V., Minaenko V. N. Reshenie zadach sektorizacii rajona upravlenija vozdušnym dvizheniem [Solving the problems of sectorization of the air traffic control area]. III. Razrabotka algoritmov opredelenija konfiguracij i vremennogo grafika raboty sektorov upravlenija [Development of algorithms for determining configurations and a time schedule for the operation of control sectors]. *Izvestija Rossijskoj akademii nauk. Teorija i sistemy upravlenija [Bulletin of the Russian Academy of Sciences. Control Theory and Systems]*, 2013, no. 2, pp. 133–153.
8. Bloom M., Gupta P., Kopardekar P. Algorithms for combining airspace sectors. *Air Traffic Control Quarterly*, September 2009, vol. 17, no. 3.
9. Drew M. C. A method of optimally combining sectors. *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO), Hilton Head, South Carolina, 21–23 September 2009*. American Institute of Aeronautics and Astronautics, 2009, p. 7057.
10. Martello S., Toth P. *An Knapsack Problems. Algorithms and Computer Implementations*. John Wiley & Sons, 1990, 296 p.
11. Stanley R. *Enumerative Combinatorics*. Vol. 1. Cambridge University Press, 2012, 440 p.
12. Garey M., Johnson D. *Computers and Intractability*. New York, W. H. Freeman and Company, 1979, 340 p.

Информация об авторах

Рубанов Игорь Владимирович, старший преподаватель кафедры естественнонаучных дисциплин, Белорусская государственная академия авиации, Минск, Беларусь.
E-mail: irubanov@inbox.ru

Ковалев Михаил Яковлевич, доктор физико-математических наук, профессор, член-корреспондент НАН Беларуси, заместитель генерального директора, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.
E-mail: kovalyov_my@newman.bas-net.by

Information about the authors

Igor V. Rubanov, Senior Lecturer of the Department of Science Disciplines, Belarusian State Academy of Aviation, Minsk, Belarus.
E-mail: irubanov@inbox.ru

Mikhail Y. Kovalyov, Dr. Sci. (Phys.-Math.), Professor, Corresponding Member of the National Academy of Sciences of Belarus, Deputy General Director, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
E-mail: kovalyov_my@newman.bas-net.by

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

УДК 519.6, 004.94
<https://doi.org/10.37661/1816-0301-2020-17-4-22-35>

Поступила в редакцию 20.03.2020
Received 20.03.2020

Принята к публикации 12.10.2020
Accepted 12.10.2020

Вейвлет-преобразование на конечном интервале

В. М. Романчук

Белорусский национальный технический университет, Минск, Беларусь
E-mail: Romanchak@bntu.by

Аннотация. Рассматриваются интегральные преобразования на конечном интервале с сингулярным базисным вейвлетом. С помощью последовательности таких преобразований решается задача непараметрической аппроксимации функции. Традиционно считается, что для базисного вейвлета должно выполняться условие допустимости, т. е. среднее значение вейвлета должно равняться нулю. Существует ряд сингулярных вейвлетов, для которых условие допустимости не выполняется. В этом случае в качестве базисного вейвлета можно использовать дельтообразные функции, которые участвуют в оценках Парзена – Розенблатта и Надарая – Ватсона. Исследование ряда вейвлет-преобразований на конечном интервале проводится только в одном частном случае из-за технических сложностей при попытке непосредственно решить эту задачу. Реализуется идея периодического продолжения вейвлет-преобразования с конечного интервала на всю числовую ось, которая позволяет сформулировать достаточные условия сходимости. Приводится пример аппроксимации с помощью суммы дискретных вейвлет-преобразований.

Ключевые слова: вейвлет, вейвлет-преобразование, окно Парзена – Розенблатта, непараметрическая аппроксимация, ядерная оценка Надарая – Ватсона

Для цитирования. Романчук, В. М. Вейвлет-преобразование на конечном интервале / В. М. Романчук // Информатика. – 2020. – Т. 17, № 4. – С. 22–35. <https://doi.org/10.37661/1816-0301-2020-17-4-22-35>

Wavelet transformation on a finite interval

Vasily M. Romanchak

Belarusian National Technical University, Minsk, Belarus
E-mail: Romanchak@bntu.by

Abstract. Integral transformations on a finite interval with a singular basis wavelet are considered. Using a sequence of such transformations, the problem of nonparametric approximation of a function is solved. Traditionally, it is assumed that the validity condition must be met for a basic wavelet (the average value of the wavelet must be zero). The paper develops the previously proposed method of singular wavelets when the tolerance condition is not met. In this case Delta-shaped functions that participate in Parzen – Rosenblatt and Nadaraya – Watson estimations can be used as a basic wavelet. The set of wavelet transformations for a function defined on a numeric axis, defined locally, and on a finite interval were previously investigated. However, the study of the convergence of the decomposition on a finite interval was carried out only in one particular case. It was due to technical difficulties when trying to solve this problem directly. In the paper the idea of evaluating the periodic continuation of a function defined initially on a finite interval is implemented. It allowed to formulate sufficient convergence conditions for the expansion of the function in a series. An example of approximation of a function defined on a finite interval using the sum of discrete wavelet transformations is given.

Keywords: wavelet, wavelet transform, the Parzen – Rosenblatt window method, nonparametric estimator, Nadaraya – Watson kernel regression

For citation. Romanchak V. M. Wavelet transformation on a finite interval. *Informatics*, 2020, vol. 17, no. 4, pp. 22–35 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-22-35>

Введение. Методы непараметрической аппроксимации в настоящее время широко распространены. Непараметрическую аппроксимацию можно использовать для построения различных математических моделей [1]. В прикладных работах часто применяют аппроксимацию ядерными функциями [2, 3] и вейвлетами [4, 5]. Если ядерные функции использовать в качестве базисного вейвлета в интегральном вейвлет-преобразовании, то получим метод сингулярных вейвлетов. При аппроксимации методом сингулярных вейвлетов происходит суммирование ядерных оценок Надарая – Ватсона [6] по параметру размытости, поэтому отпадает необходимость решать задачу оптимального выбора этого параметра. Преобразование для бесконечного промежутка рассматривалось в работах [7, 8]. Оценка сходимости суммы вейвлет-преобразований для бесконечного промежутка получена в работе [9], для конечного промежутка – в работе [10].

В настоящей статье рассматриваются более общие условия сходимости для конечного промежутка. Для этого определяется периодическое продолжение вейвлет-преобразования с конечного интервала на всю числовую ось. В качестве примера приложения теории показано, что в некоторых случаях из сигнала можно выделить медленную и быструю компоненты.

Целью исследования является развитие метода сингулярных вейвлетов, в частности получение достаточных условий сходимости ряда, составленного из вейвлет-преобразований на конечном промежутке. В работе доказываются две теоремы, которые содержат формулировки условий сходимости ряда из вейвлет-преобразований. При доказательстве используется определение периодического продолжения функции $f(x)$ с конечного интервала $[A, B]$ на бесконечный интервал $(-\infty, +\infty)$.

Преобразование с сингулярным вейвлетом

Определение. Пусть для функции $\psi(x)$ выполняется условие на бесконечности

$$|\psi(x)| \leq \frac{d}{1+x^2}, \quad (1)$$

где d – некоторая константа, $d > 0$, и для функции $\psi(x)$ существует конечное среднее значение

$$C_\psi = \int_{-\infty}^{\infty} \psi(\tau) d\tau. \quad (2)$$

Тогда функция $\psi(x)$ принадлежит пространству $L^1(R)$. Такую функцию $\psi(x)$ назовем *базисным вейвлетом*. Обычно для базисного вейвлета должно выполняться условие допустимости: среднее определяемое по формуле (2) должно равняться нулю, $C_\psi = 0$. Если считать, что для вейвлета среднее не равно нулю, $C_\psi \neq 0$, то такой базисный вейвлет будем называть *сингулярным* [7]. Например, сингулярным вейвлетом является функция плотности нормального распределения. Базисный вейвлет может быть задан с точностью до постоянного множителя. Поэтому считаем, что среднее сингулярного вейвлета равно единице, $C_\psi = 1$. Преобразование для бесконечного промежутка определяется формулой [8, 9]

$$W(f(x) - f)(x, a) = \frac{1}{a} \int_{-\infty}^{\infty} (f(x) - f(\tau)) \psi\left(\frac{\tau - x}{a}\right) d\tau, \quad (3)$$

где $a > 0$, $b \in R$.

В частном случае, если $C_\psi = 0$, преобразование (3) имеет вид

$$W(f)(x, a) = \frac{1}{a} \int_{-\infty}^{\infty} f(\tau) \psi\left(\frac{\tau - x}{a}\right) d\tau.$$

Здесь преобразование (3) с точностью до постоянной совпадает со стандартным вейвлет-преобразованием [1]. При $C_\psi = 1$ преобразование (3) можно записать в виде

$$W(f(x) - f)(x, a) = f(x) - \frac{1}{a} \int_{-\infty}^{\infty} f(\tau) \psi\left(\frac{\tau - x}{a}\right) d\tau.$$

Введение сингулярных вейвлетов расширяет возможности теории вейвлетов и ядерных оценок типа Надарая – Ватсона [10]. Определим преобразование на конечном промежутке:

$$W(f(x) - f)(x, a) = f(x) - Wf(x, a), \quad (4)$$

где $Wf(x, a)$ – вейвлет-преобразование вида

$$Wf(x, a) = \frac{1}{aC(x)} \int_A^B f(\tau) \psi\left(\frac{\tau-x}{a}\right) d\tau. \quad (5)$$

Здесь $C(x) = \int_A^B \psi\left(\frac{\tau-x}{a}\right) d\tau$, где $\psi(x)$ – сингулярный вейвлет.

Последовательность преобразований. Используя сингулярный вейвлет $\psi(x)$ и формулы (4), (5), запишем последовательность преобразований на конечном интервале:

$$F_{k+1}(x) = F_k(x) - WF_k(x, a_k), \quad (6)$$

где

$$WF_k(x, a_k) = \frac{1}{a_k C_k(x)} \int_A^B F_k(\tau) \psi\left(\frac{\tau-x}{a_k}\right) d\tau, \quad (7)$$

$$C_k(x) = \int_A^B \psi\left(\frac{\tau-x}{a_k}\right) d\tau, \quad F_0(x) = f(x), \quad k = 0, 1, 2, \dots$$

Непосредственно доказываем утверждение, что если функция $F_k(x)$ принадлежит пространству $L^1[A, B]$, то функция $F_{k+1}(x) = W(f - f)(x, a)$ также принадлежит пространству $L^1[A, B]$ [10]. С помощью последовательности (6) можно аппроксимировать функцию $f(x)$. Если вейвлет $\psi(x)$ является неотрицательной функцией, то в формуле (8) функция $C_k(x)$ неотрицательна, $C_k(x) > 0$, и преобразования (6), (7) существуют для любых положительных a_k .

Формула разложения. Если функция $f(x)$ на промежутке $[A, B]$ принадлежит пространству $L^1[A, B]$, то ее можно представить в виде суммы вейвлет-преобразований и остаточного члена:

$$f(x) = \sum_{k=0}^K WF_k(x, a_k) + F_{K+1}(x), \quad (8)$$

где a_k – произвольные положительные действительные числа;

$F_k(x)$ – элемент последовательности преобразований (6), $F_0(x) = f(x)$;

$WF_k(x, a)$ – вейвлет-преобразование, определяемое по формуле (7);

$F_{K+1}(x)$ – остаточный член разложения, который находится по формуле (7);

k – порядковый номер вейвлет-преобразования, $0 \leq k \leq K + 1$;

$K + 1$ – порядок приближения, $K \geq 0$.

Ограничимся случаем, когда вейвлет $\psi(x)$ является неотрицательной функцией, т. е. существует последовательность вейвлет-преобразований (6), (7). Тогда формула (8) выполняется тождественно для любых положительных значений a_k . Например, при $K = 1$ из рекуррентной формулы (6) следуют равенства

$$F_1(x) = F_0(x) - WF_0(x, a_0), \quad F_2(x) = F_1(x) - WF_1(x, a_1).$$

Суммируя эти равенства, получим $F_2(x) = F_0(x) - WF_0(x, a_0) - WF_1(x, a_1)$. Учитывая, что $F_0(x) = f(x)$, будет выполняться формула

$$f(x) = WF_0(x, a_0) + WF_1(x, a_1) + F_2(x). \quad (9)$$

При произвольном порядке K формула (9) доказывается аналогично. Для оценки остаточного члена $F_{k+1}(x)$ понадобятся вспомогательные определения. Сумму в выражении (8) можно рассматривать как частичную сумму ряда. Если остаточный член разложения стремится к нулю, ряд сходится.

Вспомогательные определения

Приращения высших порядков. Пусть функция f задана на промежутке $[A, B]$. Будем говорить, что задано приращение первого порядка в точке $z \in [A, B]$, если определено выражение

$$\Delta f(\Delta_0, z) = f(\Delta_0 + z) - f(z), \quad (10)$$

где $z \in [A, B]$, $\Delta_0 + z \in [A, B]$. Приращение второго порядка в точке z найдем по формуле

$$\Delta^2 f(\Delta_0, \Delta_1, z) = \Delta f(\Delta_0, \Delta_1 + z) - \Delta f(\Delta_0, z), \quad (11)$$

где $\Delta_1 + z \in [A, B]$.

В общем случае приращение в точке z порядка $k + 1$ зададим равенством

$$\Delta^{k+1} f(\Delta_0, \dots, \Delta_{k-1}, \Delta_k, z) = \Delta^k f(\Delta_0, \dots, \Delta_{k-1}, \Delta_k + z) - \Delta^k f(\Delta_0, \Delta_1, \dots, \Delta_{k-1}, z),$$

где $\Delta_k + z \in [A, B]$. Приращение k -го порядка является симметрической функцией своих параметров $\Delta_0, \Delta_1, \dots, \Delta_k$, т. е. при любой их перестановке значение функции не меняется. В частности,

$$\Delta^2 f(\Delta_0, \Delta_1, z) = \Delta^2 f(\Delta_1, \Delta_0, z),$$

$$\Delta^{k+1} f(\Delta_0, \Delta_1, \dots, \Delta_k, z) = \Delta^{k+1} f(\Delta_k, \dots, \Delta_1, \Delta_0, z).$$

Условия Липшица высших порядков. Пусть функция f задана на промежутке $[A, B]$. Будем говорить, что для функции $f(z)$ выполняется условие Липшица первого порядка в точке $z \in [A, B]$, если определена разность первого порядка и верно неравенство

$$|\Delta f(\Delta_0, z)| \leq L|\Delta_0|.$$

Условие Липшица второго порядка в точке $z \in [A, B]$ выполняется, если для разности второго порядка выполняется неравенство

$$\Delta^2 f(\Delta_0, \Delta_1, z) \leq L|\Delta_0||\Delta_1|.$$

Аналогично определим в точке $z \in [A, B]$ условие Липшица порядка $m + 1$:

$$\Delta^{m+1} f(\Delta_0, \Delta_1, \dots, \Delta_m, z) \leq L|\Delta_0||\Delta_1|, \dots, |\Delta_m|.$$

С помощью условия Липшица первого порядка можно оценить разность второго порядка:

$$|\Delta^2 f(\Delta_0, \Delta_1, z)| \leq |\Delta^1 f(\Delta_1, z + \Delta_0)| + |\Delta^1 f(\Delta_1, z)| \leq 2L_1|\Delta_1|.$$

Подобным образом можно получить оценку для разности произвольного порядка:

$$\Delta^{k+1} f(\Delta_0, \Delta_1, \dots, \Delta_k, z) \leq 2^k |\Delta_k| L, \quad (12)$$

где $z \in [A, B]$. Согласно условию Липшица второго порядка получим оценку для разности третьего порядка

$$\left| \Delta^3 f(\Delta_0, \Delta_1, \Delta_2, z) \right| \leq \left| \Delta^2 f(\Delta_1, \Delta_2, z + \Delta_0) \right| + \left| \Delta^2 f(\Delta_1, \Delta_2, z) \right| \leq 2L |\Delta_1| |\Delta_2|.$$

Для разности произвольного порядка будет выполняться неравенство

$$\left| \Delta^{k+1} F(\Delta_0, \dots, \Delta_{k-1}, \Delta_k, z) \right| \leq 2^{k-1} L |\Delta_k| |\Delta_{k-1}|. \quad (13)$$

С помощью условия Липшица порядка m можно оценить разность порядка k ($k > m$):

$$\left| \Delta^{k+1} F(\Delta_0, \dots, \Delta_{k-1}, \Delta_k, z) \right| \leq 2^{k-m+1} L |\Delta_{k-m+1}|, \dots, |\Delta_{k-1}| |\Delta_k|. \quad (14)$$

Остаточный член разложения. Вначале для функции $f(x)$ и базисного вейвлета $\psi(x)$ определим периодические продолжения. Положим, что функция $f(x)$ задана на промежутке $[A, B]$. Дополнительно будем рассматривать продолжение функции $f(x)$, $x \in [A, B]$, на числовую полуось $[A, \infty)$ по формуле

$$f(x+T) = f(x),$$

где $T = B - A$. Введем также периодическое продолжение базисного вейвлета $\psi(u_k)$, определенного на промежутке $[0, T_k)$, где $T_k = (B - A)/a_k$, $a_k > 0$, на числовую полуось $[0, \infty)$ по формуле

$$\psi(u_k + T_k) = \psi(u_k).$$

Сделав замену переменных $\tau = x + a_k u_k$ в выражениях (6) и (7), получим преобразование

$$F_{k+1}(x) = \frac{1}{C_k} \int_{A_k - H_k}^{B_k - H_k} (F_k(x) - F_k(x + a_k u_k)) \psi(u_k) du_k, \quad (15)$$

где $C_k = \int_{A_k - H_k}^{B_k - H_k} \psi(u_k) du_k$, $A_k = A/a_k$, $B_k = B/a_k$, $H_k = x/a_k$.

Можно считать, что для соотношения (15) определено периодическое продолжение функций $F_k(x + a_k u_k)$ и $\psi(u_k)$ с периодом $T_k = (B - A)/a_k$ по переменной u_k . Длина промежутка интегрирования в преобразовании (15) равна T_k и совпадает с периодом подынтегральной функции. Поэтому равенство (15) выполняется для произвольной постоянной H_k и следующее преобразование можно записать в виде равенства

$$F_{k+1}(x) = \frac{1}{C_k} \int_0^{T_k} (F_k(x) - F_k(x + a_k u_k)) \psi(u_k) du_k, \quad (16)$$

где $C_k = \int_0^{T_k} \psi(u_k) du_k$, $T_k = (B - A)/a_k$. Используя рекуррентную последовательность (16), получим формулу для остаточного члена в разложении (8). Формула (16) для $k = 1$ имеет вид

$$F_2(x) = \frac{1}{C_1} \int_0^{T_1} (F_1(x) - F_1(x + a_1 u_1)) \psi(u_1) du_1. \quad (17)$$

С учетом разности первого порядка (10) для $k = 0$ и на основании формулы (16) получим выражение

$$F_1(x) = -\frac{1}{C_0} \int_0^{T_0} \Delta F_0(a_0 u_0, x) \psi(u_0) du_0. \quad (18)$$

Подставив (18) в равенство (17) и используя определение разности второго порядка (11), запишем

$$\begin{aligned} F_2(x) &= \frac{1}{C_0 C_1} \int_0^{T_1} \int_0^{T_0} (\Delta F_0(a_0 u_0, x + a_1 u_1) - \Delta F_0(a_0 u_0, x)) \psi(u_0) \psi(u_1) du_0 du_1 = \\ &= \frac{1}{C_0 C_1} \int_0^{T_1} \int_0^{T_0} \Delta^2 F_0(a_0 u_0, a_1 u_1, x) \psi(u_0) \psi(u_1) du_0 du_1. \end{aligned}$$

Таким образом, в общем случае остаточный член разложения имеет вид

$$F_{k+1}(x) = \frac{(-1)^k}{C_0 \dots C_k} \int_0^{T_k} \dots \int_0^{T_0} \Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x) \psi(u_0) \dots \psi(u_k) du_0 \dots du_k, \quad (19)$$

где $T_k = (B - A) / a_k$, $a_k > 0$, $C_k = \int_0^{T_k} \psi(u_k) du_k$.

Если условие Липшица выполняется для функции $f(z)$ в точке $z \in [A, B]$, то оно будет выполняться в этой точке и для периодического продолжения функции $f(z)$ на множество $[A, +\infty)$. Следовательно, условие Липшица можно применять для разности порядка $k + 1$ в выражении под знаком интеграла в формуле (19).

Достаточные условия равномерной сходимости. Применение периодического продолжения позволяет получить оценки остаточного члена разложения в форме (19). Докажем две теоремы.

Теорема 1. Пусть для функции $f(x)$ в точке $x \in [A, B]$ выполняется условие Липшица порядка m , $\psi(x)$ – сингулярный неотрицательный вейвлет. Будем считать, что интеграл в выражении

$$I_\psi = \int_{-\infty}^{\infty} |x| \psi(x) dx$$

сходится и параметры вейвлет-преобразований (7) образуют геометрическую прогрессию, такую, что $a_k = a_0 q^k$, $k = 0, 1, \dots$, $a_0 > 0$. Тогда остаточный член разложения (8) стремится к нулю в точке x для таких значений q , что $0 < q < 2^{-1/m}$.

Доказательство. Ограничимся доказательством теоремы для продолжения функции $f(x)$ на числовую полуось $[A, \infty)$:

1. Разберем случай, когда для периодического продолжения функции $f(x)$ в точке $x \in [A, B]$ выполняется условие Липшица первого порядка (12) и верно неравенство

$$\Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x) \leq 2^k a_k |u_k| L. \quad (20)$$

Следовательно, для остаточного члена в форме (19) выполняется оценка

$$|F_{k+1}(x)| \leq \frac{2^k a_k L}{C_0 \dots C_k} \int_0^{T_k} \dots \int_0^{T_0} u_k \Psi(u_0) \dots \Psi(u_k) du_0 \dots du_k.$$

Тогда верно неравенство

$$|F_{k+1}(x)| \leq \frac{2^k a_k L}{C_k} \int_0^{T_k} u_k \Psi(u_k) du_k \leq \frac{2^k a_k L}{C_0} \int_{-\infty}^{\infty} |u_k| \Psi(u_k) du_k = \frac{2^k a_k L}{C_0} I_{\Psi}. \quad (21)$$

Здесь учитывается, что выполняется соотношение $C_k > C_0$, так как $B_k - A_k > B_0 - A_0$. Соответственно,

$$|F_{k+1}(x)| \leq c' 2^k q^k, \quad (22)$$

где c' – постоянная, не зависящая от номера k и выбора точки x . Таким образом, остаточный член стремится к нулю при $0 < q < 2^{-1}$ и теорема для условия Липшица первого порядка доказана.

2. Пусть выполняется условие Липшица второго порядка (13) для периодического продолжения функции $f(x)$, тогда справедливо неравенство

$$\Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x) \leq 2^{k-1} a_k a_{k-1} |u_k| |u_{k-1}| L.$$

Для остаточного члена (19) выполняется оценка

$$|F_{k+1}(x)| \leq \frac{2^{k-1} a_k a_{k-1} L}{C_k C_{k-1}} \int_0^{T_k} \int_0^{T_{k-1}} u_{k-1} u_k \Psi(u_{k-1}) \Psi(u_k) du_{k-1} du_k \leq 2^{k-1} a_k a_{k-1} \frac{I_{\Psi}^2}{C_0^2} L.$$

Следовательно, для значений $a_k = a_0 q^k$ верно неравенство

$$|F_{k+1}(x)| \leq c' 2^k q^{2k}, \quad (23)$$

где c' – постоянная. Таким образом, ряд в разложении сходится при $q < 2^{-1/2}$ и теорема для условия Липшица второго порядка доказана.

3. Пусть выполняется условие Липшица (14) порядка m для функции $f(x)$ в точке $x \in [A, B]$. Можно показать, что остаточный член разложения стремится к нулю при выполнении условия $0 < q < 2^{-1/m}$. Случай продолжения функции $f(x)$ на числовую полуось $(-\infty, B]$ рассматривается аналогично. Теорема доказана.

Можно получить оценку для остаточного члена при других предположениях. Например, если не требовать существования интеграла I_{Ψ} в выражении (20).

Теорема 2. Пусть для функции $f(x)$ в точке $x \in [A, B]$ выполняется условие Липшица порядка m , $\Psi(x)$ – сингулярный неотрицательный вейвлет и параметры a_k вейвлет-преобразований (7) образуют геометрическую прогрессию, $a_k = a_0 q^k$, $a_k > 0$, $k = 0, 1, \dots$. Тогда остаточный член формулы разложения (8) стремится к нулю в точке x для таких значений q , что $0 < q < 2^{-3/m}$.

Доказательство. Ограничимся доказательством теоремы для продолжения функции $f(x)$, определенной для $x \in [A, B]$, на числовую полуось $[A, \infty)$, предполагая, что выполняются условия Липшица первого порядка, второго порядка и для общего случая:

1. Пусть выполняется условие Липшица первого порядка для функции $f(x)$ в точке $x \in [A, B)$. Так как $a_k |u_k| < B - A$, из выражения (20) следует неравенство

$$|\Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x)| \leq 2^k L(B - A), \quad (24)$$

которое остается верным для продолжения функции $f(x)$ на числовую полуось $[A, \infty)$.

Введем обозначения:

$$\Psi_k^\alpha(x) = \begin{cases} \psi(x), & |x| \leq M_k, \\ 0, & |x| > M_k; \end{cases} \quad (25)$$

$$\Psi_k^\beta(x) = \begin{cases} 0, & |x| \leq M_k, \\ \psi(x), & |x| > M_k, \end{cases} \quad (26)$$

где M_k – некоторое положительное число, $\psi(t)$ – базисный вейвлет. Тогда верно равенство

$$\Psi(x) = \Psi_k^\alpha(x) + \Psi_k^\beta(x). \quad (27)$$

В точке $x \in [A, B)$ для функции $f(x)$ с продолжением на числовой полуоси $[A, \infty)$ выполняется неравенство

$$|\Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x) \Psi(u_k)| \leq 2^k a_k |u_k| L \Psi_k^\alpha(u_k) + 2^k L(B - A) \Psi_k^\beta(u_k),$$

которое следует из равенства (27) и неравенств (20) и (24).

Если учесть неравенство (1) и оценку $|u_k| L \Psi_k^\alpha(u_k) \leq |M_k| L \Psi_k^\alpha(u_k)$, получим выражение

$$|\Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x) \Psi(u_k)| \leq 2^k L \left(a_k M_k \Psi_k^\alpha(u_k) + \frac{d(B - A)}{M_k^2} \right). \quad (28)$$

Следовательно, для остаточного члена разложения в форме (19) выполняются соотношения

$$\begin{aligned} |F_{k+1}(x)| &\leq \frac{2^k L}{C_0 \dots C_k} \int_0^{T_k} \dots \int_0^{T_0} \left(a_k M_k \Psi_k^\alpha(u_k) + \frac{d(B - A)}{M_k^2} \right) \Psi(u_0) \dots \Psi(u_{k-1}) du_0 \dots du_k \leq \\ &\leq \frac{2^k L}{C_k} \int_0^{T_k} \left(a_k M_k \Psi(u_k) + \frac{d(B - A)}{M_k^2} \right) du_k \end{aligned}$$

и

$$|F_{k+1}(x)| \leq 2^k L \left(a_k M_k + \frac{d(B - A)^2}{a_k C_0 M_k^2} \right). \quad (29)$$

Оценка (29) найдена с применением соотношений

$$\frac{1}{C_k} \int_0^{T_k} \psi(u_k) du_k \leq \frac{1}{C_k} \int_0^\infty \psi(u_k) du_k = 1, \quad \int_0^{T_k} du_k = T_k = \frac{B-A}{a_k}, \quad C_k \geq C_0.$$

Тогда, если считать, что $M_k = (a_k)^{-2/3}$ и $a_k = a_0 q^k$, из выражения (29) следуют оценки

$$|F_k(x)| \leq c_1 2^k (a_k)^{\frac{1}{3}} \leq c_2 2^k q^{\frac{k}{3}},$$

где c_1, c_2 – положительные постоянные. Следовательно, последовательность $F_k(x)$ будет стремиться к нулю для $0 < q < 2^{-3}$.

2. Считаем, что выполняется условие Липшица второго порядка (13) для функции $f(x)$ в точке $x \in [A, B)$. Используя функции $\psi_k^\alpha(x)$ и $\psi_k^\beta(x)$, которые определены формулами (25) и (26), получим неравенство по аналогии с нахождением соотношения (28):

$$\begin{aligned} & \left| \Delta^{k+1} f(a_0 u_0, \dots, a_k u_k, x) \psi(u_k) \psi(u_{k-1}) \right| \leq \\ & \leq 2^{k-1} L \left(a_k M_k \psi_k^\alpha(u_k) + \frac{d(B-A)}{M_k^2} \right) \left(a_{k-1} M_{k-1} \psi_{k-1}^\alpha(u_{k-1}) + \frac{d(B-A)}{M_{k-1}^2} \right). \end{aligned}$$

Для остаточного члена разложения (19) выполняется условие, верность которого можно подтвердить по аналогии с доказательством неравенства (29):

$$|F_{k+1}(x)| \leq 2^{k-1} L \left(a_k M_k + \frac{d(B-A)^2}{a_k C_0 M_k^2} \right) \left(a_{k-1} M_{k-1} + \frac{d(B-A)^2}{a_{k-1} C_0 M_{k-1}^2} \right). \quad (30)$$

Выбрав в выражении (30) значения $M_k = (a_k)^{-2/3}$, $M_{k-1} = (a_{k-1})^{-2/3}$ и $a_k = a_0 q^k$, получим оценку для остаточного члена

$$|F_{k+1}(x)| \leq c_1 2^k (a_k)^{\frac{2}{3}} \leq c_2 2^k q^{\frac{2k}{3}},$$

где c_1, c_2 – положительные постоянные. Следовательно, для $q < 2^{-3/2}$ разложение функции $f(x)$ в ряд сходится.

3. Пусть выполняется условие Липшица (14) порядка m для функции $f(x)$ в точке $x \in [A, B)$. Подобным образом можно показать, что для сходимости ряда достаточно выполнить условие

$$|F_{k+1}(x)| \leq c_1 2^k q^{\frac{m_k}{3}}, \quad (31)$$

где c_1 – положительная постоянная. Следовательно, остаточный член разложения будет стремиться к нулю для таких значений q , что $0 < q < 2^{-3/m}$. Случай продолжения функции $f(t)$ на числовую полуось $(-\infty, B]$ рассматривается аналогично. Теорема доказана.

Алгоритм дискретной аппроксимации функции. Рассмотрим алгоритм дискретной аппроксимации функции, для которой не выполняются условия Липшица. Пусть $x_i, i = 1, \dots, n$, принадлежат интервалу $[-1, 1]$ и известны значения функции $y_i = f(x_i)$ в этих точках.

1. Присваиваем начальные значения y_i коэффициентам вейвлета нулевого порядка: $F_{0,i} = y_i, i = 1, \dots, n$.

2. Вычисляем коэффициенты вейвлет-преобразования, используя дискретный аналог формул (6) и (7):

$$F_{k+1,i} = F_{k,i} - WF_k(x_i, a_k), \quad (32)$$

где $k = 1, \dots, K$; $F_{k,i}$ – значения коэффициентов вейвлета k -го порядка в точке x_i , $i = 1, \dots, n$; $a_k = \alpha 2^{-k}$, α – постоянная. Здесь

$$WF_k(x, a_k) = \frac{\sum_i F_{k,i} \Psi\left(\frac{x_i - x}{a_k}\right)}{\sum_i \Psi\left(\frac{x_i - x}{a_k}\right)}. \quad (33)$$

3. Восстанавливаем функцию $f_K(x) \approx f(x)$ во всех точках интервала $[A, B]$, используя аналог формулы (8):

$$f_K(x) = \sum_{k=0}^K WF_k(x, a_k). \quad (34)$$

Например, в частном случае $\psi(t) = \frac{1}{1+t^2}$, $x_i \in [-1, 1]$, $y_i = f(x_i)$, где $f(x) = \begin{cases} 1, & \text{если } x < 0; \\ 0, & \text{если } x = 0; \\ -1, & \text{если } x > 0. \end{cases}$

Значения коэффициентов вейвлет-преобразования $W_{k,i}$, найденные по формуле (33), представлены в таблице. Номеру строки m , $m = 1, 2, \dots, 6$, соответствует номер преобразования $k = m - 1$. Столбец с номером $i + 1$, $i = 0, 1, 2, \dots, 8$, соответствует координате точки $x_i = -1 + 2i/n$, $n = 16$.

Коэффициенты вейвлет-преобразования $W_{k,i}$

-1,00	-1,00	-1,00	-1,00	0,00	1,00	1,00	1,00	1,00
-0,30	-0,34	-0,46	-0,68	0,00	0,68	0,46	0,34	0,30
-0,00	-0,02	-0,11	-0,38	0,00	0,38	0,11	0,02	0,00
0,01	0,02	0,00	-0,15	0,00	0,15	-0,00	-0,02	-0,01
0,00	0,00	0,01	-0,02	0,00	0,02	-0,01	-0,00	-0,00
-0,00	0,00	0,00	-0,00	0,00	0,00	-0,00	-0,00	0,00

Значения аппроксимирующей функции были рассчитаны по формуле (34). На рис. 1 показаны графики функций $f_k(x)$, $x \in [-1, 1]$, для $K = 4$ и $K = 6$ и $y = f(x)$, а также точки x_i , в которых заданы значения функции y_i , $i = 0, 1, 2, \dots, 8$. На рис. 2 изображен график аппроксимации функции $f(x)$ полиномами степени $n = 8$ и $n = 12$, коэффициенты которых найдены по методу наименьших квадратов.

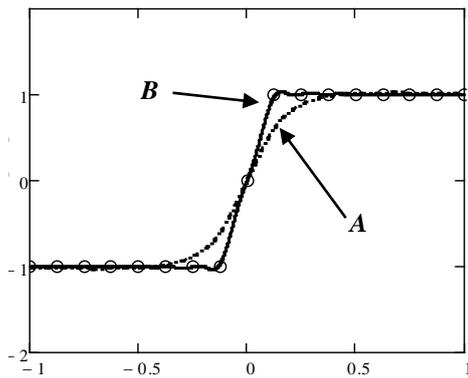


Рис. 1. Аппроксимация разрывной функции вейвлет-преобразованиями:
A – функция $f_4(x)$, B – функция $f_{12}(x)$

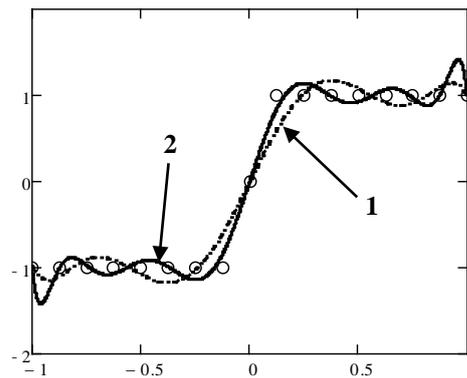


Рис. 2. Аппроксимация разрывной функции полиномом по методу наименьших квадратов:
1 – степень полинома, равная 5; 2 – степень полинома, равная 11

Аппроксимацию функцией $f_3(x)$ можно интерпретировать как результат сглаживания данных, аппроксимацию $f_6(x)$ – как интерполяцию (квазиинтерполяцию). Применяя различные порядки аппроксимации K , можно получить разные степени сглаживания функции, которые можно использовать в задаче фильтрации сигнала. Для этого функцию $f_k(x)$ удобно представить в виде суммы двух компонент:

$$f_k(x) = S(x) + T(x), \quad (35)$$

где $S(x) = \sum_{k=0}^m WF_k(x, a_k)$, $T(x) = \sum_{k=m}^K WF_k(x, a_k)$; преобразование $WF_k(x, a_k)$ определено по формуле (32). Таким образом, выполняется равенство $f(x) = S(x) + T(x) + F_{K+1}(x)$. Остаток $F_{K+1}(x)$ в узлах x_i можно найти по формуле

$$F_{K+1}(x_i) = f(x_i) - S(x_i) - T(x_i).$$

Параметры вейвлет-разложения (35) можно выбрать так, чтобы компонента $S(x)$ изменялась медленно, а компонента $T(x)$ – быстро и остаток был достаточно мал.

Пример приложения сингулярных вейвлетов. Исследуем временной ряд рождаемости в Республике Беларусь в период с 1950 по 2019 г. Пусть x_i – год, для которого определен уровень рождаемости, $x_i = 1950, 1951, \dots, 2019$; y_i – уровень рождаемости за год x_i . Аппроксимируем дискретный ряд наблюдений y_1, y_2, \dots, y_n непрерывной функцией $f(x)$, такой, что $f(x_i) = y_i$. Используя формулу (34), представим функцию $f_k(x)$ в виде суммы медленной $S(x)$ и быстрой $T(x)$ компонент. Параметры для вейвлет-разложения выбраны следующие: $m = 1$, $K = 5$, $a_k = 2^{-k}$; сингулярный вейвлет $\psi(t) = e^{-t^2}$. Результаты расчетов представлены на рис. 3.

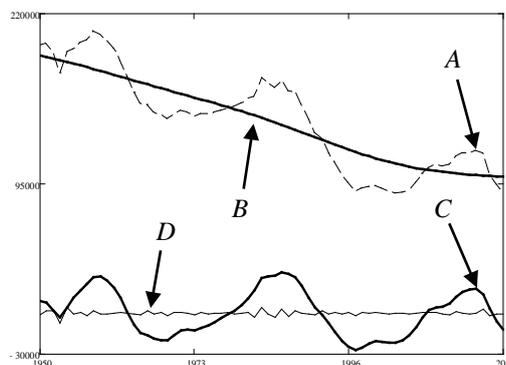


Рис. 3. Схема выделения из сигнала быстрой и медленной компонент:
A – график исходных данных, y_i ; B – график медленной компоненты, $S(x)$;
C – график быстрой компоненты, $T(x)$; D – остаток

Проведем спектральный анализ Фурье исходных данных и быстрой компоненты.

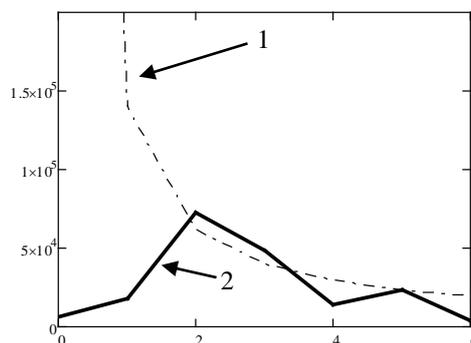


Рис. 4. Амплитудный спектр сигнала:

1 – спектр исходных данных; 2 – спектр быстрой компоненты

В результате спектрального анализа исходных данных (рис. 4) не была обнаружена периодическая аддитивная компонента. Спектр быстрой компоненты указывает на наличие периодической составляющей с периодом около 35 лет.

Графики на рис. 4. можно использовать как иллюстрацию влияния различных факторов на демографические показатели. В частности, специалистов в области демографии может заинтересовать возможность непосредственного анализа графиков.

В качестве примера приведем вариант трактовки графиков. Показатель рождаемости можно представить в виде суммы медленной и быстрой компонент. Обе компоненты имеют естественную интерпретацию. Первая обусловлена воздействием факторов, которые медленно изменяют показатели рождаемости; вторая – наличием факторов, оказывающих влияние на быстрое изменение рождаемости. Быстрая компонента, судя по графику на рис. 4, связана с военными действиями на территории страны (стрессовой ситуацией, приведшей к всплеску рождаемости в послевоенный период). На основании данных спектрального анализа можно сделать вывод, что воздействие этого всплеска рождаемости повторяется каждые 35 лет, хотя и становится менее выраженным.

Возможен следующий прогноз на период начиная с 2020 г.: значения медленной и быстрой компонент в ближайшие 10 лет будут убывать, поэтому показатель рождаемости будет сокращаться. Это обусловлено совпадением результатов воздействия двух групп факторов. Вместе с тем скорость спада медленной компоненты в последние 10 лет несколько замедлилась, что свидетельствует о появлении новых тенденций, которые оказывают влияние на медленную компоненту.

Заключение. Исследован ряд вейвлет-преобразований с сингулярным вейвлетом, для которого условие допустимости не выполняется. Показано, что вейвлет-преобразования можно использовать для аппроксимации функциональных зависимостей. Сформулированы и доказаны достаточные условия сходимости ряда вейвлет-преобразований, проведено исследование реального временного ряда. Показано, что если функция $f(x)$ на промежутке $[A, B]$ принадлежит пространству $L^1[A, B]$, то ее можно представить в виде суммы вейвлет-преобразований и остаточного члена. Для остаточного члена должно выполняться условие (31). Из условия сходимости следует, что скорость сходимости тем выше, чем выше порядок условия Липшица для функции $f(x)$.

Приведен пример аппроксимации с помощью суммы дискретных сингулярных вейвлет-преобразований. Выполнен сравнительный анализ аппроксимации рядом, составленным с помощью сингулярных вейвлетов, и методом наименьших квадратов. В качестве функции $f(x)$ на промежутке $[A, B]$ выбрана разрывная функция. Для разрывной функции условия Липшица не выполняются, поэтому не выполняются и доказанные в работе условия сходимости. Тем не менее сумма дискретных вейвлет-преобразований с сингулярным вейвлетом позволяет построить аппроксимацию функции $f(x)$, максимальное отклонение которой от точного значения существен-

но ниже, чем у приближения, полученного методом наименьших квадратов. В качестве примера приложения исследован временной ряд рождаемости в Республике Беларусь в период с 1950 по 2019 г. Для этого функция представлена в виде суммы двух компонент и остаточного члена формулы (35). Параметры вейвлет-разложения можно выбрать так, чтобы одна компонента изменялась медленно, а вторая – быстро и остаток вейвлет-разложения был достаточно мал. В этом случае можно провести спектральный анализ для каждой компоненты, причем результаты спектрального анализа каждой компоненты допускают вполне ясную интерпретацию.

Существует распространенная точка зрения, которая нашла свое отражение даже в учебной литературе, что вейвлет-преобразования обладают практически всеми достоинствами преобразований Фурье. Основанием для таких выводов является тот факт, что вейвлеты Морле [2] можно использовать для анализа сигнала, в частности периодического сигнала. Такой анализ позволяет судить о наличии периодической компоненты, однако в целом дает искаженное представление о спектре сигнала. Это вдвойне опасно, так как обработку экспериментальных данных часто проводят специалисты, которые далеки от предметной области, в то время как выводы делают исследователи, хорошо разбирающиеся в предметной области, но некритично воспринимающие результаты анализа сигнала. В рассмотренном случае результаты спектрального анализа быстрой и медленной компонент свободны от искажений. Представленный в работе пример позволяет сделать вывод о том, что сочетание вейвлет-анализа с анализом Фурье дает возможность получить объективную характеристику сигнала.

Список использованных источников

1. Хардле, В. Прикладная непараметрическая регрессия : пер. с англ. / В. Хардле. – М. : Мир, 1993. – 349 с.
2. Чуи, К. Введение в вейвлеты : пер. с англ. / К. Чуи. – М. : Мир, 2001. – 412 с.
3. Добеши, И. Десять лекций по вейвлетам : пер. с англ. / И. Добеши. – Ижевск : Регулярная и хаотическая динамика, 2001. – 464 с.
4. Watson, G. S. Smooth regression analysis / G. S. Watson // *Sankhya: The Indian Journal of Statistics, Ser. A.* –1964. – Vol. 26. – P. 359–372.
5. Parzen, E. On estimation of a probability density function and mode / E. Parzen // *The Annals of Mathematical Statistics.* – 1962. – Vol. 33, no. 3. – P. 1065–1076.
6. Надарая, Э. А. Об оценке регрессии / Э. А. Надарая // *Теория вероятностей и ее применение.* – 1964. – Т. 9, № 1. – С. 157–159.
7. Серенков, П. С. Система сбора данных о качестве как техническая основа функционирования эффективных систем менеджмента качества / П. С. Серенков, В. М. Романчак, В. Л. Соломахо // *Докл. Нац. акад. наук Беларуси.* – 2006. – Т. 50, № 4. – С. 100–104.
8. Романчак, В. М. Аппроксимация экспертных оценок сингулярными вейвлетами / В. М. Романчак, П. М. Лаппо // *Вестник Гродненского гос. ун-та. Сер. 2. Математика. Физика. Информатика, вычислительная техника и управление.* – 2017. – Т. 7, № 1. – С. 132–139.
9. Романчак, В. М. Аппроксимация сингулярными вейвлетами / В. М. Романчак // *Системный анализ и прикладная информатика.* – 2018. – № 2. – С. 23–28.
10. Романчак, В. М. Сингулярные вейвлеты на конечном интервале / В. М. Романчак // *Информатика.* – 2018. – Т. 15, № 4. – С. 39–49.

References

1. Härdle W. *Applied Nonparametric Regression.* Cambridge, Cambridge University Press, 1992, 434 p.
2. Chui C. *An Introduction to Wavelets.* San Diego, Academic Press, 1992, 266 p.
3. Daubechies I. *Ten Lectures on Wavelets.* Philadelphia, Society for Industrial and Applied Mathematics, 1992, 377 p.
4. Watson G. S. Smooth regression analysis. *Sankhya: The Indian Journal of Statistics, Ser. A,* 1964, vol. 26, pp. 359–372.
5. Parzen E. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics,* 1962, vol. 33, no. 3, pp. 1065–1076.

6. Nadaraya E. A. Ob ocenke regressii [About a regression assessment]. Teorija verojatnostej i ee primenenie [*Probability Theory and Its Application*], 1964, vol. 9, no. 1, pp. 157–159 (in Russian).

7. Serenkov P. S., Romanchak V. M., Solomakho V. L. Sistema sbora dannyh o kachestve kak tehničeskaja osnova funkcionirovanija jeffektivnyh sistem menedzhmenta kachestva [System of collection of data on quality as technical basis of functioning of effective systems of quality management]. Doklady Nacional'noj akademii nauk Belarusi [*Doklady of the National Academy of Sciences of Belarus*], 2006, vol. 50, no. 4, pp. 100–104 (in Russian).

8. Romanchak V. M., Lappo P. M. Approksimacija jekspertnyh ocenok singuljarnymi vejvletami [Approximation of expert estimates by singular wavelets]. Vestnik Grodnenskogo gosudarstvennogo universiteta. Ser. 2. Matematika. Fizika. Informatika, vychislitel'naja tehnika i upravlenie [*Bulletin of the Grodno State University. Series 2: Mathematics. Physics. Informatics, Computer Science and Management*], 2017, vol. 7, no. 1, pp. 132–139 (in Russian).

9. Romanchak V. M. Approksimacija singuljarnymi vejvletami [Approximation by singular wavelets]. Sistemnyj analiz i prikladnaja informatika [*Systems Analysis and Applied Informatics*], 2018, no. 2, pp. 23–28 (in Russian).

10. Romanchak V. M. Singuljarnye vejvlety na konečnom intervale [Singular wavelets on a finite interval]. Informatika [*Informatics*], 2018, vol. 15, no. 4, pp. 39–49 (in Russian).

Информация об авторе

Романчак Василий Михайлович, кандидат технических наук, доцент кафедры инженерной математики, Белорусский национальный технический университет, Минск, Беларусь.
E-mail: Romanchak@bntu.by

Information about the author

Vasily M. Romanchak, Cand. Sci. (Eng.), Associate Professor of the Department of Engineering Mathematics, Belarusian National Technical University, Minsk, Belarus.
E-mail: Romanchak@bntu.by

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

УДК 004.421.5 + 517.938
<https://doi.org/10.37661/1816-0301-2020-17-4-36-47>

Поступила в редакцию 06.11.2020
Received 06.11.2020

Принята к публикации 23.11.2020
Accepted 23.11.2020

Цифровая модель генератора псевдослучайных чисел на основе непрерывной хаотической системы

Е. А. Дрыбин, С. В. Садов, В. С. Садов✉

Белорусский государственный университет, Минск, Беларусь
✉E-mail: sadov@bsu.by

Аннотация. Показано, что выбор параметра временной дискретизации цифровой модели непрерывной динамической системы с хаотическими режимами на основе ее динамики позволяет управлять характеристиками выходной последовательности, в том числе избегать коротких циклов и периодических режимов поведения. На примере системы Лоренца проведен анализ закона движения хаотической системы, линеаризованной в окрестностях точек устойчивого и неустойчивого равновесия. На основании этого закона выбраны параметры математической модели генератора псевдослучайных чисел. Выходная последовательность чисел, порождаемая предложенным в работе подходом, подвергнута статистическому и корреляционному анализу. Согласно результатам проведенных тестов полученные псевдослучайные последовательности на основе непрерывных хаотических систем обладают статистически случайными свойствами и могут быть использованы в системах стеганографической и криптографической защиты данных.

Ключевые слова: временная дискретизация, детерминированный хаос, система Лоренца, случайные последовательности, цифровая модель

Для цитирования. Дрыбин, Е. А. Цифровая модель генератора псевдослучайных чисел на основе непрерывной хаотической системы / Е. А. Дрыбин, С. В. Садов, В. С. Садов // Информатика. – 2020. – Т. 17, № 4. – С. 36–47. <https://doi.org/10.37661/1816-0301-2020-17-4-36-47>

Digital model of a pseudo-random number generator based on a continuous chaotic system

Yaugen A. Drybin, Siarhei V. Sadau, Vasili S. Sadau ✉

Belarussian State University, Minsk, Belarus
✉E-mail: sadov@bsu.by

Abstract. It is shown that the choice of the time sampling parameter of the digital model of a continuous dynamic system with chaotic modes based on its dynamics makes it possible to control the characteristics of the output sequence, including avoiding short cycles and periodic behavior modes. On the example of the Lorentz system, the analysis of the law of motion of a chaotic system, linearized in the vicinity of points of stable and unstable equilibrium, is carried out, on the basis of which the parameters of the mathematical model of the generator of pseudo-random numbers are selected. The output sequence of numbers generated in proposed way is subjected to statistical and correlation analysis. Based on the results of the tests carried out, we can say that the obtained pseudo-random sequences based on continuous chaotic systems have statistically random properties and can be used in steganographic and cryptographic systems.

Keywords: time sampling, deterministic chaos, Lorentz system, random sequences, digital model

For citation. Drybin Y. A., Sadau S. V., Sadau V. S. Digital model of a pseudo-random number generator based on a continuous chaotic system. *Informatics*, 2020, vol. 17, no. 4, pp. 36–47 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-36-47>

Введение. В практических задачах криптографии и стеганографии генерирование случайных данных осуществляется с помощью генераторов псевдослучайных чисел (ГПСЧ), которые на входе используют короткий ключ (семя), а на выходе предоставляют псевдослучайную последовательность чисел (ПСП), имеющую закон распределения, мало отличающийся от равномерного распределения. В настоящей статье предлагается вариант получения ПСП с использованием хаотических систем, которые во многом схожи с криптографическими алгоритмами, поскольку и в одних, и в других системах осуществляется нелинейное преобразование информации. При этом данное нелинейное преобразование, с одной стороны, является детерминированным процессом, а с другой – должно быть практически непредсказуемо для стороннего наблюдателя. Такая особенность была отмечена К. Шенноном в работах, написанных еще до обнаружения явления детерминированного хаоса, где он предлагает перемешивающие преобразования, зависящие от аргумента, и описывает базовый механизм образования детерминированного хаоса путем растяжения и складывания [1].

Методы формирования хаотических последовательностей. Схемы традиционных ГПСЧ реализуются на основе целочисленной арифметики, но в настоящее время достаточно широкое распространение получили и генераторы, использующие арифметику с плавающей запятой. Арифметика с плавающей запятой (АПЗ) дает возможность применять в качестве ГПСЧ нелинейные динамические системы с хаотическими процессами и рекуррентные отображения, поведение которых во многом схоже с нелинейными динамическими системами.

В качестве ГПСЧ чаще всего предлагается использовать дискретные отображения, такие как отображения Хенона, кусочно-линейные отображения, отображение (сдвиг) Бернулли и т. д. Программные реализации подобных ГПСЧ в настоящее время встречаются в различных пакетах прикладных программ для математического моделирования.

Основными недостатками использования дискретных отображений для получения требуемых ПСП, которые делают их неэффективными в качестве источников псевдослучайного сигнала, являются [2, 3] периодическое поведение, проявляющееся при использовании АПЗ и обусловленное невозможностью представления иррационального числа в двоичной записи, а также наличие коротких циклов, состоящих из крайне ограниченного числа возможных состояний системы, при определенных начальных условиях. Это не позволяет эффективно использовать дискретные отображения для получения ПСП, так как при формировании набора входных параметров (являющегося семенем), как правило, заранее неизвестно, какой орбите и с какой длиной периода он соответствует. Для преодоления явления возникновения коротких орбит в хаотических системах, реализованных программно с использованием АПЗ, в последнее время предлагаются две группы методов:

- 1) увеличение размерности хаотических систем [4] или использование более сложных хаотических систем [5, 6];
- 2) нарушение итерационного процесса хаотической системы с использованием другой хаотической системы с малой размерностью [7].

Отметим, что первая группа методов лишь замедляет скорость деградации хаотической системы (явление возникновения коротких орбит проявляется, но с большим числом итераций), однако не предоставляет фундаментального решения данной проблемы. Вторая группа методов с поэтапным нарушением итерационного процесса хаотической системы позволяет уменьшить погрешность вычислений с использованием АПЗ и предотвратить деградацию хаотической системы, но выходная последовательность (и длина соответствующей ей периодической орбиты) первичной хаотической системы будет определяться хаотическим отображением с малой размерностью, которая используется для нарушения итерационного процесса.

Хаотические системы с непрерывным временем. С целью преодоления ограничений дискретных хаотических отображений для формирования ПСП предлагается использовать многомерные нелинейные динамические системы с непрерывным временем, в которых возможно существование принципиально отличных структурных множеств, называемых аттракторами. Аттрактор представляет собой множество траекторий в фазовом пространстве, к которым притягиваются все траектории из некоторой окрестности аттрактора. Возникновение аттракторов возможно только в диссипативных динамических системах, а простейшими примерами аттракторов могут служить неподвижные точки или замкнутые траектории (орбиты).

В трехмерных диссипативных нелинейных системах (в общем случае в системах с размерностью больше двух) возможно существование сложного режима колебаний, называемого странным аттрактором [8]. Спектр колебаний системы в режиме странного аттрактора очень близок к спектру случайного процесса. Отметим также, что дискретные отображения, рассмотренные выше, порождают процессы, качественно схожие с процессами в нелинейных динамических системах.

При программном формировании хаотических сигналов или с использованием цифровых сигнальных процессоров неизбежно происходит дискретизация временной переменной, а для моделирования хаотических процессов применяются методы численного интегрирования систем дифференциальных уравнений. При дискретизации временного параметра непрерывных систем математические модели соответствующих генераторов сводятся к итерационной функции вида

$$x_{n+1} = f(x_n, K), \quad (1)$$

где K – некоторый управляющий параметр.

Таким образом, при цифровой обработке выходной сигнал, порождаемый непрерывной хаотической системой, во многом становится эквивалентным сигналу дискретной системы. Однако в отличие от характеристик дискретного сигнала характеристики выходного сигнала могут управляться посредством изменения параметра K , который фактически является параметром временной дискретизации системы нелинейных дифференциальных уравнений.

Кроме того, цифровая реализация непрерывной хаотической системы для представления значений чисел, как правило, использует АПЗ.

Действительное число x может быть записано как бесконечная десятичная дробь в двоичном представлении $b_m b_{m-1} \dots b_1 \cdot a_1 a_2 \dots a_n$, где a_i, b_j – биты, $b_m b_{m-1} \dots b_1$ соответствует целой части числа, а $a_1 a_2 \dots a_n$ – дробной части.

При вычислениях с конечной точностью итерационную функцию $x_{n+1} = f(x_n)$ можно представить в виде

$$x_{n+1} = \text{round}_k(f(x_n)), \quad (2)$$

где $\text{round}_k(x)$ – функция округления, которая может быть задана следующим выражением:

$$\text{round}_k(x) = b_m b_{m-1} \dots b_1 \cdot a_1 a_2 \dots a_{k-1} (a_k + a_{k+1}). \quad (3)$$

Одной из существенных проблем программной реализации непрерывных динамических систем является накопление ошибки округления. Функция $\text{round}_k(x)$ применяется на каждой итерации, и ошибка округления может накапливаться. Траектории движения исходной и аппроксимированной систем могут расходиться очень быстро благодаря чувствительности хаотической системы к начальным условиям. Таким образом, любая математическая модель, основанная на АПЗ, не является абсолютно достоверной реализацией непрерывной хаотической системы.

Помимо некорректного асимптотического поведения, проявляющегося по истечении относительно большого отрезка времени, аппроксимированные системы обладают рядом «опасных» свойств, которые могут проявиться уже в начале траектории. Одной из возможностей их минимизации является округление состояния системы таким образом, что траектория ее движения безвозвратно покидает странный аттрактор и переходит к периодической моде. Например, во многих нелинейных системах переменные, определяющие режим возникающих колебаний, могут иметь бесконечно малые, но не равные нулю значения. В случае «неудачного» округления такой переменной до нуля хаотическое поведение системы прекращается и она переходит в нехаотический режим. Поэтому при реализации хаотической системы с использованием АПЗ важно знать закон ее движения, что позволит ограничить используемые в модели области фазового пространства и избежать нежелательных эффектов.

Характеристики хаотических сигналов, полученных с использованием цифровой реализации непрерывной динамической системы, зависят от величины шага временной дискретизации. Выбор и оценку шага временной дискретизации Δt с целью получения требуемых характеристик выходной последовательности необходимо проводить при помощи параметра дискретизации K , определяемого следующим образом [9]:

$$K = \frac{T}{\Delta t}, \quad (4)$$

где T – период квазирезонансных колебаний динамической системы.

Уменьшение величины K приводит к изменению статистических характеристик сигналов в нелинейных системах с детерминированным хаосом, но за счет уменьшения количества отсчетов на период колебаний становится возможным улучшить эффективность (быстродействие) цифровых генераторов хаотических последовательностей. Поэтому при разработке цифровых генераторов хаоса необходимо оценить предельные значения параметра K и выработать рекомендации по выбору минимального значения параметра дискретизации K_{\min} . При значениях $K < K_{\min}$ в процессе численного интегрирования может происходить переопределение переменных координат точки в фазовом пространстве, из-за чего траектория движения системы может покинуть странный аттрактор.

Построение ГПСЧ на основе непрерывной хаотической системы. В качестве нелинейной динамической системы, которая ляжет в основу ГПСЧ, будем использовать систему Лоренца

$$\begin{aligned} \dot{X} &= -\sigma X + \sigma Y, \\ \dot{Y} &= -XZ + rX - Y, \\ \dot{Z} &= XY - bZ, \end{aligned} \quad (5)$$

где X, Y, Z – переменные системы; r, σ, b – параметры системы Лоренца.

Данный выбор основан на том, что система Лоренца обладает малым количеством параметров, проста и надежна в схемотехнической реализации и в то же время характеризуется максимальной энтропией Колмогорова ($K > 2$). Это указывает на сложность ее внутреннего устройства и протекающих в ней процессов в сравнении как с дискретными отображениями, так и непрерывными хаотическими системами Чуа и Ресслера. Динамика системы Лоренца детально исследована и описана в литературе, что позволяет легко определять диапазоны допустимых параметров и области фазового пространства, в которых система ведет себя хаотически. Кроме того, система Лоренца имеет наименьшее количество областей с регулярным поведением [10], а существование коротких орбит в хаотических режимах не описано в литературе (предполагается их возможное существование в моделях с использованием АПЗ). Знание закона движения системы позволит правильно выбрать рабочие диапазоны параметров для ее программной или аппаратной моделей.

Исследуем динамику системы Лоренца при фиксированных значениях параметров $b = 8/3$ и $\sigma = 10$ и при изменении параметра r , который обычно называют управляющим.

При значениях $r > 1$ аттрактор в системе Лоренца, представляющий собой неподвижную точку в начале координат, становится неустойчивым и появляются два новых состояния равновесия:

$$\begin{aligned} X_{01} &= \pm\sqrt{b(r-1)}, \\ Y_{01} &= \pm\sqrt{b(r-1)}, \\ Z_{01} &= r-1. \end{aligned} \quad (6)$$

Существование странного аттрактора в системе Лоренца возможно, если $r > 24,06$. При этом существование устойчивых точек, соответствующих стационарному режиму, сохраняется до значения

$$r = \frac{\sigma(\sigma + b + 3)}{\sigma - b - 1} = 24,74, \quad (7)$$

после чего странный аттрактор становится единственным притягивающим множеством в фазовом пространстве системы.

Процесс перехода системы Лоренца к странному аттрактору сопровождается квазипериодическим движением по раскручивающейся спирали увеличивающегося радиуса. При малых отклонениях от состояния равновесия период квазирезонансных колебаний в системе Лоренца можно оценить с помощью выражения

$$T \approx 4\pi[8\sigma(r - 1) - (\sigma + r)^2]^{-1/2}. \quad (8)$$

Проведем аналитическое определение диапазона значений параметра временной дискретизации K для системы Лоренца, исследуя динамику решения системы дифференциальных уравнений (5), линеаризованной в окрестностях точек устойчивого и неустойчивого равновесия.

Фазовая траектория нелинейной динамической системы в хаотическом режиме может быть разбита на отдельные характерные участки для каждой из мод в этом режиме. Спектральные характеристики колебаний, порождаемых системой при движении по данному участку траектории в фазовом пространстве, позволят оценить верхнюю граничную частоту сигналов. Воспользовавшись теоремой отсчетов, можно оценить и максимальное значение времени следования отсчетов.

В фазовых траекториях системы Лоренца можно отметить характерные участки в окрестностях точки $C_0 = (0, 0, (r - 1))$ и двух точек неустойчивого равновесия системы $C_{1,2}$, координаты которых определяются выражениями (6).

В окрестностях точки $C_0 = (0, 0, (r - 1))$ при выполнении соотношения $XY \ll Z$ систему (5) можно привести к следующему виду:

$$\ddot{X} + \dot{X} \left(1 + \sigma + \frac{X^2}{b} \right) - X\sigma(r - 1) + \frac{X^3\sigma}{b} = \frac{XZ\sigma}{b}. \quad (9)$$

Пренебрегая малыми членами, приведем соотношение (9) к линейному дифференциальному уравнению

$$\ddot{X} + \dot{X}(1 + \sigma) - X\sigma(r - 1 - Z(t)) \approx 0, \quad (10)$$

где $Z(t) \approx Z_0 e^{-bt}$.

При условии $Z(t) > (r - 1)$ выражение (10) можно привести к виду

$$T\ddot{X} + \dot{X}(1 + \sigma) + X\sigma(r - 1) \approx 0, \quad (11)$$

при $Z(t) < (r - 1)$ – к виду

$$\ddot{X} + \dot{X}(1 + \sigma) - X\sigma(r - 1) \approx 0. \quad (12)$$

Рассмотрим решение уравнения в окрестностях точки C_0 в обоих случаях. В случае $Z(t) > (r - 1)$ решение можно записать следующим образом:

$$\begin{aligned}
 X_0(t) &= e^{a_0 t} \sin \omega_0 t, \\
 a_0 &= \frac{-(1 + \sigma)}{2}, \\
 \omega_0 &= \frac{1}{2} [4\sigma(r - 1) - (1 + \sigma)^2].
 \end{aligned} \tag{13}$$

С помощью преобразования Фурье получим выражение для спектра сигнала $X_0(t)$:

$$\begin{aligned}
 X_0(\omega) &= \frac{(-\omega_0 \cos T\omega_0 + a_0 \sin T\omega_0 - j\omega \sin T\omega_0)}{(a_0 - j\omega)^2 + \omega_0^2} \cdot e^{T(a_0 - j\omega)} + \\
 &+ \frac{\omega_0}{(a_0 - j\omega)^2 + \omega_0^2},
 \end{aligned} \tag{14}$$

где длительность сигнала $X_0(t)$ выбрана исходя из $T \geq 20\pi/\omega_0$.

В случае $Z(t) < (r - 1)$ решение можно записать в виде

$$\begin{aligned}
 X_0(t) &= e^{a_0 t}, \\
 a_0(t) &= \frac{-(1 + \sigma) - \sqrt{4\sigma(r - 1) + (1 + \sigma)^2}}{2}.
 \end{aligned} \tag{15}$$

При этом спектр сигнала определяется равенством

$$X_0(\omega) = -\frac{1}{a_0 + j\omega} \cdot e^{T(a_0 + j\omega)} + \frac{1}{a_0 + j\omega}. \tag{16}$$

В окрестностях точек $C_{1,2}$ решение можно записать формулой

$$X_{1,2}(t) = e^{a_{1,2} t} \sin \omega_{1,2} t. \tag{17}$$

Спектр сигнала на данном участке с учетом $\omega_{1,2} = 2\pi/T$ примет следующий вид:

$$\begin{aligned}
 X_{1,2}(\omega) &= \frac{(-\omega_{1,2} \cos T\omega_{1,2} + a_{1,2} \sin T\omega_{1,2} - j\omega \sin T\omega_{1,2})}{(a_{1,2} - j\omega)^2 + \omega_{1,2}^2} \cdot e^{T(a_{1,2} - j\omega)} + \\
 &+ \frac{\omega_{1,2}}{(a_{1,2} - j\omega)^2 + \omega_{1,2}^2}.
 \end{aligned} \tag{18}$$

Спектры сигнала решений линеаризованной системы Лоренца в окрестностях состояний равновесия при значениях параметров $r = 28$, $\sigma = 10$ и $b = 8/3$ показаны на рис. 1.

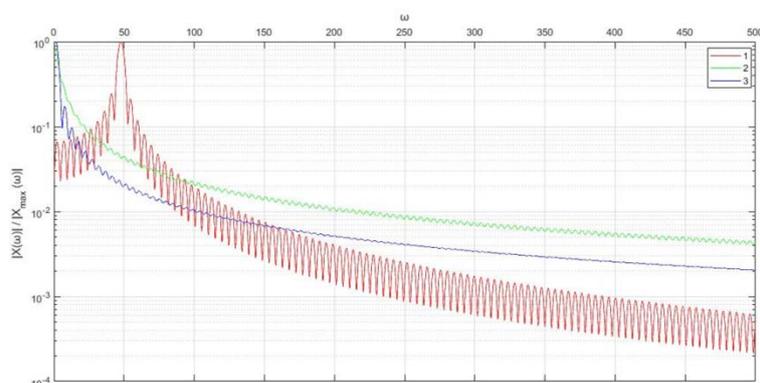


Рис. 1. Спектры сигнала решений линеаризованной системы Лоренца: 1 – рассчитанные согласно выражению (14); 2 – выражению (16); 3 – выражению (18)

Согласно зависимостям, изображенным на рис. 1, определим по уровню 0,01 верхнюю граничную частоту сигнала $\omega_{\max} = 223,28$, что соответствует минимальному значению параметра временной дискретизации:

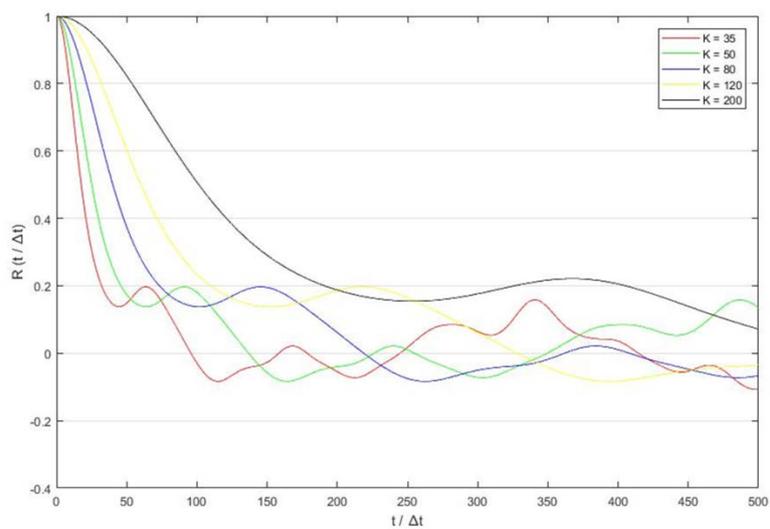
$$K_{\min} = \frac{T\omega_{\max}}{\pi} = 33,38, \quad (19)$$

где T – период квазирезонансных колебаний, который в системе Лоренца определяется согласно выражению (8).

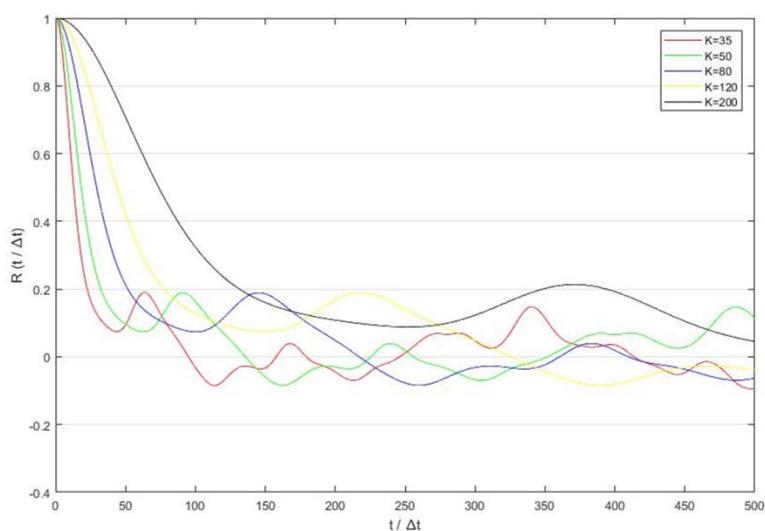
Таким образом, минимальный параметр дискретизации при численном решении системы Лоренца должен быть $K > K_{\min}$. Вместе с тем необходимо провести дополнительные экспериментальные исследования характеристик хаотического сигнала, порождаемого цифровой реализацией непрерывной нелинейной динамической системы, в зависимости от выбранного параметра K . Для этого целесообразно использовать автокорреляционную функцию. Спиральный хаос в автоколебательных системах имеет схожие свойства со свойствами зашумленных квазигармонических колебаний. При этом скорость расцепления корреляций в начале траектории зависит как от поведения мгновенной амплитуды, так и от фазы колебаний. По мере движения системы по фазовой траектории огибающая автокорреляционной функции в большей степени определяется диффузией мгновенной фазы [11].

Графики автокорреляционной функции для сигналов X , Y , Z , порождаемых смоделированной системой Лоренца при различных значениях параметра временной дискретизации K , показаны на рис. 2. Для математического моделирования и анализа хаотических процессов использовался пакет программ MATLAB. Для решения системы дифференциальных уравнений, описывающих хаотический процесс, был выбран численный метод Дорманда – Принса. Несмотря на то что данный метод относится к группе явных и, соответственно, может быть неустойчивым, он хорошо подходит для решения системы Лоренца и позволяет получить достаточно точное решение за короткое время [10]. На рис. 2 видно, что при меньших значениях параметра K автокорреляционная функция спадает быстрее и в дальнейшем держится ближе к нулевому значению. Зависимость падения автокорреляционной функции реализации Z отличается от аналогичных зависимостей для реализаций X и Y .

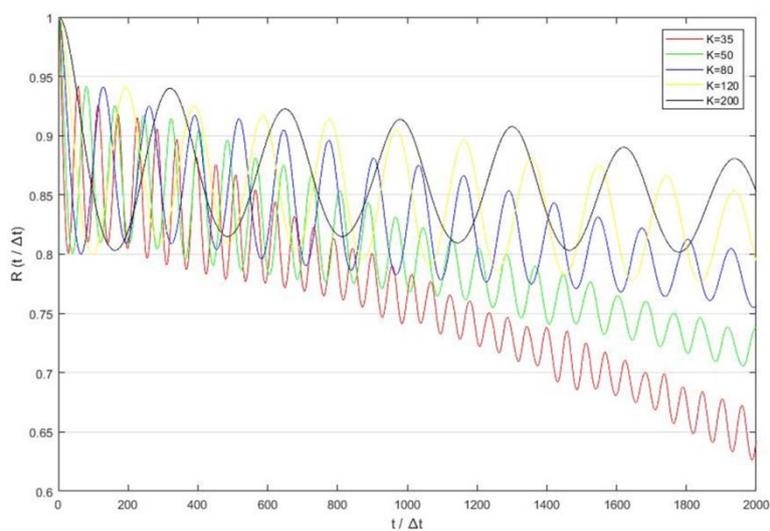
Одним из способов формирования импульсных псевдослучайных последовательностей на основе непрерывных нелинейных динамических систем является сравнение хаотического сигнала с пороговым уровнем. Оценить быстродействие генераторов такого типа можно по количеству пересечений порогового уровня хаотическим сигналом. Для системы Лоренца подобной мерой быстродействия может служить количество переходов фазовой траектории из одной области фазового пространства в другую область с отличными состояниями равновесия, которые определяются выражением (6). Зависимость количества переходов фазовой траектории между различными состояниями равновесия от параметра K в системе Лоренца при стандартных значениях параметров и неизменной продолжительности сигналов показана на рис. 3. На рисунке видно, что при уменьшении параметра временной дискретизации K (соответственно, при увеличении шага временной дискретизации и росте интенсивности шумов дискретизации) растет количество переходов фазовой траектории между различными состояниями равновесия при том же количестве временных отсчетов. Полученная зависимость также показывает, что количество переходов фазовой траектории между различными состояниями равновесия практически не изменяется при больших значениях параметра K . Кроме того, увеличение значения параметра временной дискретизации K влечет за собой и увеличение затрат на проведение численного интегрирования нелинейной системы. Следовательно, при малых значениях параметра временной дискретизации K улучшаются динамические характеристики хаотических сигналов, порождаемых системой Лоренца. Исходя из вышеизложенного, для построения программной реализации ГПСЧ на основе системы Лоренца целесообразно выбрать параметр дискретизации $K = 35$.



а)



б)



в)

Рис. 2. Автокорреляционная функция системы Лоренца для сигнала X (а), сигнала Y (б) и сигнала Z (в)

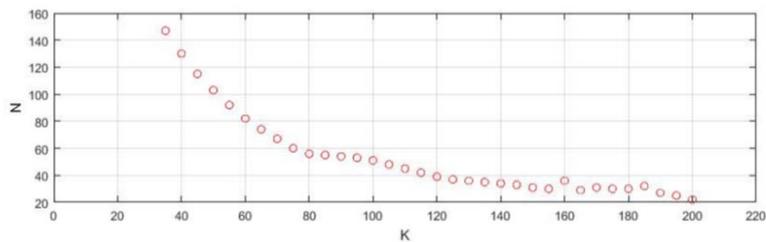


Рис. 3. Зависимость количества переходов фазовой траектории между различными состояниями равновесия от параметра K в системе Лоренца

Оценка качества выходной последовательности ГПСЧ. Для проверки статистических характеристик порождаемой двоичной последовательности используем набор статистических тестов NIST STS (NIST Statistical Test Suite), для проверки ПСП на непредсказуемость – Q-тест Льюнга – Бокса.

Для проведения тестирования построенного ГПСЧ сгенерируем $N = 1000$ двоичных последовательностей (при различных начальных условиях, заданных случайным образом так, что $X \in [1, 10]$, $Y \in [1, 10]$, $Z \in [1, 10]$) длиной $n = 1\,048\,576$ бит. Используем следующий простой и эффективный способ хеширования для отображения i -го 64-битного отсчета в i -й элемент двоичной последовательности, принадлежащий множеству $\{0,1\}$:

$$B_i = \text{mod} \left(\sum_{j=1}^{64} b_{ij}, 2 \right), \quad (20)$$

где b_{ij} – j -й бит i -го отсчета выходной последовательности сигнала X ; B_i – i -й элемент выходной двоичной последовательности ГПСЧ.

Согласно результатам тестирования NIST STS при уровне значимости $\alpha = 0,01$ (таблица) построенный ГПСЧ на основе сигнала X системы Лоренца успешно прошел статистическое тестирование пакетом NIST STS, а порождаемые им ПСП обладают статистическими свойствами случайных последовательностей.

Результаты статистических тестов NIST STS

Статистический тест	Рекомендуемые NIST параметры тестирования	Фактические параметры тестирования	Значение вероятности P (выборка)
Монобитный (частотный побитовый) тест	$n \geq 100$	$n = 1\,048\,576$	0,045 675
Частотный блочный тест	$n \geq 100$; $M = 10$	$n = 1\,048\,576$; $M = 10$	0,085 587
Тест пробегов	$n \geq 100$	$n = 1\,048\,576$	0,021 999
Тест на самую длинную последовательность единиц в блоке	$n \geq 6272$; $M = 128$	$n = 1\,048\,576$; $M = 128$	0,955 835
Тест рангов бинарных матриц	$n \geq 38\,912$ или $n \geq 38MQ$; $M = Q = 32$	$n = 1\,048\,576$; $M = Q = 32$	0,816 537
Спектральный тест на основе дискретного преобразования Фурье	$n \geq 1000$	$n = 1\,048\,576$	0,262 249
Тест на совпадение неперекрывающихся шаблонов	$n = 2^{20} = 1\,048\,576$; $m = 9$	$n = 1\,048\,576$; $m = 9$	0,162 606 0,534 146 0,494 392 0,678 686 0,616 305 0,419 021 0,191 687 0,759 756 0,275 709 ...

Продолжение таблицы

Статистический тест	Рекомендуемые NIST параметры тестирования	Фактические параметры тестирования	Значение вероятности P (выборка)
Тест на совпадение перекрывающихся шаблонов	$n \geq 1\,000\,000$; $m = 9$	$n = 1\,048\,576$; $m = 9$	0,162 606
Универсальный статистический тест Маурера	$n \geq 904\,960$; $L = 7$; $Q = 1280$	$n = 1\,048\,576$; $L = 7$; $Q = 1280$	0,657 933
Тест на линейную сложность	$n \geq 1\,000\,000$; $500 \leq M \leq 5000$	$n = 1\,048\,576$; $M = 1000$	0,964 295
Тест на подпоследовательности	$m < \log_2 n - 2$	$n = 1\,048\,576$; $m = 2$	0,016 717 0,017 912
Тест приближительной энтропии	$m < \log_2 n - 2$	$n = 1\,048\,576$; $m = 2$	0,816 537
Тест кумулятивных сумм	$n \geq 100$	$n = 1\,048\,576$	0,616 305 0,719 747
Тест на произвольные отклонения	$n \geq 1\,000\,000$	$n = 1\,048\,576$	0,337 162 0,170 294 0,015 065 0,723 129 0,287 306 0,517 442 0,170 294 0,264 458
Разновидность теста на произвольные отклонения	$n \geq 1\,000\,000$	$n = 1\,048\,576$	0,585 209 0,585 209 0,170 294 0,517 442 0,105 618 0,900 104 0,242 986 0,997 147 0,551 026 ...

Из таблицы видно, что построенный ГПСЧ на основе сигнала X системы Лоренца успешно прошел статистическое тестирование пакетом NIST STS, а порождаемые им ПСП обладают статистическими свойствами случайных последовательностей.

Q-тест Льюнга – Бокса (данные получены на тех же $N = 1000$ последовательностях длиной $n = 1\,048\,576$, которые использовались в предыдущем тестировании) выдает значение вероятности $P = 0,4485$, что свидетельствует об уровне корреляций между отсчетами порождаемых ПСП, допустимом для применения последовательностей в задачах обеспечения информационной безопасности.

Заключение. В статье рассмотрены вопросы влияния выбора параметров временной дискретизации непрерывных хаотических систем при формировании ПСП на их основе. На примере системы Лоренца показано, что, выбирая параметр временной дискретизации исходя из периода квазирезонансных колебаний динамической системы и ее динамики в окрестностях точек устойчивого и неустойчивого равновесия, можно избежать периодического поведения и наличия коротких циклов, характерных для дискретных отображений.

Проведенный анализ статистических характеристик выходных последовательностей цифровой реализации ГПСЧ на основе детерминированного хаоса позволяет говорить об их пригодности к использованию в задачах информационной безопасности.

Список использованных источников

1. Шеннон, К. Математическая теория связи / К. Шеннон // Работы по теории информации и кибернетике. – М. : ИИЛ, 1963. – С. 243–332.
2. Chaotic optimization algorithm based on Tent map / L. Shan [et al.] // *Control and Decision*. – 2005. – Vol. 20, no. 2. – P. 179–182.
3. Kocarev, L. Logistic map as a block encryption algorithm / L. Kocarev, G. Jakimoski // *Physics Letters A*. – 2001. – Vol. 289, no. 4–5. – P. 199–206.
4. Pareek, N. K. Cryptography using multiple one-dimensional chaotic maps / N. K. Pareek, V. Patidar, K. K. Sud // *Physics Letters A*. – 2003. – Vol. 309, no. 1–2. – P. 75–82.
5. Wong, W. K. A modified chaotic. Cryptographic method / W. K. Wong, L. P. Lee. // *Computer Physics Communications*. – 2001. – No. 138. – P. 234–236.
6. A unified approach to fuzzy modelling and robust synchronization of different hyperchaotic systems / H. G. Zhang [et al.] // *Chinese Physics B*. – 2008. – Vol. 17, no. 11. – P. 529–533.
7. Cryptanalysis and improvement on a block cryptosystem based on iteration a chaotic map / Y. Wang [et al.] // *Physics Letters A*. – 2007. – Vol. 363, no. 4. – P. 277–281.
8. Крот, А. М. Спектральный анализ хаотических колебаний в имитационной модели схемы Чжуа, разработанной на основе матричной декомпозиции / А. М. Крот, В. А. Сычев // *Информатика*. – 2019. – Т. 16, № 1. – С. 7–23.
9. Капранов, М. В. Анализ фазовых траекторий в окрестностях особых точек 2-D и 3-D нелинейных систем / М. В. Капранов, А. И. Томашевский. – М. : Изд-во МЭИ, 2003. – 80 с.
10. Аливер, В. Ю. Хаотические режимы в непрерывных динамических системах / В. Ю. Аливер // *Вестник МГТУ им. Н. Э. Баумана. Сер. «Приборостроение»*. – 2006. – № 1. – С. 65–84.
11. Корреляционный анализ режимов детерминированного и зашумленного хаоса / В. С. Анищенко [и др.] // *Радиотехника и электроника*. – 2003. – Т. 48, № 7. – С. 824–835.

References

1. Shannon C. A mathematical theory of communication. *Bell System Technical Journal*, 1948, vol. 27, iss. 3, pp. 379–423.
2. Shan L., Qiang H., Li J., Wang Z. Chaotic optimization algorithm based on Tent map. *Control and Decision*, 2005, vol. 20, no. 2, pp. 179–182.
3. Kocarev L., Jakimoski G. Logistic map as a block encryption algorithm. *Physics Letters A*, 2001, vol. 289, no. 4–5, pp. 199–206.
4. Pareek N. K., Patidar V., Sud K. K. Cryptography using multiple one-dimensional chaotic maps. *Physics Letters A*, 2003, vol. 309, no. 1–2, pp. 75–82.
5. Wong W. K., Lee L. P. A modified chaotic. Cryptographic method. *Computer Physics Communications*, 2001, no. 138, pp. 234–236.
6. Zhang H. G., Zhao Y., Yu W., Yang D. S. A unified approach to fuzzy modelling and robust synchronization of different hyperchaotic systems. *Chinese Physics B*, 2008, vol. 17, no. 11, pp. 529–533.
7. Wang Y., Liao X., Xiang T., Wong K. W., Yang D. Cryptanalysis and improvement on a block cryptosystem based on iteration a chaotic map. *Physics Letters A*, 2007, vol. 363, no. 4, pp. 277–281.
8. Krot A. M., Sychou U. A. A spectral analysis of chaotic oscillations in simulation model of Chua’s circuit developed with use of matrix decomposition. *Informatics*, 2019, vol. 16, no. 1, pp. 7–23 (in Russian).
9. Kapranov M. V., Tomashevskiy A. I. Analiz fazovykh traektoriy v okrestnostyah osobykh tochek 2-D i 3-D nelineynykh system. *Analysis of Phase Trajectories in the Vicinity of Singular Points of 2-D and 3-D Nonlinear Systems*. Moscow, Izdatel'stvo Moskovskogo jenergeticheskogo instituta, 2003, 80 p. (in Russian).
10. Aliver V. Y. Naoticheskie rezhimy v nepreryvnykh dinamicheskikh systemah [Chaotic regimes in continuous dynamical systems]. *Vestnik Moskovskogo Gosudarstvennogo Tehnicheskogo Universiteta im. N. Je. Baumana. Serija «Priborostroenie» [Herald of the Bauman Moscow State Technical University. Series Instrument Engineering]*, 2006, no. 1, pp. 65–84 (in Russian).
11. Anischenko V. S., Vadivasova T. E., Okrokverchov G. A., Strelkova G. I. Korreljacionnyj analiz rezhimov determinirovannogo i zashumlennogo haosa [Correlation analysis of the modes of deterministic and noisy chaos]. *Radiotekhnika i elektronika [Radio Engineering and Electronics]*, 2003, vol. 48, no. 7, pp. 824–835 (in Russian).

Информация об авторах

Дрыбин Евгений Александрович, аспирант, факультет радиофизики и компьютерных технологий, Белорусский государственный университет, Минск, Беларусь.
E-mail: ydrybin@gmail.com

Садов Сергей Васильевич, аспирант, старший преподаватель кафедры интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский государственный университет, Минск, Беларусь.
E-mail: seregasadov@gmail.com

Садов Василий Сергеевич, кандидат технических наук, доцент, профессор кафедры интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский государственный университет, Минск, Беларусь.
E-mail: sadov@bsu.by

Information about the authors

Yaugen A. Drybin, Postgraduate Student, Department of Intelligent Systems, Faculty of Radiophysics and Computer Technologies, Belarusian State University, Minsk, Belarus.

E-mail: ydrybin@gmail.com

Siarhei V. Sadau, Postgraduate Student, Senior Lecturer, Department of Intelligent Systems, Faculty of Radiophysics and Computer Technologies, Belarusian State University, Minsk, Belarus.

E-mail: seregasadov@gmail.com

Vasili S. Sadau, Cand. Sci. (Eng.), Associate Professor, Professor of the Department of Intelligent Systems, Faculty of Radiophysics and Computer Technologies, Belarusian State University, Minsk, Belarus.

E-mail: sadov@bsu.by

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

ОБРАБОТКА ИЗОБРАЖЕНИЙ, РЕЧИ И ТЕКСТА
IMAGE, SPEECH AND TEXT PROCESSING

УДК 004.9
<https://doi.org/10.37661/1816-0301-2020-17-4-48-60>

Поступила в редакцию 17.08.2020
Received 17.08.2020

Принята к публикации 07.09.2020
Accepted 07.09.2020

**Компьютеризированная диагностика рака простаты
на основе полнослайдовых гистологических изображений
и методов глубокого обучения**

В. А. Ковалев^{1, 2✉}, Д. М. Войнов², В. Д. Малышев^{1, 2}, Е. Д. Лапо^{1, 2}

¹*Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь*
✉E-mail: vassili.kovalev@gmail.com

²*Белорусский государственный университет, Минск, Беларусь*

Аннотация. Представлены результаты экспериментальных исследований и разработки средств автоматического анализа и распознавания гистологических изображений с целью получения количественных оценок наличия и степени агрессивности рака простаты в общепринятых шкалах Глисона и ISUP. В качестве исходных данных использовались 10 616 полнослайдовых гистологических изображений с размером большей стороны до 100 000 пикселей и 22 089 их фрагментов размером 256×256 пикселей. Проведена оценка эффективности решения задачи с применением как традиционных методов, так и методов глубокого обучения. В качестве финальных выбраны два решения. Первое решение основано на последовательном анализе фрагментов изображений и включает выделение признаков с использованием сети ResNet50 и последующим обобщением частных результатов распознавания с помощью небольшой сверточной сети. Второе решение базируется на одновременном анализе отобранных информативных участков, представленных в виде промежуточного псевдоизображения, и последующем его распознавании с использованием ансамбля из четырех вариантов сверточных сетей с архитектурой EfficientNetB0. В результате независимого тестирования на закрытом наборе изображений, недоступных авторам, достигнута точность предсказания финальной оценки по шкале ISUP, равная 0,9277.

Ключевые слова: рак простаты, гистология, полнослайдовые изображения, глубокое обучение, сверточные нейронные сети

Благодарности. Работа была выполнена при частичной финансовой поддержке проекта ГКНТ Беларуси, договор № 225/4/2019. Авторы выражают глубокую благодарность Д. А. Павленко за помощь в поиске и анализе дополнительной информации, необходимой для выполнения данной работы.

Для цитирования. Компьютеризированная диагностика рака простаты на основе полнослайдовых гистологических изображений и методов глубокого обучения / В. А. Ковалев [и др.] // Информатика. – 2020. – Т. 17, № 4. – С. 48–60. <https://doi.org/10.37661/1816-0301-2020-17-4-48-60>

Computerized diagnosis of prostate cancer based on whole slide histology images and deep learning methods

Vassili A. Kovalev^{1,2✉}, Dmitry M. Voynov², Valery D. Malyshau^{1,2}, Elizabeth D. Lapo^{1,2}

¹*The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus*

✉*E-mail: vassili.kovalev@gmail.com*

²*Belarusian State University, Minsk, Belarus*

Abstract. This paper presents the results of an experimental study and the development of tools for automatic analysis and recognition of histological images in order to obtain quantitative estimates of the presence and degree of aggressiveness of prostate cancer in the commonly used Gleason and ISUP scales. The input data consisted of 10 616 whole-slide histological images with the size of the largest side up to 100 000 pixels and 22 089 of their image tiles of 256×256 pixels in size. Two solutions were chosen as the final ones. The first solution is based on sequential analysis of image fragments and includes feature extraction using the ResNet50 network and the subsequent generalization of particular recognition results using a small convolutional network. The second solution is based on the simultaneous analysis of the selected informative sections, presented in the form of an intermediate pseudo-image, and its subsequent recognition using an ensemble of four variants of convolutional networks with the EfficientNetB0 architecture. Being independently tested on an unknown image dataset that was not available for authors, these approaches achieved the prediction accuracy of 0,9277 according to the ISUP scale.

Keywords: prostate cancer, histology, whole slide histology, deep learning, convolutional neural networks

Acknowledgements. The work was carried out with partial financial support from the project of the State Committee for Science and Technology of Belarus, contract no. 225/4/2019. The authors are also deeply grateful to D. A. Pavlenko for help in finding and analyzing additional information that was necessary for this work.

For citation. Kovalev V. A., Voynov D. M., Malyshau V. D., Lapo E. D. Computerized diagnosis of prostate cancer based on whole slide histology images and deep learning methods. *Informatics*, 2020, vol. 17, no. 4, pp. 48–60 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-48-60>

Введение. Рак простаты является одним из распространенных видов онкологических заболеваний у мужчин. Каждый год в мире диагностируется более 1 млн новых случаев заболевания и умирает более 350 тыс. человек [1]. Одним из общепризнанных способов уменьшения смертности от рака простаты является его раннее обнаружение и правильная оценка степени агрессивности. В качестве способа количественной оценки заболевания используется его рейтинг по так называемой шкале Глисона (Gleason Score, [2]), предложенной Дональдом Глисоном. Показатель Глисона определяется визуально по гистологическим изображениям высокого разрешения, которые снимаются с образцов ткани, получаемых в результате биопсии. Существующие варианты данной шкалы незначительно отличаются друг от друга в зависимости от исторически сложившихся предпочтений и некоторых других причин. Однако общей основой оценки является базовая шкала с целыми значениями {0, 3, 4, 5}. При этом уровень 0 соответствует здоровой ткани, уровень 3 назначается участку биопсии с минимальной злокачественностью и положительным прогнозом выживаемости при соответствующем лечении, а уровень 5 обозначает максимальную агрессивность рака с вероятным (иногда уверенным) отрицательным исходом. Соответственно, уровень 4 является промежуточным между 3 и 5. Учитывая то, что изображение образца ткани может быть очень неоднородным как по своим визуальным свойствам, так и по показателям Глисона, в используемой международной системе принято указывать два уровня. Первый уровень соответствует участку ткани, площадь которого является наибольшей на изучаемом образце ткани, а второй – участку, следующему за ним по площади. Указанные номера уровней принято записывать через знак «+». Так, например, образец ткани, большая часть которой представляет собой начальную стадию заболевания, а меньшая – его наиболее агрессивную форму, кодируется как «3+5»; образец, на котором полностью доминирует последняя стадия, представляется в виде «5+5» и т. д. В дальнейшем сводный показатель вычисляется либо просто как сумма указанных чисел, либо, что чаще всего, используется как

некоторый логичный с медицинской точки зрения показатель, определяемый в виде таблично заданной функции типа $0 + 0 = 0$, $3 + 3 = 1$, $3 + 4 = 2$, $4 + 3 = 3$, $3 + 5 = 4$, $4 + 5 = 5$ и $5 + 5 = 5$. Указанная интегрированная шкала получила название ISUP (ISUP Score, [2]). Очевидная «нелинейность» шкалы ISUP отражает значительно большую опасность высоких значений, особенно максимального значения показателя Глисона, равного пяти. Это обусловлено тем фактом, что в случае присутствия такой агрессивной ткани, даже в небольшом ее количестве, существует весьма высокая вероятность быстрого развития заболевания в негативном направлении.

С точки зрения машинного обучения в рассматриваемой задаче требуется предсказать некоторое число, которое может иметь одно из нескольких значений шкалы ISUP в интервале 0–5. Поскольку ISUP представляет собой шкалу агрессивности рака и имеет место упорядоченность возможных значений, важно предсказывать не только правильный ответ в виде некоторого абстрактного номера класса, но и значение, наиболее близкое к нему. В первом случае, т. е. без учета наличия отношения порядка на номерах классов, имеет место обычная постановка задачи классификации, в то время как при учете данного факта более корректной является регрессионная постановка.

Вместе с тем следует отметить, что диагностика рака простаты представляет собой достаточно сложную задачу даже для опытных врачей-патологов [1, 3]. Как следствие, в ряде случаев может наблюдаться значительная рассогласованность заключений различных экспертов по одному и тому же пациенту. Сложность диагностики обусловлена целым рядом различных факторов, к числу которых относится большая неоднородность образца ткани, высокая вариабельность морфологического строения и результирующих пространственно-цветовых паттернов даже внутри одного показателя Глисона, существенные различия гистологической картины у разных пациентов, наличие артефактов подготовки, окраски и сканирования образцов ткани и ряд других [4, 5].

К настоящему времени уже был выполнен ряд работ по проблеме компьютеризированной диагностики и оценки степени злокачественности рака простаты [5–8]. Так, в работах [5, 6] с помощью методов глубокого обучения были получены хорошие результаты при решении задач бинарной классификации ткани простаты (рак против нормы, ранняя стадия рака против поздней). В то же время уровень определения конкретного показателя Глисона был значительно хуже при применении компьютеризированных систем не только в качестве самостоятельного предсказателя ([5, 8], каппа 0,62 и 0,70 соответственно), но и в качестве вспомогательного модуля в процессе поддержки принятия решений гистопатологами ([7], каппа 0,733). При этом во всех рассмотренных работах количество полнослайдовых изображений (ПСИ), использованных для обучения нейронных сетей, было существенно меньше, чем в настоящем исследовании. Тем не менее среди них следует отметить работу [6], в которой исследование проводилось на значительном количестве ПСИ. Однако большинство изображений (11 429) использовалось все же только для тестирования.

Целью данной работы являлось экспериментальное исследование и разработка средств автоматического анализа и распознавания гистологических ПСИ, а также получение общепринятых количественных оценок наличия и степени агрессивности рака простаты для их последующего использования в качестве «второго мнения» врачами-патологами. Работа выполнялась в рамках международного соревнования по диагностике рака простаты PANDA (URL: <https://panda.grand-challenge.org/>), в котором принимали участие более тысячи команд из разных стран. По результатам независимого тестирования программного обеспечения профильными специалистами на недоступном авторам наборе изображений данная работа заняла 18-е место из 1010 возможных, что соответствует верхней части списка (топ 1,8 %) лауреатов серебряной медали (URL: <https://www.kaggle.com/c/prostate-cancer-grade-assessment/leaderboard>).

1. Исходные данные

Исходный набор гистологических изображений (табл. 1, акроним DS_RAW). Данный набор состоял из 10 616 ПСИ (рис. 1, а) с соответствующими показателями в шкалах ISUP и Глисона. Изображения были предоставлены Университетом Неймегена, Голландия (источник А, увеличение $\times 20$, пиксел 0,24 мкм, второй уровень с пикселом 0,48 мкм, рис. 1, б) и Каролинским институтом, Швеция (источник Б, увеличение $\times 20$, пиксел 0,48 мкм, рис. 1, в). Следует

отметить, что очевидные цветовые различия изображений обусловлены особенностями технологии подготовки и окраски образцов ткани, характерными для каждого медицинского учреждения, и напрямую не связаны с наличием или отсутствием злокачественных опухолей. Для каждого ПСИ имелись маски, на которых указаны опухолевые области и их показатели Глисона. Из этого набора данных были исключены 671 ПСИ, которые содержали артефакты или у которых информация на маске и показатели не соответствовали друг другу (URL: <https://www.kaggle.com/dannellyz/collection-of-600-suspicious-slides-data-loader>). В результате очистки был получен набор данных DS_CLEAN. В качестве тестового набора данных использовались два закрытых набора ПСИ. Первый из них включал 420 ПСИ, а второй – 580 ПСИ, доступных только организаторам соревнования.

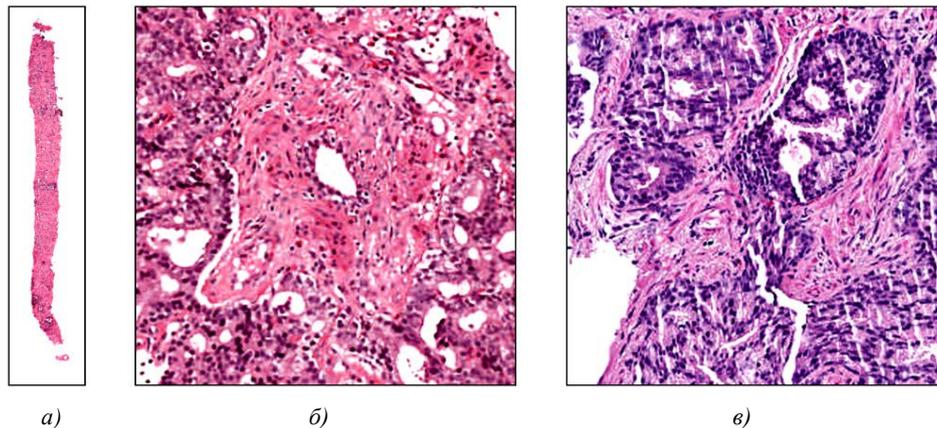


Рис. 1. Примеры исходных изображений: а) ПСИ размером 18 176×4096 пикселей; б), в) тайлы размером 256×256 пикселей из источников А и Б соответственно

Таблица 1

Количественный состав наборов данных

Акроним	Количество ПСИ	Количество тайлов	Баланс показателей, %										
			ISUP							Глисона			
			0	1	2	3	4	5	0	3	4	5	
DS_RAW	10 616	–	27	25	13	12	12	11	18	51	45	13	
DS_CLEAN	9945	–	27	25	13	12	12	11	18	51	45	13	
DS_SMAL	3150	6000	11	15	15	16	13	30	39	22	24	20	
DS_MAIN	5067	22 089	51	32	–	–	15	2	43	25	19	13	

В процессе разметки входных данных были допущены ошибки, из-за которых до 30 % всех ПСИ имеют неверные показатели в шкалах ISUP и Глисона (URL: <https://panda.grand-challenge.org/>).

Предварительный набор изображений (табл. 1, акроним DS_SMAL). Все ПСИ на уровне максимального разрешения были разбиты на так называемые тайлы, т. е. некоторые элементарные участки квадратной формы размером 256×256 пикселей. Тайлы, содержащие фон, удалялись. В результате разбиения было получено около 8 млн тайлов. Так как только небольшое количество тайлов относилось к третьему, четвертому и пятому классам Глисона, для получения сбалансированного набора данных, необходимого для обучения нейронных сетей, было выбрано их некоторое подмножество, обеспечивающее равномерное представительство классов. Кроме того, вычисления с использованием полного набора тайлов заняли бы слишком много времени, что делает нецелесообразным использование всех 8 млн. С целью ускорения предварительных вычислительных экспериментов из указанного исходного набора случайным образом было выбрано 6000 тайлов.

Основной набор изображений (табл. 1, акроним DS_MAIN). Данный набор использовался для обучения нейронных сетей, которые определяли оценку по шкале Глисона на уровне одиночного тайла. Во избежание ошибок неверного назначения показателя Глисона из-за наличия

сразу двух его значений у каждого ПСИ нужные тайлы отбирались из ПСИ, которые имели два равных показателя: 3 + 3, 4 + 4, 5 + 5, а также с ПСИ, содержащих нормальную ткань и, соответственно, размеченных как 0 + 0. При использовании такого подхода выявился недостаток тайлов из источника Б с пятым показателем Глисона. Чтобы сбалансировать набор данных, в этот класс были добавлены тайлы из ПСИ с показателями 4 + 5 и 5 + 4 по шкале Глисона. Такой подход уменьшает точность разделения четвертого и пятого классов, но в целом повышает значение тестовой статистической метрики классов каппа [9], вычисляемой следующим образом:

$$\kappa_w = 1 - \frac{q_0}{q_e},$$

$$q_0 = \sum_i \sum_j v_{ij} p_{ij},$$

$$q_e = \sum_i \sum_j v_{ij} p_i^* p_j,$$

где p_i^* – доля объектов выборки, которые нейронная сеть отнесла к категории i по значениям шкалы ISUP i ; p_j – доля объектов из категории j ; p_{ij} – доля объектов, относящихся к категории j , которые нейронная сеть предсказала как объекты из категории i , $v_{ij} = (i - j)^2$. Все тайлы были выбраны из слоя ПСИ с увеличением, в четыре раза меньшим, чем максимальное, и имели размер 256×256 пикселей. Информация о количественном составе описанных выше наборов изображений представлена в табл. 1.

Следует отметить, что некоторые тайлы, как и целые ПСИ, могут иметь области с различными показателями Глисона, т. е. даже изображения размером 256×256 пикселей могут одновременно содержать участки, представляющие разные классы.

Как следует из табл. 1, в наборе данных DS_MAIN на один ПСИ в среднем приходится четыре тайла, которые не всегда являются смежными. Таким образом, степень увеличения количества данных как таковых за счет перекрытия тайлов несущественная. Более того, это может внести дополнительные неточности из-за неизбежного увеличения доли фона на тайлах, так как при игольчатой биопсии простаты ширина образца ткани соответствует всего одному-трем тайлам. Отсутствие перекрытия тайлов частично компенсируется их аугментацией, включающей повороты и отражения в процессе обучения нейронных сетей.

2. Общая схема решения задачи. Как уже упоминалось ранее, ПСИ могут достигать весьма значительных размеров, вплоть до 100 000 пикселей по каждому измерению с объемом занимаемой памяти до 10 Гб (при обработке без сжатия). При таких размерах использование традиционных методов достаточно затруднительно, а существующие методы глубокого обучения вовсе не приспособлены для анализа данных больших размеров. Кроме того, алгоритмы распараллеливания вычислений на графических ускорителях (GPU) при глубоком обучении не позволяют размещать требуемое количество изображений в графической памяти одновременно даже при использовании самых современных устройств указанного типа. Поэтому для решения данной задачи необходимо проводить определенную предобработку, которая либо сокращает размер исходной задачи (при этом теряется определенная информация), либо делит ее на несколько меньших подзадач, решаемых существующими методами, с последующим объединением частных решений.

Эффективной реализацией *первого подхода* к предобработке является метод, основанный на разбивке исходного изображения по схеме «ПСИ → полотно (мозаика) тайлов». При этом под тайлом здесь понимается некоторый небольшой участок изображения, который содержит фрагмент изучаемой ткани. Сама предобработка включает в себя следующие три шага:

- разделение ПСИ на непересекающиеся тайлы таким образом, чтобы каждый пиксел изображения принадлежал ровно одному тайлу;
- выбор определенным способом среди полученных тайлов заранее заданного числа некоторых представительных тайлов;
- расположение представительных тайлов в прямоугольнике (квадрате) и сохранение их как единого изображения.

При *втором подходе* исходное ПСИ последовательно анализируется и классифицируется тайл за тайлом. По завершении процесса полученные частные результаты распознавания обобщаются с использованием некоторого алгоритма и соответствующий шкале ISUP класс назначается всему ПСИ.

3. Предварительные эксперименты. Предварительные эксперименты проводились для оценки сложности решения поставленной задачи распознавания изображений; выявления различий изображений, связанных с их съемкой в разных странах; получения сравнительных оценок ошибок классификации, характерных для разных классов; выяснения представительства и существующих дисбалансов классов в обучающей выборке на уровне отдельных тайлов, а также для других целей, характерных для этапа разведочного анализа данных.

Оценка эффективности традиционных методов распознавания. Оценка эффективности традиционных подходов, включающих вычисление информативных признаков изображений с последующим обучением классификаторов для решения задачи распознавания, выполнялась на усеченном, «пробном» наборе изображений DS_SMAL. Учитывая тот факт, что гистологические изображения представляют собой некоторый специфический вид цветных текстур, в качестве количественных параметров описания изображений использовались следующие известные текстурные признаки:

- текстурные признаки Габора [10] в виде статистических моментов выходных сверток с 40 фильтрами, пятью масштабными уровнями, восемью направлениями;
- гистограммы ориентации градиентов (HOG) с шестью направлениями в интервале 0–180°;
- локальные бинарные шаблоны (LBP) с восемью соседями и радиусами окрестностей $R = 1, 2$;
- улучшенная версия матриц совместной встречаемости цветов пикселей изображений [11], представленных в данном случае в оптимальной палитре из 128 цветов, выбранных с использованием одного из вариантов метода кластеризации на базе K-средних;
- площади и статистические моменты яркостей суперпикселей, получаемых при адаптивном разбиении изображений на суперпиксели методом простой линейной итеративной кластеризации SLIC [12].

В ситуациях, когда признаки изображений были представлены короткими векторами, например в случае ориентационных гистограмм HOG, они подавались на вход классификаторов напрямую. В случае большого количества элементов дескрипторов изображений, как это имеет место, например, при использовании матриц совместной встречаемости цветов, исходные признаки редуцировались до достаточно малого числа некоррелированных переменных с применением метода главных компонент с порогом отбора 95 % объясняемой вариабельности входных данных. Исключение составляли лишь эксперименты по поиску наиболее похожих изображений по образцу с целью автоматизации процессов формирования обучающих выборок. В этом случае векторизованные версии матриц совместной встречаемости сравнивались напрямую с помощью метрики L1 [13]:

$$L1(p, q) = \sum_{i=1}^n |p_i - q_i|,$$

где p, q – векторизованные версии матриц совместной встречаемости. Метод главных компонент со встроенным нормированием переменных также использовался при объединении и последующем совместном использовании разносортных признаков.

В качестве классификаторов применялись такие известные методы, как SVM, Random Forests и kNN. В силу очевидных приоритетов разведочного анализа, мотивированных необходимостью изучения свойств исходных данных как таковых, а не эффективности различных методов классификации и распознавания, другие классификаторы, равно как и оптимизация управляющих параметров указанных классификаторов, не использовались, т. е. опция сеточной грид-оптимизации параметров классификаторов при вызове реализующих их программных модулей была выключена. Все вычислительные эксперименты, рассматриваемые в разд. 3, проводились с использованием языка R, одноименной платформы разработки (URL: <https://www.R-project.org/>) и входящих в нее библиотек.

Основные типы признаков изображений, перечисленных выше, показаны на рис. 2. Общие результаты предварительных экспериментальных исследований данного направления приведены в табл. 2. Для каждого типа признаков были проведены четыре серии экспериментов по точности предсказания наличия рака и его агрессивности в следующих четырех шкалах: значение ISUP для большого сегмента изображения, указанного в маске; значение Глисона 1; значение Глисона 2; значение в шкале ISUP. Кроме точного предсказания в виде конкретного значения класса, задача предсказания также решалась для некоторых объединений соседних классов, что имеет определенный медицинский смысл. Например, рак отсутствует (0), рак умеренно агрессивен (3u4) или это крайне агрессивная форма рака (5). Поскольку точность предсказания в зависимости от типа используемых классификаторов варьировала незначительно (порядка 0,5–1,5 %), тип классификатора в табл. 2 не указывается.

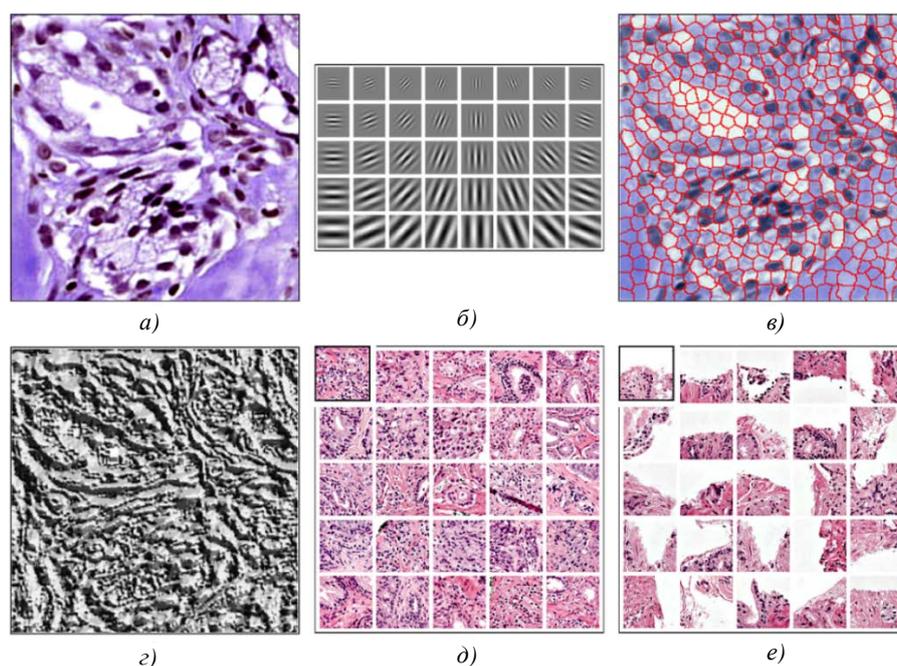


Рис. 2. Исходное тестовое гистологическое изображение (тайл 256×256 пикселей) (а); используемые фильтры Габора (б); суперпиксельное представление исходного изображения (в); карта локальных бинарных шаблонов (г); примеры поиска изображений по двум образцам, представленным в левых верхних углах с применением матриц совместной встречаемости цветов (д, е)

Таблица 2

Точность предсказания опасности рака традиционными методами

Тип признаков, схема объединения классов	Тип шкалы, количество классов			
	Большой сегмент, пять классов	Глисон 1, четыре класса	Глисон 2, четыре класса	ISUP, пять классов
Габор и HOG	0,418	0,493	0,506	0,491
Суперпиксели	0,435	0,532	0,534	0,528
Габор и суперпиксели	0,448	0,539	0,534	0,532
Совместная встречаемость цветов	0,452	0,535	0,531	0,530
Все признаки	0,466	0,550	0,545	0,539
Все признаки, (0u1u2), (3u4), (5)	0,597	–	–	0,603
Все признаки, (0), (3u4), (5)	–	0,603	0,611	–

Представленные в табл. 2 результаты показывают, что при использовании традиционных методов точность предсказания наличия и степени агрессивности рака достаточно невысокая. Вероятнее всего, причинами этого является целый ряд факторов, включая ориентацию каждого типа традиционных признаков на конкретные, наперед заданные свойства изображений; отсутствие механизма адаптации к особенностям изображений, заданных обучающей выборкой;

слабую способность традиционных классификаторов к обобщению пространства признаков (generalization); малую «емкость» набора обучаемых параметров (десятки и сотни против сотен тысяч и миллионов у сверточных нейронных сетей) и др. Кроме того, при интерпретации результатов предсказания, приведенных в табл. 2, следует учитывать, что разметка классов в пробном наборе изображений DS_SMAL сильно зашумлена, т. е. количество ошибок назначения классов медиками точно неизвестно.

Предварительные эксперименты с использованием методов глубокого обучения. Для решения задачи классификации тайлов был обучен ряд общедоступных архитектур сверточных сетей, включая ResNet50, MobileNetV2, VGG16. Процесс обучения сетей выполнялся с применением оптимизатора RMSprop. Когда потери на валидационной выборке изображений переставали уменьшаться, скорость обучения (learning rate) снижалась в 10 раз. К входным данным применялась аугментация, включающая в себя стандартные геометрические операции, такие как отражение, случайный поворот на 90° один или более раз, отражение по диагонали. В связи с тем что изображения из представленных источников имеют различные цветовые палитры, к ним была применена цветовая нормализация методом Масенко [14] с использованием маски, основанная на декомпозиции по цвету. Нормализованные тайлы из источников данных А и Б показаны на рис. 3. В обоих случаях верхние восемь тайлов получены из источника А, а нижние восемь – из источника Б. В ходе тестирования наилучшие результаты были получены с помощью архитектуры ResNet50, поэтому она была выбрана для всех последующих экспериментов. Результаты проведенных экспериментов приведены в табл. 3.

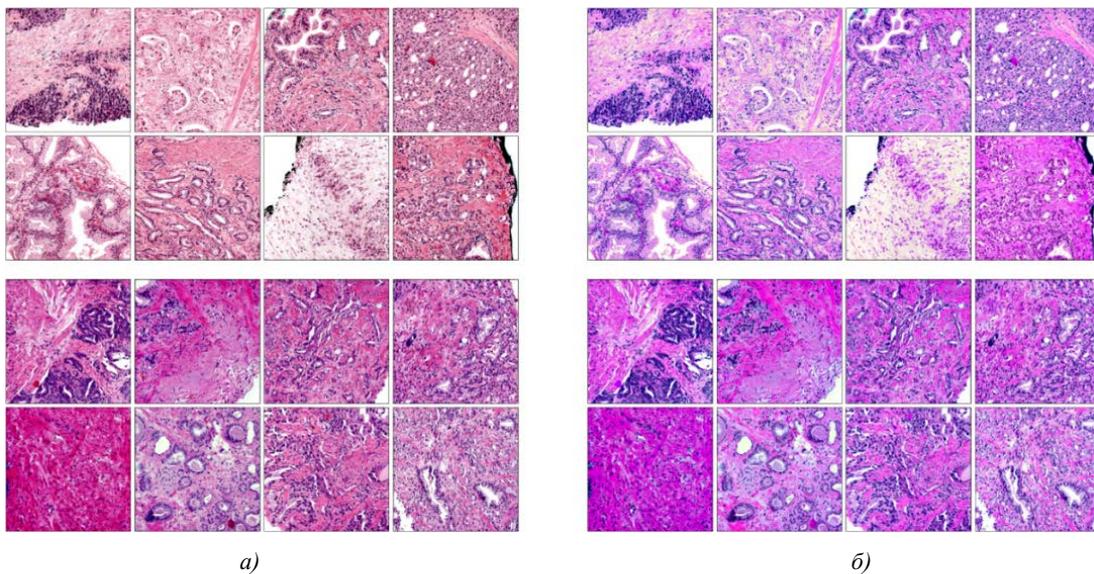


Рис. 3. Примеры исходных изображений (а) и их нормализованные версии (б)

Таблица 3

Результаты предварительных экспериментов по распознаванию тайлов

Архитектура сети, дополнительные условия	Точность классификации	Взвешенная каппа-метрика
ResNet50	0,87	0,87
MobileNetV2	0,82	0,81
VGG16	0,68	0,71
ResNet50 (с MaxPooling)	0,84	0,83
ResNet50 (с удалением ошибочных тайлов)	0,86	0,83

В качестве эксперимента также были сделаны попытки обучения сети с каппа-потерями (каппа-loss) в качестве функции потерь, однако такой метод не позволил получить значение по метрике «точность» (accuarcy) выше 0,59. Вариант с усреднением категориальной функции потерь и каппа-потерями оказался тоже неэффективным.

В связи с тем что сеть, обученная для классификации тайлов, не использовалась явно для предсказания, а выступала в роли алгоритма извлечения признаков изображений (feature extractor), с целью потенциального улучшения представления получаемой карты признаков была обучена нейронная сеть, основанная на архитектуре ResNet50, где предпоследний слой GlobalAveragePooling был заменен на MaxPooling. Данная модель не превзошла рассмотренные на этапе классификации тайлов модели и не дала видимых преимуществ на последующих этапах решения задачи.

При детальном изучении результатов тестирования было отмечено, что среди тестовых тайлов, вырезанных из ПСИ, которые получены из источника Б, есть достаточно большой процент тех, что были помечены как раковые, но ошибочно классифицировались сетью как норма. Этот результат мог получиться за счет того, что для изображений из источника Б была предоставлена только приблизительная область, в которой находится опухоль, однако точная локализация раковых клеток не была выполнена. Следовательно, при выделении тайлов из такой области некоторые из них могли действительно не содержать образцов ткани с раком. Для очистки данных от подобных тайлов применялась следующая последовательность действий:

- 1) обучение сети на всех данных;
- 2) предсказание обучающей и валидационной выборки с помощью обученной сети;
- 3) исключение из выборок тайлов с показателем Глисона 3, 4, 5, принадлежащих источнику Б, для которых была предсказана оценка по шкале Глисона 0;
- 4) обучение сети с использованием новых данных.

Представленный подход позволил повысить значение метрики «отзыв» (recall) для показателя Глисона 0, однако значение данной метрики для остальных оценок по шкале Глисона упало.

Реализация всех методов решения задачи с помощью второго подхода (см. разд. 2), основанного на классификации изображений с помощью нейронных сетей, была осуществлена на языке программирования Python с использованием библиотеки Keras [15], которая позволяет интегрировать такие упомянутые выше архитектуры нейронных сетей, как ResNet50, MobileNetV2, VGG16.

Предварительные эксперименты также проводились для отработки первого подхода к решению задачи (см. разд. 2), который предполагал формирование промежуточного изображения-полотна из представительных тайлов. В качестве представительных выбирались тайлы, которые содержали наибольшую долю (площадь), занимаемую непосредственно тканью простаты. Примеры подобных участков представлены на рис. 2, *д*, в то время как неподходящие тайлы с малой площадью ткани – на рис. 2, *е*. Такой способ показал хорошую эффективность ввиду специфики анализируемых данных: исходные ПСИ содержали снимки тонкой полоски ткани простаты, а большую часть изображения занимал фон.

В результате предобработки описанным способом каждому ПСИ из исходного набора ставится в соответствие гораздо меньшее изображение, которое тем не менее является достаточно представительным в отношении разнообразия участков ткани. Благодаря этому набор данных, состоящий из пар типа «ПСИ – ISUP score», преобразуется в набор пар типа «полотно тайлов – ISUP score», в котором полотно из тайлов можно подвергнуть анализу существующими методами глубокого обучения. Учитывая наличие некоторой упорядоченности значений шкалы ISUP, задача диагностики рака простаты представлялась как задача ординарной регрессии. Для этого метка каждого изображения отображалась в вектор из пяти чисел так, что для метки M в полученном векторе первые M элементов принимали значение 1, а остальные элементы – значение 0. Например, значение 0 представлялось как $[0, 0, 0, 0, 0]$, а значение 3 – как $[1, 1, 1, 0, 0]$.

Полученная задача решалась известными методами глубокого обучения, а именно с помощью глубокой нейронной сети в роли предсказателя. В качестве архитектуры нейронной сети была выбрана EfficientNet-B0 ввиду ее высокой эффективности при достаточно малом размере и относительно небольшом времени, необходимом для ее обучения. С учетом того что была поставлена задача ординарной регрессии, построенная нейросетевая модель обладала следующими особенностями:

- последний слой сети состоял из пяти нейронов;

– для того чтобы элементы выходного вектора находились в интервале от 0 до 1, в качестве функции активации была выбрана логистическая функция (сигмоид);

– для получения значения шкалы ISUP выход такой сети суммировался и округлялся. Например, вектор предсказаний [0,9, 0,7, 0,8, 0,1, 0,2] соответствовал значению 3 шкалы ISUP.

С учетом приведенных выше особенностей в качестве функции потерь была выбрана бинарная кросс-энтропия [16]. В случае, когда предсказание идет для некоторых двух классов А и Б, кросс-энтропию можно вычислить по формуле

$$L_{\text{cross-entropy}} = -(y \log(p) + (1 - y) \log(1 - p)),$$

где y – бинарный индикатор 0 или 1 в зависимости от того, является ли метка класса А правильной классификацией для текущего наблюдения, а p – предсказанная вероятность того, что текущее наблюдение принадлежит классу А.

Таблица 4

Результаты предсказания значения шкалы ISUP по изображению-полотну

Входной набор данных	Размер полотна тайлов	Взвешенная каппа-метрика
1. DS_CLEAN	4×4	0,884
2. DS_CLEAN	5×5	0,892
3. DS_CLEAN	6×6	0,881
4. DS_RAW	5×5	0,883
5. DS_RAW	6×6	0,883

Из табл. 4 видно, что изменение размеров изображения-полотна несущественно влияло на качество конечных результатов распознавания. Кроме того, проводились эксперименты с решением исходной задачи в классификационной, а не регрессионной постановке. Однако качество получаемых результатов в этом случае было существенно ниже.

4. Финальные эксперименты

Решение, основанное на последовательной классификации тайлов ПСИ. В данном случае задача решалась согласно второму подходу в два этапа. На первом этапе проводился последовательный анализ тайлов, далее на втором этапе полученные результаты анализа обобщались для всего ПСИ. Оба этапа выполнялись с помощью соответствующих нейронных сетей. На первом этапе использовалась сеть, основанная на архитектуре ResNet50 (см. разд. 3). Результатом ее работы были признаки, необходимые для предсказания показателей Глисона для каждого текущего тайла анализируемого ПСИ. На втором этапе применялась некоторая небольшая сверточная нейронная сеть, входом которой являлись упомянутые признаки показателей Глисона, а выходом – финальные показатели в шкале ISUP. Выполнение подхода было обусловлено тем, что в существующей медицинской практике показатель ISUP определяется на основе соотношения областей на ПСИ с различными показателями Глисона.

В техническом плане в качестве входных признаков второй, обобщающей нейронной сети были выбраны 2048 признаков с предпоследнего слоя первой сети. Из этих признаков формировалось некоторое псевдоизображение. Таким образом, каждому ПСИ соответствовало псевдоизображение размером $WT \times HT \times 2048$, где WT и HT – это количество тайлов в ширину и высоту ПСИ соответственно. Именно такие псевдоизображения подавались на вход второй сверточной нейронной сети для получения выходного показателя ISUP.

Поскольку пространственное разрешение псевдоизображений, получаемых после работы первой сети, было слишком мало, использовать на втором этапе все преимущества и обобщающие возможности таких широко известных нейронных сетей, как VGG16, ResNet, EfficientNet и др., не представлялось возможным. Поэтому была предложена некоторая небольшая сеть, архитектура которой детально представлена в табл. 5. В этой нейронной сети между всеми сверточными слоями использовалась функция ReLU и пакетная нормализация.

Для предсказания показателя ISUP с помощью второй нейронной сети применялись такие функции потерь, как категориальная кросс-энтропия, каппа-функция или их комбинация.

Таблица 5

Сверточная сеть, используемая для предсказания показателей ISUP

Тип слоя	Количество каналов	
	входных	выходных
Сверточный слой $3 \times 3 \times 2$ + MaxPool	2048	1024
Сверточный слой $3 \times 3 \times 2$ + MaxPool	1024	512
Сверточный слой $3 \times 3 \times 2$ + MaxPool	512	256
Сверточный слой $3 \times 3 \times 2$ + MaxPool	256	128
Полносвязный слой	128	32
Полносвязный слой	32	16
Полносвязный слой	16	6

Поскольку категориальная кросс-энтропия не учитывает порядок в распределении классов, с ее помощью не удалось добиться высоких показателей тестовой метрики. С другой стороны, каппа-функция в качестве функции потерь является довольно чувствительной к выбросам, что не позволило добиться эффективного процесса обучения. Поэтому было принято решение использовать комбинированную функцию, которая представляет собой сумму категориальной кросс-энтропии, и каппа-функцию. В процессе обучения энтропия быстро получает оптимальное значение, после чего нейронная сеть оптимизирует в основном каппа-функцию, что позволяет достичь лучшие показатели точности.

В результате реализации описанной методики были получены следующие значения точности предсказания показателей ISUP ПСИ: 0,807 на доступной тестовой выборке и 0,811 на неизвестной авторам тестовой выборке.

Решение, основанное на промежуточном изображении из информативных тайлов. Задача решалась согласно первому подходу, т. е. методом формирования промежуточного изображения-полотна из информативных тайлов. В качестве последних выбирались тайлы, которые содержали наибольшую долю (площадь), занимаемую непосредственно тканью простаты. В отличие от предыдущего решения в данном случае задача решалась не в классификационной, а регрессионной постановке. Для повышения качества распознавания использовались различные приемы, основные из них представлены ниже:

1. Построение ансамбля моделей. Среди имеющихся обученных нейронных сетей выбиралось некоторое их подмножество. Каждая модель из подмножества совершала предсказание значений шкалы ISUP каждого изображения из тестовой выборки, в качестве ответа выбиралось среднее значение их предсказаний. Стоит отметить, что существуют более сложные методы построения ансамблей моделей, такие, например, как взвешенное среднее, обучение модели по предсказаниям и др. Однако в рамках данной работы выбранный способ показал достаточно большой прирост качества предсказания наличия рака и его стадию.

2. Сдвинутая нарезка тайлов. Помимо нарезки тайлов с верхнего левого угла ПСИ также проводилась нарезка с точки, сдвинутой от этого угла на 64, 128, 192 пиксела вниз и влево. После такой нарезки конструировалось новое изображение-полотно и процесс предсказания выполнялся заново. В результате получались четыре различных предсказания, из которых затем вычислялось среднее. Этот подход позволяет найти различные текстурные особенности изображения, которые могли остаться незамеченными в результате оригинальной нарезки.

3. Аугментация тестовой выборки. В последнее время данный метод стабилизации предсказаний получил очень широкое распространение. Суть его заключается в проведении различных трансформаций изображений и предсказания как для исходной тестовой выборки, так и для трансформированной с последующим усреднением результатов. Назначение метода аналогично аугментации при обучении – это ослабление эффекта переобучения под какие-либо визуальные особенности изображений, не имеющие значения при решении задачи.

В окончательном варианте был применен прием ансамблирования. Среди пяти обученных нейронных сетей, представленных в табл. 4, для построения финального ансамбля были выбра-

ны обученные сети 1, 3, 4 и 5. Сеть 2 не выбиралась по причине более низкого качества ансамблей, получившихся при ее использовании. В качестве метода аугментации был выбран поворот изображений на 90°.

Решение, построенное по описанной методике, получило следующие значения взвешенной каппа-метрики предсказания финальных значений по шкале ISUP: 0,8863 на публичной тестовой выборке и 0,9277 на закрытой выборке, недоступной авторам. Стоит отметить, что аналогичное решение, в котором для ансамбля выбирались только сети 1 и 3, обученные на очищенной выборке, показало значения качества 0,8836 и 0,9297 соответственно. Однако это было обнаружено уже после завершения упомянутого выше международного соревнования по диагностике рака простаты и поэтому данные результаты официально приняты не были.

Заключение. Методы распознавания изображений, основанные на глубоком обучении, обеспечивают существенно более высокое качество решения задач распознавания изображений по сравнению с традиционными подходами, основанными на вычислении признаков (feature extraction) и последующем применении известных классификаторов типа SVM, Random Forests, kNN и др.

Подход, основанный на последовательном распознавании фрагментов изображений (тайлов) с последующим обобщением результатов, показал относительно низкое качество решения задачи и оказался неконкурентоспособным по сравнению с альтернативным подходом.

Наличие определенных отношений порядка типа лучше-хуже на выходных классах должно быть учтено при постановке задачи распознавания и ее реализации в виде нейронной сети, сконфигурированной для работы в классификационной или регрессионной постановке.

С учетом специфики исходных данных замена полнослайдовых изображений на подмножество представительных тайлов позволила построить нейросетевые модели высокого качества, несмотря на существенную потерю информации.

Применяемые в настоящей работе техники стабилизации предсказаний типа ансамблей нейронных сетей и аугментации тестовых данных давали большой прирост качества, в то время как тренировка более сложных нейросетевых моделей чаще всего заканчивалась их переобучением.

References

1. Rawla P. Epidemiology of prostate cancer. *World Journal of Oncology*, 2019, vol. 10, no. 2, pp. 63–89.
2. Epstein J. I., Allsbrook W. C. Jr, Amin M. B., Egevad L. L. ISUP Grading Committee. The 2005 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma. *The American Journal of Surgical Pathology*, 2005, vol. 29, iss. 9, pp. 1228–1242.
3. Camparo P., Egevad L., Algaba F., Berney D. M., Boccon-Gibod L., ..., Varma M. Utility of whole slide imaging and virtual microscopy in prostate pathology. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, 2012, vol. 120, iss. 4, pp. 298–304.
4. Goldenberg S. L., Nir G., Salcudean S. E. A new era: artificial intelligence and machine learning in prostate cancer. *Nature Reviews Urology*, 2019, vol. 16, pp. 391–403.
5. Ström P., Kartasalo K., Olsson H., Solorzano L., Delahunt B., ..., Eklund M. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *The Lancet Oncology*, 2020, vol. 21, iss. 2, pp. 222–232.
6. Pantanowitz L., Quiroga-Garza G., Bien L., Heled R., Laifenfeld D., ..., Dhir R. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. *The Lancet Digital Health*, 2020, vol. 2, iss. 8, pp. e407–e416.
7. Bulten W., Balkenhol M., Belinga J.-J. A., Brilhante A., Çakır A., ..., Litjens G. Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Modern Pathology*, 2020. Available at: <https://arxiv.org/abs/2002.04500> (accessed 06.08.2020).
8. Nagpal K., Foote D., Liu Y., Chen P.-H. C., Wulczyn E., ..., Stumpe M. C. Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *Nature Partner Journal Digital Medicine*, 2019, vol. 2, iss. 48, pp. 1–10.
9. Schuster C. A note on the interpretation of Weighted Kappa and its relations to other rater agreement statistics for metric scales. *Educational and Psychological Measurement*, April 2004, vol. 64, no. 2, pp. 243–253.
10. Luan S., Chen C., Zhang B., Han J., Liu J. Gabor convolutional networks. *IEEE Transactions on Image Processing*, 2018, vol. 27, no. 9, pp. 4357–4366.

11. Kovalev V., Volmer S. Color co-occurrence descriptors for querying-by-example. *International Conference on Multimedia Modeling, Lausanne, Switzerland, 12–15 October 1998*. Lausanne, 1998, pp. 32–38.
12. Achanta R., Shaji A., Smith R., Lucchi A., Fua P., Susstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on PAMI*, 2012, vol. 34, no. 11, pp. 2274–2282.
13. Horn R. A., Johnson C. R. Matrix Analysis. Part 5. *Norms for Vectors and Matrices*. England, Cambridge University Press, 1990.
14. Macenko M., Niethammer M., Marron J., Borland D., Woosley J. T., ..., Thomas N. E. A method for normalizing histology slides for quantitative analysis. *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, USA, 28 June – 1 July 2009*. Boston, 2009, pp. 1107–1110.
15. Gulli A., Sujit P. *Deep learning with Keras*. Packt Publishing Ltd, 2017, 318 p.
16. Rubinstein R. Y., Kroese D. P. *The Cross Entropy Method: a Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation (Information Science and Statistics)*. Berlin, Heidelberg, Springer-Verlag, 2004, 321 p.

Информация об авторах

Ковалев Василий Алексеевич, кандидат технических наук, заведующий лабораторией анализа биомедицинских изображений, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь, доцент, Белорусский государственный университет, Минск, Беларусь.
E-mail: vassili.kovalev@gmail.com

Войнов Дмитрий Михайлович, магистрант, Белорусский государственный университет, Минск, Беларусь.
E-mail: voynovdd@gmail.com

Мальшев Валерий Дмитриевич, инженер-программист лаборатории анализа биомедицинских изображений, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь, магистрант, Белорусский государственный университет, Минск, Беларусь.
E-mail: malyshevalery@gmail.com

Лапо Елизавета Дмитриевна, инженер-программист лаборатории анализа биомедицинских изображений, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь, магистрант, Белорусский государственный университет, Минск, Беларусь.
E-mail: lilibetlapo@gmail.com

Information about the authors

Vassili A. Kovalev, Cand. Sci. (Eng.), Head of the Laboratory of Biomedical Images Analysis, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus, Associated Professor, Belarusian State University, Minsk, Belarus.
E-mail: vassili.kovalev@gmail.com

Dmitry M. Voynov, Undergraduate, Belarusian State University, Minsk, Belarus.
E-mail: voynovdd@gmail.com

Valery D. Malyschau, Software Engineer of the Laboratory of Biomedical Image Analysis, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus, Undergraduate, Belarusian State University, Minsk, Belarus.
E-mail: malyshevalery@gmail.com

Elizabeth D. Lapo, Software Engineer of the Laboratory of Biomedical Image Analysis, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus, Undergraduate, Belarusian State University, Minsk, Belarus.
E-mail: lilibetlapo@gmail.com

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

УДК 004.8
<https://doi.org/10.37661/1816-0301-2020-17-4-61-72>

Поступила в редакцию 17.07.2020
Received 17.07.2020

Принята к публикации 21.09.2020
Accepted 21.09.2020

Моделирование языка и двунаправленные представления кодировщиков: обзор ключевых технологий

Д. И. Качков

Белорусский государственный университет, Минск, Беларусь
E-mail: dmitriydikanskiy@gmail.com

Аннотация. Представлен очерк развития технологий обработки естественного языка, которые легли в основу BERT (Bidirectional Encoder Representations from Transformers) – языковой модели от компании Google, демонстрирующей высокие результаты на целом классе задач, связанных с пониманием естественного языка. Две ключевые идеи, реализованные в BERT, – это перенос знаний и механизм внимания. Модель предобучена решению нескольких задач на обширном корпусе неразмеченных данных и может применять обнаруженные языковые закономерности для эффективного дообучения под конкретную проблему обработки текста. Используемая архитектура Transformer основана на внимании, т. е. предполагает оценку взаимосвязей между токенами входных данных. В статье отмечены сильные и слабые стороны BERT и направления дальнейшего совершенствования модели.

Ключевые слова: информатика, информационные технологии, языковые модели, обработка естественного языка, механизм внимания, архитектура Transformer, модель BERT

Для цитирования. Качков, Д. И. Моделирование языка и двунаправленные представления кодировщиков: обзор ключевых технологий / Д. И. Качков // Информатика. – 2020. – Т. 17, № 4. – С. 61–72. <https://doi.org/10.37661/1816-0301-2020-17-4-61-72>

Language modeling and bidirectional coders representations: an overview of key technologies

Dzmitry I. Kachkou

Belarusian State University, Minsk, Belarus
E-mail: dmitriydikanskiy@gmail.com

Abstract. The article is an essay on the development of technologies for natural language processing, which formed the basis of BERT (Bidirectional Encoder Representations from Transformers), a language model from Google, showing high results on the whole class of problems associated with the understanding of natural language. Two key ideas implemented in BERT are knowledge transfer and attention mechanism. The model is designed to solve two problems on a large unlabeled data set and can reuse the identified language patterns for effective learning for a specific text processing problem. Architecture Transformer is based on the attention mechanism, i.e. it involves evaluation of relationships between input data tokens. In addition, the article notes strengths and weaknesses of BERT and the directions for further model improvement.

Keywords: informatics, information technology, language models, natural language processing, attention mechanism, transformer architecture, model BERT.

For citation. Kachkou D. I. Language modeling and bidirectional coders representations: an overview of key technologies. *Informatics*, 2020, vol. 17, no. 4, pp. 61–72 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-61-72>

Введение. Естественный язык является основным средством коммуникации для человека, и проблема его автоматической обработки – одно из актуальных направлений научных исследований в области искусственного интеллекта. Системы, способные обрабатывать естественный язык, позволяют решать широкий спектр задач, в том числе извлечения информации об окружающем мире из огромных массивов текстов, автоматического перевода с одного языка на другой без участия человека, разработки эффективных интерфейсов взаимодействия между человеком и компьютером и т. д.

Среди различных инструментов обработки естественного языка можно выделить класс средств, базирующихся на языковых моделях – системах, которые были предобучены на больших корпусах текстов рассчитывать распределение вероятности появления тех или иных токенов (слов) на заданных позициях в предложении. Как показали исследования, нейронные языковые модели могут эффективно дообучаться под широкий спектр задач, связанных с интерпретацией информации на естественном языке. По состоянию на май 2020 г. большинство систем, демонстрирующих наилучший результат при решении конкретных задач обработки естественного языка, так или иначе основаны на языковой модели BERT, разработанной в компании Google.

Механизм внимания. В 2014 г. был предложен механизм seq2seq (sequence-to-sequence), идея которого описана в работе [1]. Решение состоит из двух рекуррентных нейронных сетей, представляющих собой кодировщик (encoder) и декодировщик (decoder) соответственно. Задача кодировщика заключается в том, чтобы поставить в соответствие исходной последовательности переменной длины некоторый вектор состояния фиксированной размерности. Декодировщик, в свою очередь, разворачивает этот вектор в целевую последовательность, не имеющую фиксированной длины. Обучение проходит одновременно для обеих сетей: ставится задача максимизировать условную вероятность целевой последовательности по заданной исходной последовательности.

В терминах задачи о сопоставлении двух последовательностей могут быть описаны различные проблемы естественной обработки языка, например автоматического перевода [1, 2] или генерация ответных реплик в диалоге [3]. Более того, источником вектора состояния может быть кодировщик любой природы, поэтому подобная архитектура оказалась применима и в таких задачах, как генерация подписи для картинки [4], построение комментария к изменениям в исходном коде программ [5] и составление связанного текста на основе данных таблицы [6].

Использование рекуррентных нейронных сетей обусловлено их способностью сохранять некоторое скрытое состояние, которое характеризует происходящее во входной последовательности. Таким образом обеспечивается обработка предложений произвольной длины [7, с. 232–233].

Уже в работе [8] отмечается основной недостаток подобной архитектуры – необходимость упаковывать произвольное количество информации, содержащейся в исходной последовательности, в вектор фиксированной длины. С увеличением длины входного предложения эффективность построенной модели снижается. Авторы предложили усовершенствовать систему, внедрив механизм внимания. Идея подобного механизма вдохновлена человеческой природой: известно, что при взгляде на рисунок человек не анализирует его полностью во всех подробностях, а фокусируется на отдельных участках, которые, как ему кажется, несут наибольшее количество полезной информации [7, с. 331–332]. Попытки имитировать внимание предпринимались в машинном обучении при анализе изображений, определение значимых участков картинки позволяло сэкономить время на обработке незначимых, периферийных фрагментов [7, с. 333]. Эта концепция переносится и на задачи обработки естественного языка: от механизма внимания требуется оценить, какие составляющие длинного исходного предложения являются наиболее существенными при решении текущей задачи.

Предложенный механизм был реализован следующим образом. Кодировщик передает декодировщику не единый вектор состояния, а множество векторов-аннотаций, построенных с помощью двунаправленной рекуррентной нейронной сети [9] для каждого слова исходного предложения. Декодировщик в ходе своей работы вычисляет актуальный вектор контекста как взвешенную сумму полученных от кодировщика векторов. Используемые при суммировании

веса – аналог внимания: они вычисляются на основе текущего внутреннего состояния рекуррентной нейронной сети декодировщика для каждого вектора-аннотации и определяют, насколько значимо соответствующее этому вектору слово на данном этапе перевода. Предложенная модель показала значительный прирост в качестве перевода длинных предложений.

Исследование нейронного машинного перевода с использованием механизма внимания проводилось в ряде работ. Например, в работе [10] рассмотрены два архитектурных решения. Первое основано на подходе, представленном в [8], и отличается от него лишь деталями. Второе решение, названное авторами локальным вниманием (local attention), предполагает построение текущего вектора контекста, основываясь на подмножестве наиболее значимых векторов-аннотаций. В качестве наиболее значимых предполагается рассматривать векторы-аннотации, соответствующие словам $[p_t - D, p_t + D]$, где p_t – центральный элемент окна, выбираемый с помощью обученного компонента системы, а D – фиксированная ширина окна. Кроме того, было отмечено, что локально близкие слова в целевом предложении с большой долей вероятности локально близки и в исходной. Поэтому на каждом шаге перевода при вычислении очередных весов внимания имеет смысл учитывать веса внимания предыдущего шага. В работе [11] была рассмотрена возможность посимвольного перевода, в [12] исследователи из Facebook применили нейронные сети с механизмом внимания для автореферирования предложений, в [13, 14] внимание используется для решения задачи распознавания речи, в [15] – для построения модели, способной отвечать на вопросы. Наконец, в работе [16] были описаны детали реализации Google's Neural Machine Translation – системы с вниманием, которая в 2016 г. легла в основу сервиса Google Translate.

С ростом научного интереса к механизму внимания приходило понимание большого потенциала данного архитектурного решения. Основная масса работ использовала его в связке с рекуррентными нейронными сетями, в частности с LSTM-ячейками [17] или их модификацией GRU [18]. Однако вычисления в рекуррентных нейронных сетях имеют низкий уровень параллелизма, поскольку сама архитектура требует последовательной обработки поступающих единиц [19]. Эта проблема мотивировала поиск путей усовершенствования подхода. Например, в работах [20, 21] вместо рекуррентных нейронных сетей использовались одномерные сверточные сети (convolutional neural network) [22]. В статье [23] исследователи из Google применили механизм внимания для декомпозиции задач при решении проблемы сопоставления двух утверждений в терминах «следствие», «противоречие» или «нейтральная связь». Авторы показали, что подобная архитектура допускает высокий уровень параллелизации вычислений, а качество полученных результатов не ниже достигнутых с помощью других методов.

Можно сказать, что механизм внимания стал результатом естественного развития архитектуры seq2seq. Текст – сложная структура, в которой удаленность двух токенов друг от друга не является исчерпывающим фактором, позволяющим однозначно определять степень их взаимосвязи. Так, например, для разрешения анафоры, т. е. для определения значения встреченного в тексте местоимения, необходимо установить, на какой именно объект оно указывает. Однако определяющее выражение может располагаться на значительном удалении от местоимения, в том числе в другом предложении. Поэтому для качественного решения задач обработки естественного языка необходим более глубокий анализ входных последовательностей. В частности, таковым мог быть переход от линейных входных последовательностей к более сложным конструкциям, таким как синтаксические деревья, выражающие зависимость одних членов предложения от других. Механизм внимания, однако, более информативен, поскольку предоставляет численное выражение связи между двумя токенами, тогда как в древовидной структуре зависимость может быть выражена только наличием либо отсутствием ребра между компонентами. Кроме того, механизм внимания более абстрактен и может быть применен в иных сферах машинного обучения, например при обработке изображений. В том числе его можно адаптировать к другим формам входных данных, не представляющих собой упорядоченную последовательность или матрицу.

С другой стороны, рассчитанные с помощью механизма внимания оценки связи между токенами зависят от качества обучающей выборки и эффективности самого процесса обучения. Таким образом, внимание является дополнительным набором параметров в архитектуре нейронной сети, подбираемых в ходе обработки обширного массива текстов, а не некоторой попыткой

смоделировать внимание человека. Здесь можно провести параллель с системами автоматического перевода: нейронные системы показали лучшую эффективность, чем архитектуры, основанные на правилах и словарях, однако с точки зрения своего устройства и процесса обучения они имеют мало общего с процессом имитации переводческой деятельности человека.

Архитектура Transformer. В 2017 г. сотрудники Google представили публикацию [24], в которой также отказались от использования рекуррентных нейронных сетей. Авторы предложили архитектуру Transformer, основанную на механизме внимания. Кратко принцип работы предложенной архитектуры изложен ниже.

Кодирующий компонент Transformer состоит из стека идентичных по структуре кодировщиков, аналогично при декодировании используется стек декодеров одинакового строения. Глубина обоих стеков совпадает, в публикации она равна шести единицам.

При обработке предложения первый кодировщик получает векторы, соответствующие входящим в него словам. Каждый вектор строится как сумма векторного представления слова и вектора, кодирующего положение рассматриваемого слова в предложении. Для каждого входящего вектора, представляющего слово w_i , путем умножения на обученные матрицы строится тройка векторов: вектор запроса q_i , вектор ключа k_i и вектор значения v_i . Скалярное произведение вектора запроса q_i и вектора ключа слова k_j рассматривается как коэффициент внимания (self-attention), которое следует уделить слову w_j при анализе слова w_i . Результирующий вектор для слова w_i вычисляется как взвешенная сумма векторов значений, где в качестве весов используются нормализованные коэффициенты внимания.

Подобное вычисление проводится параллельно несколько раз (в публикации [24] – восемь) с использованием различных наборов матриц для генерации q_i , k_i и v_i . Данный подход (multi-head attention) позволяет смоделировать различные аспекты внимания. Например, одна цепочка вычислений акцентирует внимание на семантической составляющей слов, что, в частности, позволит справиться с проблемой разрешения анафор [25], другая – на грамматической составляющей, что поможет сохранить связь в словосочетаниях. Векторы, полученные в каждой цепочке вычислений, конкатенируются в один, который после нормализации (layer normalization) [26] передается во второй слой кодировщика – нейронную сеть прямого распространения.

Вывод второго слоя кодировщика – новые векторные представления слов предложения, которые вновь нормализуются и передаются на вход следующему кодировщику в стеке. Аналогичный процесс повторяется для всех оставшихся кодировщиков.

Процесс декодирования близок к процессу кодирования и проходит в три этапа. На первом этапе применяется механизм внимания, аналогичный работе кодировщика, но с естественным ограничением: в качестве входных слов целевого предложения используются уже построенные слова. Второй этап – повторное применение механизма внимания, однако для построения вектора ключа и вектора запроса используются результаты работы стека кодировщиков. Наконец, третий этап – использование нейронной сети прямого распространения. Таким образом последовательно обрабатывают все декодировщики. Вывод последнего декодировщика проходит через линейный слой, который строит новый вектор. Размерность его равна размеру словаря, известной модели. После нормализации значения каждой компоненты такого вектора интерпретируются как условные вероятности того, что соответствующее компоненте слово должно оказаться на следующей позиции строящегося целевого предложения.

Эксперименты показали, что архитектура Transformer может обучаться значительно быстрее, чем решения с использованием рекуррентных и сверточных нейронных сетей, достигая при этом более высоких результатов.

Архитектура Transformer получила широкое распространение в области естественного языка. В частности, она применялась для генерации речи [27], автоматического реферирования текста [28], поддержания диалога [29], решения текстовых математических задач [30]. Оказалось, что подобная архитектура применима в том числе за пределами задач обработки естественного языка: на базе Transformer были построены система рекомендаций [31] и генератор музыки [32].

Были предложены также усовершенствования архитектуры [33–36]. В 2020 г. компания Microsoft объявила о создании модели на базе Transformer, содержащей 17 млрд параметров и предназначенной для синтеза речи (URL: <https://www.microsoft.com/en-us/research/blog/turing-nlg-a-17-billion-parameter-language-model-by-microsoft/>).

Языковые модели. До 2017 г. в сфере обработки естественного языка в качестве переиспользуемых компонентов обычно выступали предобученные векторные представления слов, в частности word2vec [37] и GloVe [38]. В их основе лежало явление, известное в литературе под названием «дистрибутивная гипотеза» [39]. Согласно этому явлению схожие по своему значению слова встречаются в схожих контекстах. Соответственно, механизм работы word2vec и GloVe – обучение на большом корпусе текстов системы, способной ставить в соответствие слову вектор, некоторым образом характеризующий множество контекстов, в которых это слово встречается. Таким образом, согласно гипотезе схожим по смыслу словам будут сопоставлены близкие векторы. Использование word2vec или GloVe в качестве первого слоя нейронной сети повысило качество решения различных задач, отмечают авторы в работе [40]. Тем не менее значительное количество параметров, регулирующих остальную часть системы, требовалось обучать с нуля. Кроме того, строгое сопоставление слова и вектора имело свои недостатки. В частности, как подчеркнуто в [41], эти модели строят единый вектор для многозначных слов (например, слово «киви» может обозначать фрукт, птицу, а в некоторых контекстах использоваться для обозначения жителей Новой Зеландии), что негативно сказывается на качестве работы.

В то же время в другой области применения машинного обучения – компьютерном зрении – доказал свою эффективность прием переноса знаний (transfer learning) [42, с. 43–44] между моделями. Результаты показали, что если обучить сверточную нейронную сеть классифицировать изображения на большом объеме данных, например датасете ImageNet [43], то переиспользование этой сети с небольшим количеством дополнительных слоев позволяет быстро достичь высоких результатов при решении других задач, в частности при определении позы человека на фотографии [44] и обнаружении объектов на картинке [45]. Подход оказался настолько успешен, что в скором времени использование предобученных нейронных сетей де-факто стало стандартом в исследованиях, связанных с распознаванием изображений [46].

Попытки применить схожий подход к обработке естественного языка также оказались успешными. В работе [47] было применено обучение с частичным привлечением учителя (semi-supervised learning): на первом шаге система, имеющая архитектуру seq-2-seq, обучалась без учителя, на втором – дообучалась под задачу классификации. В качестве первого шага авторы предложили два приема: предсказание следующего слова в предложении и самодекодирование, при котором декодировщику требовалось восстановить из контекстного вектора исходное предложение. В работе [40] модель seq-2-seq обучалась автоматическому переводу, после чего обученные кодировщики дообучались под решение других задач. Авторы рассматривали векторы, построенные кодировщиками, как контекстные (context vectors) – альтернативу фиксированным векторным представлениям слов word2vec и GloVe, учитывающую контекст встреченных слов. В статье [48] подобные контекстные векторы обучались с помощью двунаправленной рекуррентной нейронной сети. Идея получила свое развитие в основанных на рекуррентных нейронных сетях моделях ULMFiT [49] и ELMo (Embeddings from Language Models) [50], которые показали лучшие для своего времени результаты на целом классе различных задач обработки естественного языка. В частности, разработчики ULMFiT сравнили свою предобученную на корпусе WikiText-103 [51] модель с нейронной сетью, обучающейся на размеченных данных с нуля. Эксперименты показали, что для достижения равной точности при анализе тональности текста модели ULMFiT требуется в 10 раз меньше размеченных данных, чем непредобученной системе; для определения темы текста – в 20 раз меньше.

В 2018 г. команда OpenAI представила языковую модель Generative Pre-Training (GPT) [52], основанную на архитектуре Transformer с единым стеком преобразователей, которая была предложена в [53]. Для тренировки GPT также применялось обучение с частичным привлечением учителя. Первый шаг в обучении модели – предобучение без учителя: система тренировалась предсказывать следующее слово в предложении на большом корпусе неразмеченных текстов. На втором шаге происходила тонкая настройка (fine-tuning): к модели добавлялась еще одна

нейронная сеть, веса которой дообучались под конкретную задачу. С помощью предложенного подхода авторам удалось превзойти лучшие для своего времени результаты сразу по нескольким задачам: ответы на вопросы по входным данным (reading comprehension), проверка грамматичности (linguistic acceptability), оценка семантической схожести (semantic similarity) и проверка двух текстов на логическое соответствие (textual entailment).

Связь между значением слова и контекстом, в котором оно употреблено, очевидна. Эта связь легла в основу методов построения векторного представления слов word2vec и GloVe. Можно сказать, что языковые модели развивают данную идею, поскольку они строят контекстно-зависимые векторные представления слов. В ходе обучения языковые модели обрабатывают огромное количество различных текстов и вычлениают разнообразные зависимости между токенами. Получив общее представление о том, как обычно токены располагаются в тексте, модель может эффективно дообучаться под конкретные задачи. Можно выделить два взаимообусловленных тезиса: с одной стороны, возможность переноса знаний позволяет использовать для обучения те задачи, для которых требуется неразмеченный корпус данных; с другой стороны, возможность использовать неразмеченные данные позволяет обучаться на практически неограниченном количестве текстов, повышая качество модели и эффективность переноса знаний. Работу языковых моделей можно сравнить с решением логических задач, в которых в качестве актантов и предикатов используются вымышленные лексемы, что, однако, не мешает сделать вывод об истинности или ложности некоторого высказывания.

Модель BERT. Компания Google предложила свою модель языка, получившую название BERT [54]. BERT также базируется на архитектуре Transformer и во многом схожа с GPT. Принципиальным ее отличием является метод обучения: вместо предсказания последующего слова в последовательности модель BERT на стадии предобучения тренировалась определять закрытые маской слова в предложении. Впервые такая задача была предложена в работе [55]. Второй задачей на этапе предобучения было определение того, следовали ли два предложения в тексте одно за другим. В качестве источников неразмеченных данных в работе выступили корпус BookCorpus [56] и англоязычная Википедия.

Для проведения сравнительного анализа авторы разработали модель BERT-base, которая сопоставима по размерам и производительности с GPT. Вторая модель, BERT-large, имеет в три раза больше параметров – около 340 млн.

BERT показала впечатляющий результат, оказавшись лучшей в мире моделью для решения 11 различных задач. BERT-base по всем задачам оказалась эффективнее, чем GPT, что подтвердило эффективность используемого для предобучения подхода. На многозадачном тесте GLUE (General Language Understanding Evaluation) [57], цель которого – оценить понимание прочитанного компьютером, BERT-base и BERT-large набрали соответственно 79,6 и 82,1 балла из 100, тогда как предыдущим наивысшим достижением было 75,1 балла у GPT.

Модель BERT была выложена в открытый доступ, и в скором времени на ее основе появилось множество новых моделей:

- разработанная в Facebook модель RoBERTa (Robustly optimized BERT approach) [58], для которой был усовершенствован процесс обучения, в том числе увеличен объем неразмеченных данных;

- ALBERT (a lite BERT) [59], созданная совместно сотрудниками Google Research и Toyota Technological Institute, содержащая меньшее число параметров, чем оригинальная BERT, но при этом обучающаяся эффективнее;

- DistilBERT [60] – подвергнутая «дистилляции» (distillation) [61] BERT, которая имеет на 40 % меньше параметров и на 60 % быстрее работает, сохраняя при этом 97 % от качества работы исходной модели;

- TinyBERT [62] от Huawei – еще одна «дистиллированная» версия BERT;

- MT-DNN [63] от Microsoft, которая представляет усовершенствование многозадачной модели, предложенной в работе [64], с использованием BERT в качестве единого компонента. Эта модель на тесте GLUE превзошла средний человеческий результат (URL: <https://docs.microsoft.com/archive/blogs/stevengu/microsoft-achieves-human-performance-estimate-on-glue-benchmark>);

– StructBERT [65] от разработчиков Alibaba, при обучении которой использовались задачи предсказания порядка слов и предложений [66]. Таким образом авторы стремились сообщить модели больше информации о базовых языковых структурах;

– BioBERT [67], предназначенная для работы с текстами на биомедицинскую тематику;

– ViBERT (vision-and-language BERT) [68] – расширенная модель BERT, которая работает с парами «изображение – текст».

В мае 2020 г. среди лидеров в тесте GLUE большинство составляют модели, основанные на BERT.

Итак, можно выделить две ключевые идеи, лежащие в основе модели BERT: механизм внимания, на котором построена архитектура Transformer, и характерный для языковых моделей принцип переноса знаний. Как утверждалось выше, языковые модели обучаются искать закономерности между токенами в корпусе текстов. Механизм внимания ориентирован на поиск взаимосвязей между токенами в конкретных входных данных. Совершенно естественно, что два подхода соединились в эффективный ансамбль, ставший своего рода прорывом в сфере обработки естественного языка. Практика показывает, что увеличение числа параметров в архитектуре модели и увеличение обучающей выборки ведет к построению еще более качественных моделей. Развитие науки в этом направлении можно описать как итеративный процесс, включающий два этапа: создание более крупной и эффективной языковой модели, демонстрирующей лучшие результаты в сфере обработки естественного языка, и разработку более простых и быстрых моделей, способных продемонстрировать схожий уровень.

Слабые стороны BERT. Следует отметить, что, несмотря на очень высокий процент правильных решений серии различных задач, BERT и другие языковые модели требуют дальнейшего усовершенствования. Как показывают исследования, после обучения эти системы ориентируются в том числе на ложные эвристики, обусловленные неудачным подбором данных в используемых датасетах [69]. Хотя BERT-система решила задачу понимания аргументации (argument reasoning comprehension) с точностью 77 %, что всего на 3 % меньше среднего человеческого уровня, авторы работы [69] утверждают, что в области понимания аргументации BERT не обучается ничему. При тестировании систем на новых, специально подготовленных тестовых данных эффективность языковых моделей оказывается существенно ниже [70, 71].

Представление о том, как слова взаимосвязаны друг с другом, позволяет на высоком уровне взаимодействовать с текстом и решать большинство задач естественной обработки языка. Тем не менее по своей сути это скорее копирование и вставка информации, а не понимание.

Работа BERT и других языковых моделей напоминает работу персонажа мысленного эксперимента о китайской комнате. В этом мысленном эксперименте, предложенном философом Джоном Сёрлом, описывается человек, не знающий китайских иероглифов. Он находится в запертой комнате, в которой также имеется подробная инструкция по манипуляции иероглифами. Вне комнаты находится наблюдатель, который через щель передает в комнату некоторое сообщение на китайском языке. Находящийся в комнате человек получает это сообщение и в соответствии с инструкцией перерисовывает некоторый ответ, который возвращает наблюдателю. В указанных условиях у наблюдателя может сложиться представление, что человек в комнате владеет китайским языком. На самом же деле он не имеет ни малейшего представления о теме разговора и просто выполняет инструкции.

Примечательно, что представление о мире, полученное языковыми моделями в ходе обучения, в некотором роде перекликается с постмодернистской концепцией Жака Деррида о мире как совокупности текстов. BERT, как и другие языковые модели, теоретически может извлекать из корпуса текстов знания: например, попытка предсказать последнее слово в фразе «самой быстрой птицей на свете является» будет обусловлена закономерностью в использовании слов, но может привести к верному ответу. Тем не менее понимание в том смысле, в котором этот термин применим к человеку, для автоматической языковой модели едва ли достижимо. Одной из проблем является интерпретация дейктических единиц: завершение предложения «в настоящий момент президентом США является» оказывается гораздо более сложной задачей, поскольку в корпусе текстов разных лет схожая фраза может завершаться различным образом.

«Представление» машины о мире можно несколько расширить с помощью взаимодействия системы с реальными или абстрактными явлениями и объектами. В этом случае знания модели

будут включать не только взаимосвязь токенов между собой, но и взаимосвязь отдельных токенов с объектами мира. В некотором смысле это будет сравнимо с изучением языка ребенком, который обучается ему параллельно с исследованием предметов окружающей действительности.

Вторая проблема заключается в том, что для предварительного обучения языковой модели BERT требуются огромные объемы текстов: в оригинальной работе были использованы корпус BookCorpus и англоязычная Википедия, в последующих публикациях (см., к примеру, [58]) множество текстов было расширено. Данный подход не применим в случае, когда отобрать настолько обширную коллекцию текстов для текущей задачи невозможно. Такая проблема обнаруживается, например, при обработке малых языков, не имеющих широкого использования в Интернете и литературной традиции. Для решения задачи компьютерной обработки подобных малых языков следует разработать иные методы. В частности, научный интерес представляет эффективность, с которой ребенок изучает родной язык. Успешное моделирование онтогенеза языка могло бы обеспечить построение качественных моделей малых языков.

Заключение. Модель BERT опирается на использование механизма внимания и принцип переноса знаний. Работая совместно, эти две идеи позволяют эффективно обнаруживать закономерности между применением слов и токенов в огромной обучающей выборке размеченных текстов. Обнаруженных закономерностей оказывается достаточно, чтобы эффективно решать разнообразные задачи, связанные с автоматической обработкой естественного языка. Более того, качество работы можно увеличивать, усложняя архитектуру модели и расширяя обучающую выборку.

Главный недостаток BERT и подобных моделей заключается в том, что понимание имитируется за счет закономерностей, найденных в текстах, в том числе за счет ложных эвристик. Вторым недостатком вытекает из первого: для качественной имитации понимания требуется длительное обучение на огромной выборке.

Отметим, что оба недостатка обусловлены архитектурой модели, поэтому BERT не может быть доработана с целью их исправления. Чтобы избежать упомянутых негативных моментов, следует рассматривать принципиально иные подходы к моделированию автоматических обработчиков естественного языка. Таким подходом, в частности, может быть моделирование онтогенеза языка – процесса обучения ребенка родному языку.

References

1. Cho K., Merriënboer B. van, Gulcehre C., Bahdanau D., Bougares F., Schwenk H., Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014*, pp. 1724–1734. <https://doi.org/10.3115/v1/D14-1179>
2. Sutskever I., Vinyals O., Le Q. V. Sequence to sequence learning with neural networks. *Neural Information Processing Systems*, 2014, pp. 3104–3112. Available at: <https://arxiv.org/abs/1409.3215> (accessed 07.07.2020).
3. Serban I. V., Lowe R., Charlin L., Pineau J. Generative deep neural networks for dialogue: A short review. *Neural Information Processing Systems, Workshop on Learning Methods for Dialogue*, 2016. Available at: <https://arxiv.org/abs/1611.06216> (accessed 07.07.2020).
4. Vinyals O., Toshev A., Bengio S., Erhan D. Show and tell: A neural image caption generator. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, Massachusetts, USA, 7–12 June 2015*, pp. 3156–3164. <https://doi.org/10.1109/CVPR.2015.7298935>
5. Loyola P., Marrese-Taylor E., Matsuo Y. A Neural architecture for generating natural language descriptions from source code changes. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada, 30 July – 4 August 2017*, vol. 2, pp. 287–292. <https://doi.org/10.18653/v1/P17-2045>
6. Lebrecht R., Grangier D., Auli M. Neural text generation from structured data with application to the biography domain. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas, USA, 1–5 November 2016*, pp. 1203–1213. <https://doi.org/10.18653/v1/D16-1128>
7. Nikolenko S., Kandurin A., Arhangelskaja E. Glubokoe obuchenie. *Deep Learning*. Saint Petersburg, Piter, 2020, 480 p. (in Russian).

8. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate. *International Conference on Learning Representations*, 2015. Available at: <https://arxiv.org/abs/1409.0473> (accessed 07.07.2020).
9. Schuster M., Paliwal K. K. Bidirectional recurrent neural networks. *Signal Processing, IEEE Transactions*, 1997, vol. 45(11), pp. 2673–2681. <https://doi.org/10.1109/78.650093>
10. Luong T., Pham H., Manning C. D. Effective approaches to attention-based neural machine translation. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015*, pp. 1412–1421. <https://doi.org/10.18653/v1/D15-1166>
11. Chung J., Cho K., Bengio Y. A character-level decoder without explicit segmentation for neural machine translation. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 7–12 August 2016*, vol. 1, pp. 1693–1703. <https://doi.org/10.18653/v1/P16-1160>
12. Rush A., Chorpa S., Weston J. A neural attention model for abstractive sentence summarization. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015*, pp. 379–389. <https://doi.org/10.18653/v1/D15-1044>
13. Chorowski J., Bahdanau D., Serdyuk D., Cho K., Bengio Y. Attention-based models for speech recognition. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2015, vol. 1, pp. 577–585. Available at: <https://arxiv.org/abs/1506.07503> (accessed 07.07.2020).
14. Chan W., Jaitly N., Le Q. V., Vinyals O. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, China, 20–25 March 2016*, pp. 4960–4964. <https://doi.org/10.1109/ICASSP.2016.7472621>
15. Hermann K. M., Kočiský T., Grefenstette E., Espeholt L., Kay W., Suleyman M., Blunsom P. Teaching machines to read and comprehend. *Neural Information Processing Systems 28: 29th Annual Conference on Neural Information Processing Systems*, 2015, pp. 1693–1701. Available at: <https://arxiv.org/abs/1506.03340> (accessed 07.07.2020).
16. Wu Y., Schuster M., Chen Z., Le Q. V., Norouzi M., ..., Dean J. *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*, 2016. Available at: <https://arxiv.org/abs/1609.08144> (accessed 07.07.2020).
17. Hochreiter S., Schmidhuber, J. Long short-term memory. *Neural Computation*, 1997, vol. 9(8), pp. 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
18. Cho K., Merriënboer B. van, Bahdanau D., Bengio Y. On the properties of neural machinetranslation: Encoder-decoder approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, Doha, Qatar, 25 October 2014*, pp. 103–111. <https://doi.org/10.3115/v1/W14-4012>
19. Martin E., Cundy C. Parallelizing linear recurrent neural nets over sequence length. *International Conference on Learning Representations*, 2018. Available at: <https://arxiv.org/abs/1709.04057> (accessed 07.07.2020).
20. Kalchbrenner N., Espeholt L., Simonyan K., Oord van den A., Graves A., Kavukcuoglu K. *Neural Machine Translation in Linear Time*, 2016. Available at: <https://arxiv.org/abs/1610.10099> (accessed 07.07.2020).
21. Gehring J., Auli M., Grangier D., Yarats D., Dauphin Y. N. Convolutional sequence to sequence learning. *Proceedings of the 34th International Conference on Machine Learning*, 2017, vol. 70, pp. 1243–1252. Available at: <https://arxiv.org/abs/1705.03122> (accessed 07.07.2020).
22. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, vol. 86(11), pp. 2278–2324. <https://doi.org/10.1109/5.726791>
23. Parikh A. P., Täckström O., Das D., Uszkoreit J. A decomposable attention model for natural language inference. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas, USA, 1–5 November 2016*, pp. 2249–2255. <https://doi.org/10.18653/v1/D16-1244>
24. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., ..., Polosukhin I. Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 4–9 December 2017*, pp. 6000–6010. Available at: <https://arxiv.org/abs/1706.03762> (accessed 07.07.2020).
25. Mitkov R. *Anaphora Resolution: The State of the Art*. Paper based on the COLING'98/ACL'98 tutorial on anaphora resolution. Wolverhampton, University of Wolverhampton, 1999, 34 p.
26. Ba J. L., Kiros J. R., Hinton G. E. *Layer Normalization*, 2016. Available at: <https://arxiv.org/abs/1607.06450> (accessed 07.07.2020).
27. Li N., Liu S., Liu Y., Zhao S., Liu M., Zhou M. Neural speech synthesis with transformer network. *The AAAI Conference on Artificial Intelligence*, 2019. Available at: <https://arxiv.org/abs/1809.08895> (accessed 07.07.2020).
28. Khandelwal U., Clark K., Jurafsky D., Kaiser Ł. *Sample Efficient Text Summarization using a Single Pre-Trained Transformer*, 2019. Available at: <https://arxiv.org/abs/1905.08836> (accessed 07.07.2020).

29. Vlasov V., Mosig J. E. M., Nicho A. *Dialogue Transformers*, 2019. Available at: <https://arxiv.org/abs/1910.00486> (accessed 07.07.2020).
30. Griffith K., Kalita J. Solving arithmetic word problems automatically using transformer and unambiguous representations. *International Conference on Computational Science and Computational Intelligence, Las Vegas, USA, 5–7 December 2019*, pp. 526–532. <https://doi.org/10.1109/CSCI49370.2019.00101>
31. Kang W.-C., McAuley J. Self-attentive sequential recommendation. *IEEE International Conference on Data Mining, Singapore, 17–20 November 2018*, pp. 197–206. <https://doi.org/10.1109/ICDM.2018.00035>
32. Huang C.-Z. A., Vaswani A., Uszkoreit J., Shazeer N., Simon I., ..., Eck D. *Music Transformer*, 2018. Available at: <https://arxiv.org/abs/1809.04281> (accessed 07.07.2020).
33. Dehghani M., Gouws S., Vinyals O., Uszkoreit J., Kaiser Ł. Universal transformers. *7th International Conference on Learning Representations*, 2019. Available at: <https://arxiv.org/abs/1807.03819> (accessed 07.07.2020).
34. Dai Z., Yang Z., Yang Y., Carbonell J., Le Q. V., Salakhutdinov R. Transformer-XL: attentive language models beyond a fixed-length context. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July – 2 August 2019*, pp. 2978–2988. <https://doi.org/10.18653/v1/P19-1285>
35. So D. R., Liang C., Le Q. V. The evolved transformer. *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 5877–5886. Available at: <https://arxiv.org/abs/1901.11117> (accessed 07.07.2020).
36. Zhao C., Xiong C., Rosset C., Song X., Bennett P., Tiwary S. Transformer-XH: multi-evidence reasoning with eXtra hop attention. *8th International Conference on Learning Representations*, 2020. Available at: <https://openreview.net/forum?id=r1eLiCNYwS> (accessed 07.07.2020).
37. Mikolov T., Chen K., Corrado G., Dean J. Distributed representations of words and phrases and their compositionality. *Proceedings of the 26th International Conference on Neural Information Processing Systems*, 2013, vol. 2, pp. 3111–3119. Available at: <https://arxiv.org/abs/1310.4546> (accessed 07.07.2020).
38. Pennington J., Socher R., Manning C. D. Glove: global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014*, pp. 1532–1543. <https://doi.org/10.3115/v1/D14-1162>
39. Sahlgren M. The distributional hypothesis. From context to meaning. Distributional models of the lexicon in linguistics and cognitive science (special issue of the Italian Journal of Linguistics). *Rivista di Linguistica*, 2008, vol. 20(1), pp. 33–53.
40. McCann, B., Bradbury J., Xiong C., Socher R. Learned in translation: contextualized word vectors. *31st Conference on Neural Information Processing Systems*, 2017, pp. 6297–6308. Available at: <https://arxiv.org/abs/1708.00107> (accessed 07.07.2020).
41. Hedderich M. A., Yates A., Klakow D., Melo G. de. Using Multi-Sense Vector embeddings for reverse dictionaries. *Proceedings of the 13th International Conference on Computational Semantics – Long Papers, Gothenburg, Sweden, 23–27 May 2019*, pp. 247–258. <https://doi.org/10.18653/v1/W19-0421>
42. Ruder S. *Neural Transfer Learning for Natural Language Processing. Ph. D. Thesis*. Galway, National University of Ireland, 2019, 329 p.
43. Deng J., Dong W., Socher R., Li L.-J., Li K., Fei-Fei L. ImageNet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009*, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
44. Papandreou G., Zhu T., Kanazawa N., Toshev A., Tompson J., Bregler C., Murphy K. Towards accurate multi-person pose estimation in the wild. *IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*, pp. 3711–3719. <https://doi.org/10.1109/CVPR.2017.395>
45. He K., Gkioxari G., Dollár P., Girshick R. Mask R-CNN. *IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*, pp. 2980–2988. <https://doi.org/10.1109/ICCV.2017.322>
46. Mahajan D., Girshick R., Ramanathan V., He K., Paluri M., ..., Maaten L. van der. Exploring the limits of weakly supervised pretraining. *European Conference on Computer Vision, Munich, Germany, 8–14 September 2018*, pp. 181–196. https://doi.org/10.1007/978-3-030-01216-8_12
47. Dai A. M., Le Q. V. Semi-supervised sequence learning. *Neural Information Processing Systems 28: 29th Annual Conference on Neural Information Processing Systems 2015, Montreal, Canada, 7–12 December 2015*, vol. 2, pp. 3079–3087. <https://doi.org/10.18653/v1/P17-1161>
48. Peters M. E., Ammar W., Bhagavatula C., Power R. Semi-supervised sequence tagging with bidirectional language models. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 2017, vol. 1, pp. 1756–1765. Available at: <https://arxiv.org/abs/1705.00108> (accessed 07.07.2020).
49. Howard J., Ruder S. Universal language model fine-tuning for text classification. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 15–20 July 2018*, vol. 1, pp. 328–339. <https://doi.org/10.18653/v1/P18-1031>

50. Peters M. E., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L. Deep contextualized word representations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, New Orleans, Louisiana, USA, 1–6 June 2018*, vol. 1, pp. 2227–2237. <https://doi.org/10.18653/v1/N18-1202>
51. Merity S., Xiong C., Bradbury J., Socher R. Pointer sentinel mixture models. *5th International Conference on Learning Representations*, 2017. Available at: <https://arxiv.org/abs/1609.07843> (accessed 07.07.2020).
52. Radford A., Narasimhan K., Salimans T., Sutskever I. Improving language understanding with unsupervised learning. *Technical report*, 2018. Available at: <https://openai.com/blog/language-unsupervised/> (accessed 07.07.2020).
53. Liu P. J., Saleh M., Pot E., Goodrich B., Sepassi R., Kaiser L., Shazeer N. Generating wikipedia by summarizing long sequences. *6th International Conference on Learning Representations*, 2018. Available at: <https://arxiv.org/abs/1801.10198> (accessed 07.07.2020).
54. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MN, USA, 2–7 June 2019*, vol. 1, pp. 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
55. Taylor W. L. Cloze procedure: A new tool for measuring readability. *Journalism Bulletin*, 1953, vol. 30(4), pp. 415–433.
56. Zhu Y., Kiros R., Zemel R., Salakhutdinov R., Urtasun R., Torralba A., Fidler S. Aligning books and movies: Towards story-like visual explanations by watching movies and reading books. *Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015*, pp. 19–27. <https://doi.org/10.1109/ICCV.2015.11>
57. Wang A., Singh A., Michael J., Hill F., Levy O., Bowman S. R. GLUE: A multi-task benchmark and analysis platform for natural language understanding. *Proceedings of the 2018 EMNLP Workshop Blackbox NLP: Analyzing and Interpreting Neural Networks for NLP, Brussels, Belgium, 1 November 2018*, pp. 353–355. <https://doi.org/10.18653/v1/W18-5446>
58. Liu Y., Ott M., Goyal N., Du J., Joshi M., ..., Stoyanov V. *RoBERTa: A Robustly Optimized BERT Pretraining Approach*, 2019. Available at: <https://arxiv.org/abs/1907.11692> (accessed 07.07.2020).
59. Lan Z., Chen M., Goodman S., Gimpel K., Sharma P., Soricut R. ALBERT: a lite BERT for self-supervised learning of language representations. *8th International Conference on Learning Representations*, 2020. Available at: <https://openreview.net/forum?id=H1eA7AEtvS> (accessed 07.07.2020).
60. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *Conference on Neural Information Processing Systems*, 2019. Available at: <https://arxiv.org/abs/1910.01108> (accessed 07.07.2020).
61. Hinton G., Vinyals O., Dean J. Distilling the knowledge in a neural network. *Neural Information Processing Systems. Deep Learning and Representation Learning Workshop*, 2015. Available at: <https://arxiv.org/abs/1503.02531> (accessed 07.07.2020).
62. Jiao X., Yin Y., Shang L., Jiang X., Chen X., ..., Liu Q. *TinyBERT: Distilling BERT for Natural Language Understanding*, 2019. Available at: <https://arxiv.org/abs/1909.10351> (accessed 07.07.2020).
63. Liu X., He P., Chen W., Gao J. Multi-task deep neural networks for natural language understanding. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July – 2 August 2019*, pp. 4487–4496. <https://doi.org/10.18653/v1/P19-1441>
64. Liu X., Gao J., He X., Deng L., Duh K., Wang Y.-Y. Representation learning using multi-task deep neural networks for semantic classification and information retrieval. *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, Colorado, USA, 31 May– 5 June 2015*, pp. 912–921. <https://doi.org/10.3115/v1/N15-1092>
65. Wang W., Bi B., Yan M., Wu C., Xia J., ..., Si L. StructBERT: incorporating language structures into pre-training for deep language understanding. *8th International Conference on Learning Representations*, 2020. Available at: <https://openreview.net/forum?id=BJgQ4ISFPH> (accessed 07.07.2020).
66. Elman J. L. Finding structure in time. *Cognitive Science*, 1990, vol. 14(2), pp. 179–211.
67. Lee J., Yoon W., Kim S., Kim D., Kim S., ..., Kang J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 2020, vol. 36(4), pp. 1234–1240. <https://doi.org/10.1093/bioinformatics/btz682>
68. Lu J., Batra D., Parikh D., Lee S. *ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks*, 2019. Available at: <https://arxiv.org/abs/1908.02265> (accessed 07.07.2020).
69. Niven T., Kao H.-Y. Probing neural network comprehension of natural language arguments. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July – 2 August 2019*, pp. 4658–4664. <https://doi.org/10.18653/v1/P19-1459>

70. Zellers R., Holtzman A., Bisk Y., Farhadi A., Choi Y. HellaSwag: Can a machine really finish your sentence? *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July – 2 August 2019*, pp. 4791–4800. <https://doi.org/10.18653/v1/P19-1472>

71. McCoy T., Pavlick E., Linzen T. Right for the wrong reasons: diagnosing syntactic heuristics in natural language inference. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July – 2 August 2019*, pp. 3428–3448. <https://doi.org/10.18653/v1/P19-1334>

Информация об авторе

Качков Дмитрий Ильич, аспирант кафедры много-процессорных систем и сетей факультета прикладной математики и информатики, Белорусский государственный университет, Минск, Беларусь.
E-mail: dmitriydikanskiy@gmail.com

Information about the author

Dzmitry I. Kachkou, Postgraduate Student of Department of Multiprocessor Systems and Networks of the Faculty of Applied Mathematics and Informatics, Belarusian State University, Minsk, Belarus.
E-mail: dmitriydikanskiy@gmail.com

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

УДК 004.912
<https://doi.org/10.37661/1816-0301-2020-17-4-73-82>

Поступила в редакцию 28.10.2020
Received 28.10.2020

Принята к публикации 20.11.2020
Accepted 20.11.2020

Интернет-поиск и лексико-семантическая обработка аналогов принятых решений в различных предметных областях

С. Ф. Липницкий

*Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь
E-mail: lipn@newman.bas-net.by*

Аннотация. Предлагается математическая модель интернет-поиска и лексико-семантической обработки аналогов принятых решений в различных предметных областях. Поиск проводится по запросам, синтезированным из описаний проблемных ситуаций. Процесс синтеза осуществляется в два этапа. На первом этапе предложения из текста, содержащего описание проблемной ситуации, классифицируются с учетом информативности вербальной ассоциации между ними. На втором этапе вычисляется информативность каждого класса. Наиболее информативные из сформированных классов после их индексирования используются в качестве запросов. При лексико-семантической обработке найденные аналоги решений исследуются на тональность. Оценка тональности реализуется путем использования лингвистических словарей тонально окрашенной лексики, которые формируются на основе специальных тонально окрашенных тематических корпусов текстов. В предельном случае создаются два типа словарей. Первый тип предназначен для анализа положительной тональности в описаниях принятых решений, а второй – для анализа отрицательной тональности, т. е. в процессе лексико-семантического анализа рассматриваются главным образом положительные и отрицательные аспекты принятых решений. Результаты анализа предъявляются пользователю (лицу, принимающему решения).

Ключевые слова: интернет-поиск, лингвистические словари, математическая модель, проблемная ситуация, релевантность, синтез запросов, тональная окрашенность

Для цитирования. Липницкий, С. Ф. Интернет-поиск и лексико-семантическая обработка аналогов принятых решений в различных предметных областях / С. Ф. Липницкий // Информатика. – 2020. – Т. 17, № 4. – С. 73–82. <https://doi.org/10.37661/1816-0301-2020-17-4-73-82>

Internet search and lexical-semantic processing of analogs when making decisions in various subject areas

Stanislav F. Lipnitsky

*The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus
E-mail: lipn@newman.bas-net.by*

Abstract. A mathematical model of Internet search and lexical-semantic processing of analogs of the decisions made in various subject areas is proposed. The search is carried out on queries synthesized from descriptions of problem situations. The synthesis process is carried out in two stages. At the first stage, sentences from the text containing a description of the problem situation are classified taking into account the informativeness of the verbal association between them. At the second stage, the information content of each class is calculated. The most informative of the generated classes, after their indexing, are used as queries. In lexical-semantic processing, the found analogs of solutions are examined for sentiment. When assessing sentiment, linguistic dictionaries of tone-colored vocabulary are used, which are formed on the basis of special tone-colored thematic corpora of texts. In the extreme case, two types of dictionaries are created. The first type is intended for the analysis of the positive sentiment in the descriptions of the decisions made, and the second is intended for the analysis of

the negative sentiment, that is, in the process of lexical-semantic analysis, mainly positive and negative aspects of the decisions made are considered. The results of the analysis are presented to the user (decision-maker).

Keywords: internet search, linguistic dictionaries, mathematical model, problem situation, relevance, query synthesis, tonal coloration

For citation. Lipnitsky S. F. Internet search and lexical-semantic processing of analogs when making decisions in various subject areas. *Informatics*, 2020, vol. 17, no. 4, pp. 73–82 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-73-82>

Введение. Процессы принятия решений в различных предметных областях имеют много общего. Как правило, они включают в себя следующие основные этапы:

- описание проблемной ситуации и постановку задачи принятия решения;
- поиск вариантов (альтернатив) решения поставленной задачи;
- выбор критериев оценки альтернатив для описания вариантов решения;
- выявление ограничений на критерии;
- принятие решения с учетом результатов оценки альтернатив.

При вербальном анализе решений используется качественная (нечисловая) информация на всех его этапах [1].

В статье [2] автором предложена математическая модель информационной поддержки процесса принятия решения в части веб-поиска его альтернативных вариантов. Поиск альтернатив реализуется по запросам, синтезированным из описания проблемной ситуации.

В настоящей статье предлагаются модель и алгоритмы лексико-семантической обработки аналогов уже принятых решений, найденных в Интернете по запросам, которые были синтезированы из описаний проблемных ситуаций в соответствии с алгоритмами из работы автора [2]. В процессе лексико-семантического анализа рассматриваются главным образом положительные и отрицательные аспекты принятых решений, т. е. оценивается тональность при их обсуждении. Результаты анализа предъявляются пользователю (лицу, принимающему решения).

Архитектура информационной системы. Функциональными компонентами системы интернет-поиска и лексико-семантической обработки аналогов принятых решений являются четыре подсистемы (рис. 1): синтеза запросов, индексирования запросов, поиска аналогов в Интернете, поиска аналогов решений в Интернете, лексико-семантической обработки аналогов.

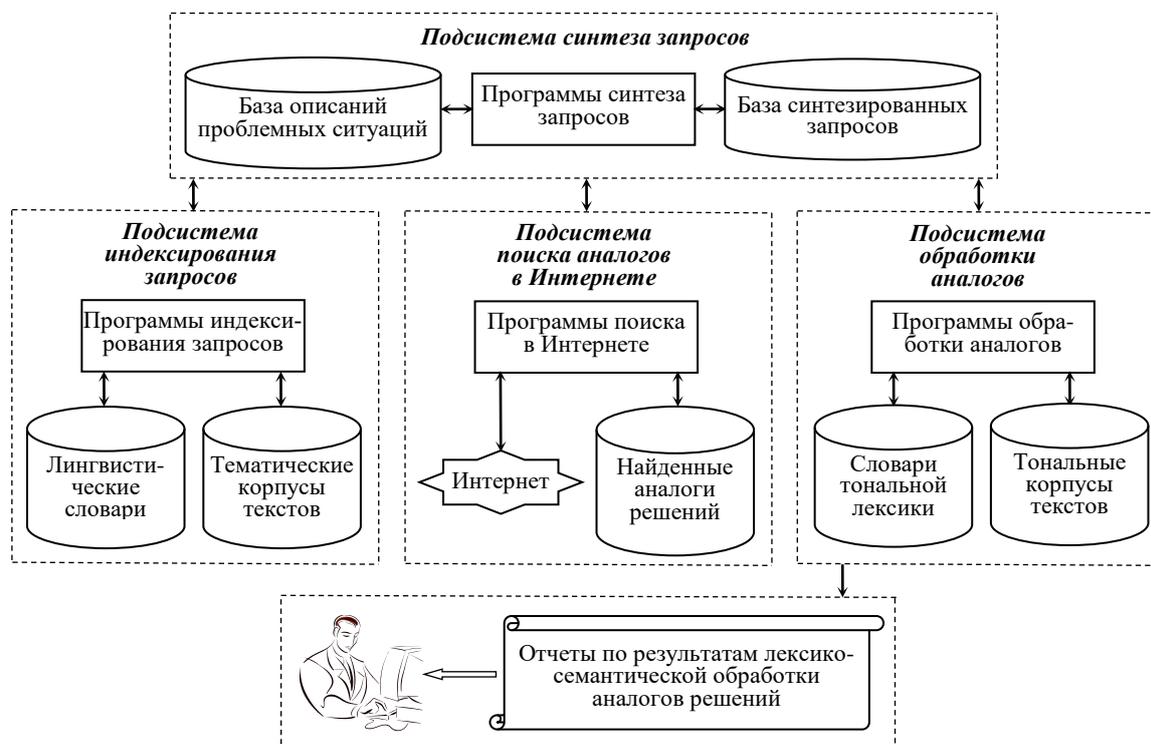


Рис. 1. Структурная схема системы интернет-поиска и лексико-семантической обработки аналогов принятых решений

Подсистема синтеза запросов включает базу описаний проблемных ситуаций при принятии решений, базу синтезированных запросов из этих описаний, а также соответствующие программные средства. Подсистема обеспечивает формирование запросов на основе исследования вербальных ассоциаций между предложениями в описании проблемной ситуации [3, 4]. Предложения из текста описания классифицируются с учетом информативности ассоциации между ними. Далее вычисляется информативность каждого класса. Информативные классы после их индексирования используются в качестве запросов на интернет-поиск аналогов принятых решений.

В состав подсистемы индексирования входит совокупность лингвистических словарей для вычисления информативности слов и вербальных ассоциаций между ними. Словари формируются из специальных наборов публикаций по каждой предметной области – тематических корпусов текстов. Поисковые образы классов предложений представляются в виде множеств слов и вербально-ассоциативных пар слов с соответствующими значениями информативности.

Подсистема поиска аналогов решений в Интернете состоит из специализированных программных агентов, основная задача которых заключается в систематическом получении и накоплении новых данных из Интернета. Поиск аналогов решений реализуется в порядке, определяемом упорядочивающим отношением, которое задается на множестве веб-страниц каждого сканируемого веб-сайта.

Подсистема лексико-семантической обработки аналогов принятых решений имеет в своем составе лингвистические словари тонально окрашенной лексики, а также тонально окрашенные тематические корпуса текстов. Они используются при формировании оценок тональности сообщений. В предельном случае создаются два типа словарей. Первый тип предназначен для анализа положительной тональности в описаниях принятых решений, а второй – для анализа отрицательной тональности.

Синтез запросов. Процесс синтеза запросов в системе интернет-поиска и лексико-семантической обработки аналогов принятых решений реализуется в два этапа. На первом этапе предложения из текста, содержащего описание проблемной ситуации, классифицируются с учетом информативности вербальной ассоциации между ними. На втором этапе вычисляется информативность каждого класса. Информативные классы после их индексирования используются в качестве запросов на интернет-поиск принятых решений.

Обозначим через $T = \langle \rho_1, \rho_2, \dots, \rho_l \rangle$ описание проблемной ситуации, где $\langle \rho_1, \rho_2, \dots, \rho_l \rangle$ – кортеж предложений текста T . При разбиении данного кортежа предложений на классы будем использовать формулы вычисления информативности вербальной ассоциации между словами предложений и между самими предложениями [2].

Информативность вербальной ассоциации между словами. Под вербальными ассоциациями в компьютерной лингвистике понимают семантические связи между словами в языке, соответствующие ассоциативным отношениям между обозначаемыми ими сущностями в реальном мире. Различают два типа таких ассоциаций: парадигматические и синтагматические [5, 6]. Парадигматические ассоциации существуют между словами языка независимо от контекста и объединяют понятия, обозначающие предметы или явления, между которыми имеется постоянная связь (например, пары слов *книга – знание, человек – дом*). В противоположность парадигматическим синтагматические ассоциации возникают в тексте, т. е. между словами и словосочетаниями каждого конкретного его предложения (например, в парах слов *технология – информационная, текст – шрифт*). Считаем, что информативность вербальной ассоциации между произвольными словами a и b некоторого предложения – это вероятность его появления в полном корпусе текстов. При практической реализации информационной системы под указанной информативностью будем понимать дробь

$$I_{Cf}^{ab} = n_{Cf}^{ab} / N_{Cf}, \quad (1)$$

где n_{Cf}^{ab} – количество предложений в полном корпусе текстов Cf , в которых одновременно присутствуют слова a и b или синонимы и словоизменения хотя бы одного из этих слов, а N_{Cf} – количество всех предложений в корпусе Cf .

В развернутом виде формулу (1) можно переписать, используя информацию, которую содержат специальные лингвистические словари:

- словарь вербально-ассоциативных пар слов $Dic_{ab} = \{ \langle (a, b), I_{Cf}^{ab} \rangle \mid a, b \in \pi, \pi \in Cf \}$, в котором каждой паре слов поставлена в соответствие информативность их вербальной ассоциации;
- словарь словоизменительных парадигм $Dic_{par} = \{ (a, Par_a) \mid a \in W_{Cf}, a \in Par_a \}$, состоящий из пар $\langle \text{словоформа}, \text{парадигма} \rangle$. В позиции парадигмы Par_a представлены все словоизменения словоформы a ;
- словарь синонимичных словоформ $Dic_{syn} = \{ (a, Syn_a) \mid a \in W_{Cf}, a \in Syn_a \}$, включающий в себя пары $\langle \text{словоформа}, \text{синонимичные словоформы} \rangle$, в которых каждой словоформе a соответствует множество ее синонимов Syn_a .

С учетом словоизменений и синонимов, зафиксированных в словарях Dic_{par} и Dic_{syn} , формулу (1) перепишем в виде

$$I_{Cf}^{ab} = \frac{n_{Cf}^{ab} + n_{Cf}^{Par_{ab}} + n_{Cf}^{Syn_{ab}}}{N_{Cf}}. \quad (2)$$

Параметр $n_{Cf}^{Par_{ab}}$ в формуле (2) указывает на число вхождений всех пар словоформ, являющихся словоизменениями соответственно слов a и (или) b и встречающихся в одном и том же предложении корпуса текстов Cf :

$$n_{Cf}^{Par_{ab}} = \sum_{\substack{c \in Par_a, d \in Par_b, \\ c \neq a \text{ и (или) } d \neq b, \\ (c, d) \in \Theta}} n_{Cf}^{cd}.$$

Аналогичное выражение справедливо для параметра $n_{Cf}^{Syn_{ab}}$:

$$n_{Cf}^{Syn_{ab}} = \sum_{\substack{d \in Syn_a, f \in Syn_b, \\ d \neq a \text{ и (или) } f \neq b, \\ (d, f) \in \Theta}} n_{Cf}^{df}.$$

Информативность вербальной ассоциации между предложениями и текстами. Рассмотрим l -мерное евклидово пространство E . Для его построения лексикографически упорядочим все пары словоформ из полного корпуса текстов Cf , т. е. сформируем кортеж $\Theta = \langle (a_1, b_1), (a_2, b_2), \dots, (a_l, b_l) \rangle$.

Пусть π и ρ – два предложения (текста) из корпуса Cf , а W_π и W_ρ – соответственно множества словоформ в этих предложениях, дополненные всеми синонимами и словоизменениями из словарей Dic_{par} и Dic_{syn} .

Построим вектор в пространстве E :

$$\mathbf{I}_{Cf}^{\pi\rho} = (I_{Cf}^{a_1 b_1}, I_{Cf}^{a_2 b_2}, \dots, I_{Cf}^{a_l b_l}). \quad (3)$$

Если словоформы a_i и b_i содержатся соответственно в множествах W_π и W_ρ (или W_ρ и W_π), то значение информативности $I_{Cf}^{a_i b_i}$ в формуле (3) определяется из словаря вербально-ассоциативных пар слов Dic_{ab} . В противном случае $I_{Cf}^{a_i b_i} = 0$.

Нормализованную информативность $I_{Cf}^{\pi\rho}$ вербальной ассоциации между предложениями (текстами) π и ρ можно интерпретировать как проекцию вектора $\mathbf{e} = (1, 1, \dots, 1)$ размерности l на направление вектора $\mathbf{I}_{Cf}^{\pi\rho}$, т. е. отношение скалярного произведения векторов $\mathbf{I}_{Cf}^{\pi\rho}$ и \mathbf{e} к длине вектора $\mathbf{I}_{Cf}^{\pi\rho}$:

$$I_{Cf}^{\pi\rho} = \frac{\mathbf{I}_{Cf}^{\pi\rho} \cdot \mathbf{e}}{|\mathbf{I}_{Cf}^{\pi\rho}|} = \frac{\sum_{i=1}^l I_{Cf}^{a_i b_i}}{\sqrt{\sum_{i=1}^l (I_{Cf}^{a_i b_i})^2}}. \quad (4)$$

При реализации алгоритма вычисления информативности вербальной ассоциации между предложениями или текстами удобно пользоваться следующей формулой, полученной из выражения (4):

$$I_{Cf}^{пр} = \frac{I_1 + I_2 + \dots + I_r}{\sqrt{(I_1)^2 + (I_2)^2 + \dots + (I_r)^2}}, \quad (5)$$

где I_1, I_2, \dots, I_r – все отличные от нуля координаты вектора $\mathbf{I}_{Cf}^{пр}$.

Описание алгоритма классификации предложений из текста описания проблемной ситуации. Алгоритм классификации предложений текста T работает следующим образом.

Обозначим через S_1 первый класс предложений. В качестве единственного элемента первого класса S_1 будем рассматривать предложение ρ_1 . Затем вычислим информативность вербальной ассоциации между предложениями ρ_1 и ρ_2 по формуле (5). Если вычисленное значение не меньше некоторой пороговой величины ρ_0 , то предложение ρ_2 поместим в класс S_1 . Далее аналогичным образом вычислим информативность вербальной ассоциации между предложениями из пар $(\rho_1, \rho_3), \dots, (\rho_1, \rho_l)$. После завершения процесса формирования класса S_1 точно так же формируются и другие классы. В итоге будем иметь совокупность классов $\{S_1, S_2, \dots, S_m\}$ ($m \leq l$). Информативные классы будем использовать в качестве запросов на поиск аналогов принятых решений.

Индексирование запросов. Запросы являются, как правило, краткими сообщениями. Их объем не позволяет выявить статистические характеристики словоформ. Поэтому индексированию запроса предшествует процесс его расширения за счет включения релевантных предложений из полного корпуса текстов.

Информативность слов. Пусть T – текстовый документ, объем которого обеспечивает вычисление статистических характеристик его словоформ и предложений. Информативность I_T^a слова a из текста T вычислим по формуле

$$I_T^a = n_T^a / n_{Cf}^a, \quad (6)$$

где n_T^a и n_{Cf}^a – частоты встречаемости (с учетом словоизменения и синонимии) словоформы a в тексте T и полном корпусе текстов Cf соответственно [7]. При вычислении будем использовать частотный словарь словоформ $Dic_a = \{\langle a, n_{Cf}^a, n_{Ct_1}^a, n_{Ct_2}^a, \dots, n_{Ct_n}^a \rangle \mid a \in W_{Cf}\}$, в котором каждой словоформе из множества W_{Cf} всех словоформ корпуса Cf приписаны частоты ее встречаемости $n_{Cf}^a, n_{Ct_1}^a, n_{Ct_2}^a, \dots, n_{Ct_n}^a$ во всех тематических корпусах текстов Ct_i ($i = \overline{1, n}; N \geq 1$).

Используя лингвистические словари Dic_{par} и Dic_{syn} , формулу (6) перепишем в виде

$$I_T^a = \frac{n_T^a + n_T^{Par_a} + n_T^{Syn_a}}{n_{Cf}^a + N_{Cf}^{Par_a} + N_{Cf}^{Syn_a}}. \quad (7)$$

В выражении (7) $n_T^{Par_a}$ – это число вхождений всех словоформ текста T , являющихся словоизменениями словоформы a , т. е. верно равенство

$$n_T^{Par_a} = \sum_{b \in Par_a, b \neq a} n_T^b.$$

Параметр $n_T^{Syn_a}$ означает количество синонимов словоформы a в тексте T :

$$n_T^{Syn_a} = \sum_{c \in Syn_a, c \neq a} n_T^c.$$

Аналогичный смысл имеют параметры $N_{Cf}^{Par_a}$ и $N_{Cf}^{Syn_a}$:

$$N_{Cf}^{Par_a} = \sum_{b \in Par_a, b \neq a} n_{Cf}^b, \quad N_{Cf}^{Syn_a} = \sum_{c \in Syn_a, c \neq a} n_{Cf}^c.$$

Пусть теперь S – краткое текстовое сообщение. Обозначим через W_S множество всех его словоформ. Вычислим информативность $J_{Cf}^{S\pi}$ вербальной ассоциации между текстом S и некоторым предложением π из полного корпуса текстов Cf . По аналогии с выражением (3) построим вектор $\mathbf{J}_{Cf}^{S\pi} = (J_{Cf}^{c_1d_1}, J_{Cf}^{c_2d_2}, \dots, J_{Cf}^{a_k b_k})$ в евклидовом пространстве. Для вычисления информативности $J_{Cf}^{S\pi}$ воспользуемся аналогом формулы (5):

$$J_{Cf}^{S\pi} = \frac{J_1 + J_2 + \dots}{\sqrt{(J_1)^2 + (J_2)^2 + \dots}}, \quad (8)$$

где J_1, J_2, \dots – все отличные от нуля координаты вектора $\mathbf{J}_{Cf}^{S\pi}$. Если информативность выражения (8) не меньше некоторого критического значения, то предложение π занесем в текст S . Аналогично поступим и с другими такими предложениями полного корпуса текстов. В результате получим расширенное множество предложений, которое снова будем считать текстом (запросом) S .

Информативность I_S^a любого слова $a \in W_S$ вычислим по формуле (7):

$$I_S^a = \frac{n_S^a + n_S^{Par_a} + n_S^{Syn_a}}{n_{Cf}^a + N_{Cf}^{Par_a} + N_{Cf}^{Syn_a}}. \quad (9)$$

Информативность предложений и текстов. При вычислении информативности предложений текста T будем также исходить из их векторного представления: $\mathbf{\Pi} = (I_\pi^{a_1}, I_\pi^{a_2}, \dots, I_\pi^{a_l})$, где $I_\pi^{a_1}, I_\pi^{a_2}, \dots, I_\pi^{a_l}$ – значения информативности слов произвольного предложения π . (Компонента вектора $\mathbf{\Pi}$ равна нулю, если соответствующего слова нет в предложении π .) Тогда аналогично формуле (8) нормализованную информативность I_T^π предложения π будем вычислять по формуле

$$I_T^\pi = \frac{I_1 + I_2 + \dots}{\sqrt{(I_1)^2 + (I_2)^2 + \dots}}, \quad (10)$$

где I_1, I_2, \dots – значения информативности всех слов предложения π .

Информативность произвольного текста T из полного корпуса текстов будем вычислять по формуле, аналогичной выражению (10):

$$I_{Cf}^T = \frac{I_T^\pi + I_T^p + \dots}{\sqrt{(I_T^\pi)^2 + (I_T^p)^2 + \dots}}, \quad (11)$$

где I_T^π, I_T^p, \dots – значения информативности всех предложений документа T .

Описание алгоритма индексирования запросов. Алгоритм индексирования запросов (классов предложений из описания проблемной ситуации) функционирует в три этапа.

На первом этапе вычисляется информативность каждого из множеств S_1, S_2, \dots, S_m по формуле (11). Класс предложений будем считать информативным, если значение информативности не меньше некоторой пороговой величины. В результате выполнения первого этапа получаем совокупность информативных классов предложений $\{U_1, U_2, \dots, U_s\}$ ($s \leq m$). Классы, имеющие недостаточный объем для вычисления статистических характеристик словоформ, т. е. являющиеся краткими сообщениями, дополняются релевантными предложениями из полного корпуса текстов Cf с использованием формулы (8). Полученные в результате такого расширения новые классы будем применять в качестве запросов на поиск аналогов принятых решений.

На втором этапе вычисляется информативность $I_{U_i}^a$ ($i = \overline{1, s}$) всех словоформ из предложений всех классов U_1, U_2, \dots, U_s по формулам (7) и (9).

На третьем этапе формируются поисковые образы

$$\text{ПП}_i = \{(a, I_{U_i}^a); (b, I_{U_i}^b); \dots | a, b, \dots \in U_i\}, \quad i = \overline{1, s}, \quad (12)$$

всех классов U_1, U_2, \dots, U_s предложений из описания проблемной ситуации. Эти поисковые образы будут использованы в качестве поисковых предписаний на поиск аналогов решений.

Интернет-поиск аналогов принятых решений. Выявление аналогов принятых решений связано с двумя видами информационного поиска: поиска веб-страниц, упорядоченных по убыванию их информативности, и фактографического поиска информативных фрагментов текстовых документов на этих страницах.

При поиске аналогов решения поставленной задачи нужно учитывать тот факт, что в Интернете индексируются не сами документы, а веб-страницы, на которых они расположены. Это обстоятельство существенным образом влияет на выбор критериев выдачи и построение алгоритмов поиска аналогов.

Критерии выдачи. Критерий выдачи – это правило, согласно которому вычисляется степень соответствия запросу веб-страниц или текстовых документов, найденных в процессе информационного поиска. В большинстве информационных систем критерии выдачи строятся на основе векторной модели описания данных в виде косинуса угла между векторами поискового предписания и поискового образа документа [8].

Пусть по-прежнему W_{Cf} – множество всех словоформ полного корпуса текстов Cf , а R – m -мерное евклидово пространство ($m = |W_{Cf}|$). Для каждой веб-страницы S построим вектор ее поискового образа в пространстве R : $\mathbf{F}_S = (I_S^{a_1}, I_S^{a_2}, \dots, I_S^{a_m})$. Аналогично запишем вектор поискового предписания (12): $\mathbf{F}_{\text{ПП}_i} = (I_{\text{ПП}_i}^{b_1}, I_{\text{ПП}_i}^{b_2}, \dots, I_{\text{ПП}_i}^{b_m})$. Тогда для поиска веб-страниц по поисковому предписанию $\mathbf{F}_{\text{ПП}_i}$ в качестве критерия выдачи используем косинус угла φ между векторами \mathbf{F}_S и $\mathbf{F}_{\text{ПП}_i}$:

$$\cos \varphi = \frac{\mathbf{F}_S \cdot \mathbf{F}_{\text{ПП}_i}}{|\mathbf{F}_S| \cdot |\mathbf{F}_{\text{ПП}_i}|} = \frac{\sum_{j=1}^m I_S^{a_j} I_{\text{ПП}_i}^{b_j}}{\sqrt{\sum_{j=1}^m (I_S^{a_j})^2} \cdot \sqrt{\sum_{i=1}^n (I_{\text{ПП}_i}^{b_j})^2}}. \quad (13)$$

Описание алгоритма поиска веб-страниц. Поиск аналогов принятых решений осуществляется в три этапа.

На первом этапе предложения текста описания проблемной ситуации разбиваются на классы.

На втором этапе реализуется индексирование информативных классов предложений. В результате индексирования формируется совокупность поисковых предписаний (12) для интернет-поиска аналогов принятых решений.

На третьем этапе по каждому поисковому предписанию ПП_i ($i = \overline{1, s}$) проводится поиск веб-страниц, содержащих аналоги принятых решений. При поиске используется критерий выдачи (13), все найденные страницы упорядочиваются по убыванию его значений.

Описание алгоритма фактографического поиска. Поиск сводится к выделению в найденных текстах информативных фрагментов, релевантных каждому классу предложений из множества классов $\{U_1, U_2, \dots, U_s\}$ ($s \leq m$). Процедура включает два этапа.

На первом этапе вычисляется информативность I_Q^a каждой словоформы a из найденного текста Q по формуле, аналогичной выражению (7):

$$I_Q^a = \frac{n_Q^a + n_Q^{Par_a} + n_Q^{Syn_a}}{n_{Cf}^a + N_{Cf}^{Par_a} + N_{Cf}^{Syn_a}}.$$

Затем определяется информативность $I_Q^{\pi}, I_Q^{\rho}, \dots$ каждого предложения текста Q по формуле (10).

На втором этапе фактографического поиска выявляется контекстное окружение всех информативных предложений текста Q путем вычисления информативности вербальной ассоциации каждого из них с другими предложениями этого текста по формуле (5).

Сформированные таким образом фрагменты текстовых документов на найденных веб-страницах могут быть использованы как аналоги принятых решений.

Лексико-семантическая обработка аналогов принятых решений. Данная обработка заключается в выявлении тонально окрашенной информации в описаниях принятых решений, найденных в Интернете. Для оценки тональности информации будем использовать совокупность специальных тематических корпусов текстов.

Тонально окрашенные тематические корпуса текстов. Для оценки тональности сообщений в Интернете будем использовать совокупность тонально окрашенных тематических корпусов текстов. Каждому корпусу соответствует некоторая оценка тональности. При n -балльной шкале оценок количество таких корпусов должно быть равно n . Всякий i -й корпус Ct_i включает текстовые документы одинаковой тональности, т. е. корпус Ct_i – эта пара $\langle Ct_i, Et_i \rangle$, где Et_i – оценка тональности каждого документа из множества Ct_i . В простейшем случае формируются два корпуса текстов с тонально окрашенной лексикой. Первый корпус создается для анализа положительной тональности в описаниях принятых решений, а второй – для анализа отрицательной тональности. На основе тонально окрашенных тематических корпусов текстов формируется частотный словарь тонально окрашенной лексики $Lex_a = \{ \langle a, n_{Cf}^a, n_{\langle Ct_1, Et_1 \rangle}^a, n_{\langle Ct_2, Et_2 \rangle}^a, \dots, n_{\langle Ct_n, Et_n \rangle}^a \rangle \mid a \in W_{Cf} \}$. Здесь a – словоформа, n_{Cf}^a и $n_{\langle Ct_i, Et_i \rangle}^a$ ($i = \overline{1, n}$) – абсолютные частоты ее появления соответственно в полном корпусе текстов Cf и в i -м тонально окрашенном корпусе, W_{Cf} – множество всех словоформ полного корпуса.

Оценка тональности аналогов принятых решений. Пусть по-прежнему Ct_i ($i = \overline{1, n}$) – тонально окрашенные тематические корпуса текстов. Каждый корпус Ct_i состоит из текстов одинаковой тональности и представляет собой пару $\langle Ct_i, Et_i \rangle$ (Et_i – оценка тональности для всех текстов из множества Ct_i). Пусть также Q – текстовое сообщение, полученное в результате формирования контекстного окружения к некоторому найденному на веб-странице информативному предложению. Построим вектор F_Q сообщения Q и векторы F_{Ct_i} ($i = \overline{1, n}$) по аналогии с построением векторов F_s и F_{III_2} . Для каждой пары (F_Q, F_{Ct_i}) вычислим косинус угла между этими векторами по формуле

$$\cos(F_Q, F_{Ct_i}) = \frac{F_Q F_{Ct_i}}{|F_Q| |F_{Ct_i}|}, i = \overline{1, n}.$$

Тогда сообщению Q будет соответствовать оценка тональности Et_i при таком значении i , при котором $\cos(F_Q, F_{Ct_i})$ принимает наибольшее значение.

По результатам интернет-поиска аналогов принятых решений и анализа тонально окрашенной информации формируется отчет (таблица).

Отчет по результатам поиска аналогов принятых решений и анализа тонально окрашенной информации на примере приобретения участка или дачи

Тонально окрашенное сообщение	Адрес веб-страницы	Информативность, %	Оценка тональности
<i>Купить участок.</i> Вариант для терпеливых, трудолюбивых и богатых. Дачу нужно будет построить. Участок – облагородить. Сад – посадить. При больших вложениях, с наймом специалистов можно построиться примерно за два года. Но чаще всего нужно лет пять. И еще лет 15–20 на сад	https://myfin.by/stati/view/13498-za-i-protiv-stoitli-pokupat-dachu	43	3 (из 10)
<i>Купить готовую дачу.</i> Выйдет где-то на 1/3 дешевле, чем было потрачено на ее строительство. И это только денежные расходы, компенсировать прошлому хозяину его труды и время не нужно		61	7 (из 10)

В таблице каждому сообщению соответствует адрес веб-страницы, на которой оно расположено, а также приведены информативность и числовое значение оценки тональности в принятой шкале.

Заключение. Разработана математическая модель интернет-поиска и лексико-семантической обработки аналогов принятых решений, найденных в Интернете по запросам, которые были синтезированы из описаний проблемных ситуаций в соответствии с алгоритмами из работы автора [2]. Промоделированы четыре этапа данного процесса. На первом этапе синтезируются запросы на основе исследования вербальных ассоциаций между предложениями в описании проблемной ситуации. На втором этапе полученные запросы индексируются. Поисковые предписания представляются в виде множеств слов с соответствующими значениями информативности. На третьем этапе реализуется поиск аналогов принятых решений в порядке, определяемом специальным упорядочивающим отношением, которое задается на множестве веб-страниц каждого сканируемого веб-сайта. На четвертом этапе проводится лексико-семантическая обработка информационных сообщений, в процессе которой найденные аналоги решений исследуются на тональность. При оценке тональности используются лингвистические словари тонально окрашенной лексики, которые формируются на основе специальных тонально окрашенных тематических корпусов текстов. В предельном случае создаются два типа словарей. Первый тип предназначен для анализа положительной тональности в описаниях принятых решений, а второй – для анализа отрицательной тональности.

Предложенная модель интернет-поиска и лексико-семантической обработки аналогов принятых решений может быть использована в тех предметных областях, где необходимо работать с крупными объемами текстовой информации. После получения результатов анализа тональности сообщений пользователю информационной системы будет проще принять окончательное решение.

Список использованных источников

1. Ларичев, О. И. Вербальный анализ решений / О. И. Ларичев. – М. : Наука, 2006. – 181 с.
2. Липницкий, С. Ф. Синтез запросов и поиск альтернатив в системе информационной поддержки принятия решений / С. Ф. Липницкий // Проблемы физики, математики и техники. – 2020. – № 2. – С. 91–95.
3. Мартинович, Г. А. Вербальные ассоциации и организация лексикона человека / Г. А. Мартинович // Филологические науки. – 1989. – № 3. – С. 39–45.
4. Еленевская, М. Н. Хранение и описание вербальных ассоциаций: словари и тезаурусы [Электронный ресурс] / М. Н. Еленевская, И. Г. Овчинникова. – Режим доступа: <https://cyberleninka.ru/article/n/hranenie-i-opisanie-verbalnyh-assotsiatsiy-slovary-i-tezaurusy/viewer>. – Дата доступа: 05.11.2020.
5. Морковкин, В. В. Идеографические словари [Электронный ресурс] / В. В. Морковкин. – Режим доступа: http://rifmovnik.ru/ideoog_book.htm. – Дата доступа: 05.11.2020.
6. Мартинович, Г. А. Вербальные ассоциации в ассоциативном эксперименте / Г. А. Мартинович. – СПб. : Изд-во СПбГУ, 1997. – 72 с.
7. Липницкий, С. Ф. Модель представления знаний в информационных системах на основе вербальных ассоциаций / С. Ф. Липницкий // Информатика. – 2011. – № 4(32). – С. 21–28.
8. Ландэ, Д. В. Поиск знаний в Internet. Профессиональная работа / Д. В. Ландэ. – М. : Диалектика-Вильямс, 2005. – 272 с.

References

1. Larichev O. I. Verbal'nyj analiz reshenij. *Verbal Analysis of Decisions*. Moscow, Nauka, 2006, 181 p. (in Russian).
2. Lipnitsky S. F. Sintez zaprosov i poisk al'ternativ v sisteme informacionnoj podderzhki prinjatija reshenij [Synthesis of queries and search for alternatives in the decision support information system]. *Problemy fiziki, matematiki i tehniki* [Problems of Physics, Mathematics and Technology], 2020, no. 2, p. 91–95 (in Russian).
3. Martinovich G. A. Verbal'nye associacii i organizacija leksikona cheloveka [Verbal associations and organization of the human lexicon]. *Filologicheskie nauki* [Philological Sciences], 1989, no. 3, p. 39–45 (in Russian).
4. Yelenevskaya M. N., Ovchinnikova I. G. Hranenie i opisanie verbal'nyh associacij: slovary i tezaurusy. *Storage and Description of Verbal Associations: Dictionaries and Thesauri* (in Russian). Available at: <https://cyberleninka.ru/article/n/hranenie-i-opisanie-verbalnyh-assotsiatsiy-slovary-i-tezaurusy/viewer> (accessed 05.11.2020).

5. Morkovkin V. V. Ideograficheskie slovari. *Ideographic Dictionaries* (in Russian). Available at: http://rifmovnik.ru/ideog_book.htm (accessed 05.11.2020).

6. Martinovich G. A. Verbal'nye associacii v associativnom jeksperimente. *Verbal Associations in an Associative Experiment*. Saint Petersburg, Izdatel'stvo Sankt-Peterburgskogo gosudarstvennogo universiteta, 1997, 72 p. (in Russian).

7. Lipnitsky S. F. Model' predstavlenija znaniy v informacionnyh sistemah na osnove verbal'nyh associacij [Model of knowledge representation in information systems based on verbal associations]. *Informatika [Informatics]*, 2011, no. 4(32), p. 21–28 (in Russian).

8. Lande D. V. Poisk znaniy v Internet. Professional'naja rabota. *Knowledge Search in Internet. Professional Work*. Moscow, Dialektika-Viliams, 2005, 272 p. (in Russian).

Информация об авторе

Липницкий Станислав Феликсович, доктор технических наук, главный научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.
E-mail: lipn@newman.bas-net.by

Information about the author

Stanislav F. Lipnitsky, Dr. Sci. (Eng.), Chief Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
E-mail: lipn@newman.bas-net.by

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

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
INFORMATION TECHNOLOGY

УДК 004.33.054
<https://doi.org/10.37661/1816-0301-2020-17-4-83-91>

Поступила в редакцию 24.08.2020
Received 24.08.2020

Принята к публикации 22.09.2020
Accepted 22.09.2020

**Международный стандарт HL7 FHIR как основа
создания единого информационного пространства
здравоохранения Республики Беларусь**

К. И. Костюк^{1✉}, А. В. Браницкий¹, В. В. Рубо¹, И. М. Нестерович²

¹Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь
✉E-mail: kostyukkarolina@newman.bas-net.by

²Унитарное предприятие «Геоинформационные системы», Минск, Беларусь

Аннотация. Рассматриваются позиции Республики Беларусь в сфере информационно-коммуникационных технологий, уровни сбора медицинских данных, а также международный стандарт для обмена медицинской информацией FHIR (Fast Healthcare Interoperability Resources), представленный международной некоммерческой организацией HL7 (Health Level Seven International) в 2011 г. и предназначенный для улучшения взаимодействия независимых медицинских информационных систем и снижения стоимости разработки интерфейсов. В FHIR введен термин «ресурс» в качестве основной единицы значимой информации для передачи и хранения данных, описывающей набор информационных ресурсов и протокол взаимодействия с хранилищем ресурсов посредством REST-сервиса или обмена сообщениями. Фиксированный базовый набор информационных ресурсов призван охватить 80 % информатизируемых сценариев. Данный стандарт начал широко применяться в медицинских информационных системах Республики Беларусь, в том числе в республиканской системе обращения электронных рецептов.

Ключевые слова: национальная система электронной медицинской карты, индекс развития информационно-коммуникационных технологий, электронная медицинская карта, медицинская информационная система, стандарт для обмена медицинскими данными HL7 FHIR

Для цитирования. Международный стандарт HL7 FHIR как основа создания единого информационного пространства здравоохранения Республики Беларусь / К. И. Костюк [и др.] // Информатика. – 2020. – Т. 17, № 4. – С. 83–91. <https://doi.org/10.37661/1816-0301-2020-17-4-83-91>

**International standard HL7 FHIR as a working framework
for development of common healthcare information space
of the Republic of Belarus**

Karolina I. Kostyuk^{1✉}, Aliaksandr V. Branitski¹, Vitaly V. Roubo¹, Irina M. Nesterovich²

¹The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus
✉E-mail: kostyukkarolina@newman.bas-net.by

²Unitary Enterprise "Geoinformation Systems", Minsk, Belarus

Abstract. The positions of the Republic of Belarus in the field of ICT, the levels of medical data store and collection, HL7 FHIR standard are considered. International standard HL7 FHIR for the exchange of medical data, designed to improve the interaction of independent medical information systems and reduce the cost of

developing interfaces. FHIR introduces the term Resource as the main unit of meaningful information for transferring and storing data, describes a set of information resources and a protocol for interacting with the resource store through a REST service or messaging. A fixed basic set of information resources is designed to cover 80 % of information scenarios. This standard is used in the national system of electronic prescriptions in the Republic of Belarus.

Keywords: national electronic health record system, index of development of information and communication technologies, electronic medical record, medical information system, standard for the exchange of medical data HL7 FHIR

For citation. Kostyuk K. I., Branitski A. V., Roubo V. V., Nesterovich I. M. International standard HL7 FHIR as a working framework for development of common healthcare information space of the Republic of Belarus. *Informatics*, 2020, vol. 17, no. 4, pp. 83–91 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-4-83-91>

Введение. Современный период развития общества характеризуется возрастающим влиянием информационных технологий (ИТ), которые проникают во все сферы человеческой деятельности и обеспечивают распространение информационных потоков в обществе, образуя глобальное информационное пространство. Трудно найти сферу, в которой сейчас не используются ИТ. Лидирующие позиции по их внедрению занимают архитектура, машиностроение, образование, банковская структура и, конечно же, медицина. Во многих медицинских исследованиях просто невозможно обойтись без компьютера и специального ПО. Умение использовать ИТ становится одним из профессиональных навыков медицинского специалиста. Применение современных технологий в медицине улучшает качество оказания медицинской помощи, совершенствует деятельность медицинских и фармацевтических организаций.

Одним из основных показателей средней ожидаемой продолжительности жизни и социально-демографического развития в странах мира является индекс уровня продолжительности жизни (life expectancy index). Результаты расчетов численности населения Земли за 2019 г. подтверждают значительные успехи в снижении смертности и повышении продолжительности жизни (URL: <https://population.un.org/wpp/Methodogy/>). Показатель ожидаемой продолжительности жизни рассчитывается ежегодно по методике Программы развития ООН на основе статистических данных, получаемых от национальных институтов и международных организаций и аккумулируемых в отделе народонаселения Департамента ООН по экономическим и социальным вопросам. В целом принято считать, что рост продолжительности жизни есть следствие экономического развития, научного прогресса (прежде всего в области медицины) и других показателей [1].

Несмотря на возросшие возможности в лечебном деле, системы здравоохранения даже самых богатых стран сталкиваются с экономическими и производственными трудностями в поддержании качества медицинской помощи для стареющего населения. ИТ способны обеспечить управление ресурсами, дистанционное оказание медицинской помощи в рамках врачебных онлайн-консультаций и удаленное наблюдение за состоянием здоровья пациентов, снижение количества врачебных ошибок, доступность медицинской помощи и современный уровень лечения для населения. В настоящее время осуществляется переход к комплексной автоматизации этих направлений медицины, лечебно-профилактических учреждений и территориального здравоохранения.

Национальная система электронной медицинской карты. При проведении в 2015 г. третьего глобального исследования ВОЗ по электронному здравоохранению (URL: https://www.who.int/goe/publications/atlas_2015/en/) особое внимание уделялось использованию электронных технологий для поддержки всеобщего охвата услугами здравоохранения, в том числе предоставлению услуг удаленным группам населения и общинам с недостаточным уровнем обслуживания посредством телемедицины, подготовке кадров здравоохранения с помощью электронного обучения, улучшению диагностики и лечения, предоставлению точной и своевременной информации о пациенте с помощью электронных медицинских карт (ЭМК) – всему тому, что способствует улучшению работы и финансовой эффективности системы здравоохранения.

Согласно третьему глобальному исследованию ВОЗ в Республике Беларусь с 2005 г. действует национальная система ЭМК и каждый гражданин Беларуси получил национальный идентификационный номер. Таким образом осуществляется возможность унифицировать истории здоровья граждан страны. (Медицинские карты не связаны с национальным удостоверением личности.)

В Республике Беларусь финансирование сферы здравоохранения в значительной степени зависит от средств республиканского и (или) местных бюджетов. Расходы на систему здравоохранения с 2013 по 2018 г. составили 3,8–4,2 % ВВП, что следует из заключений об исполнении республиканского бюджета Комитета госконтроля за эти годы (URL: <http://www.kgk.gov.by/ru/control-deyatelnost-atchiv-ru>). Суммы расходов выросли с 2,5 млрд руб. в 2013 г. до 4,89 млрд руб. в 2018 г. В последние годы расходы на здравоохранение составляют около 15 % от общих расходов консолидированного бюджета. Основная часть средств на медицину выделяется из местных бюджетов. В 2019 г. всего на здравоохранение потратили 5,75 млрд руб. В бюджете на 2020 г. было заложено 6,26 млрд руб. (URL: <http://www.minfin.gov.by/upload/bp/budget/budjet2020.pdf>).

В 2016 г. Международным банком реконструкции и развития проведено исследование в отношении предоставления кредита в размере 125 млн долл. для проекта «Модернизация системы здравоохранения Республики Беларусь» (с 2017 по 2022 г.), в том числе создания системы электронного здравоохранения, приобретения необходимых технологий, повышения квалификации работающих кадров (URL: <http://documents.worldbank.org/curated/en>). Все это помогло частично решить проблему финансирования информатизации сферы здравоохранения и создать интегрированную ЭМК пациента.

Информатизация области здравоохранения проводится в соответствии со Стратегией развития информатизации в Республике Беларусь на 2016–2022 г., утвержденной на заседании Президиума Совета Министров Республики Беларусь от 03.11.2015 № 26 (URL: <http://e-gov.by/zakony-i-dokumenty/strategiya-razvitiya-informatizacii-v-respublike-belarus-na-2016-2022-gody>). Согласно этой Стратегии основными направлениями информатизации системы здравоохранения являются:

- комплексная автоматизация медицинских учреждений на основе порталных решений и веб-технологий;
- обеспечение взаимодействия медицинских учреждений в рамках единого информационного пространства организаций здравоохранения;
- внедрение ЭМК гражданина Республики Беларусь, включая разработку правового режима ее использования;
- развитие проекта по использованию электронных рецептов;
- развитие средств телемедицины, в том числе для обеспечения возможности консультирования пациентов в режиме реального времени и удаленного мониторинга состояния здоровья хронических больных.

Индекс развития информационно-коммуникационных технологий. Индекс развития информационно-коммуникационных технологий (ИКТ) в странах мира рассчитывается и публикуется Международным союзом электросвязи (МСЭ) – специализированным подразделением ООН, определяющим мировые стандарты в области ИКТ. Основная задача данного индекса состоит в возможности определения тех показателей, значения которых необходимо улучшить для сокращения «цифрового разрыва» с наиболее развитыми странами. Он был разработан в 2007 г. на основе 11 статистических показателей, которыми МСЭ оперирует в своих оценках развития ИКТ. Индекс сводит эти показатели в единый критерий, который призван сравнивать достижения стран мира в развитии ИКТ и может быть использован в качестве инструмента для проведения сравнительного анализа на глобальном, региональном и национальном уровнях.

Приказом Министерства связи и информатизации Республики Беларусь от 15 апреля 2011 г. № 105 принята и утверждена методика расчета индекса развития ИКТ. Согласно этой методике показатели сгруппированы в три субиндекса: доступа к ИКТ, использования ИКТ, практических навыков использования ИКТ.

Субиндекс доступа к ИКТ позволяет оценить уровень развития инфраструктуры электросвязи в соответствии с показателями:

- количества телефонных линий на 100 жителей;
- количества абонентов сотовой подвижной электросвязи на 100 жителей;
- пропускной способности внешнего шлюза на одного интернет-пользователя;
- доли домохозяйств, имеющих персональный компьютер;
- доли домохозяйств, имеющих доступ в сеть Интернет.

Субиндекс использования ИКТ включает в себя показатели количества интернет-пользователей на 100 жителей, абонентов и пользователей стационарного широкополосного доступа на 100 жителей, абонентов и пользователей мобильного широкополосного доступа на 100 жителей.

Субиндекс практических навыков использования ИКТ позволяет оценить уровень развития человеческого капитала по показателям уровня грамотности взрослого населения, охвата населения образованием второй и третьей ступеней (базовым и средним).

МСЭ на регулярной основе публикует в областях индексы развития ИКТ, что позволяет странам следить за изменениями во временной динамике. С 2018 г. выпуск индекса приостановлен из-за пересмотра показателей, включенных в исследование, а также методов их измерения. МСЭ планирует возобновить ежегодную публикацию результатов исследования в 2020 г.

В соответствии с отчетом МСЭ за 2017 год (URL: <http://www.itu.int/mis2017>, <http://www.itu.int/idi2017>) Беларусь занимает 32-ю позицию в общем списке стран мира и является лидером по развитию ИКТ в регионе СНГ (7,55 балла). По сравнению с 2016 г. удалось добиться прогресса по субиндексу доступа к ИКТ (плюс три позиции), субиндексу использования ИКТ (плюс две позиции), а также сохранить позиции в числе мировых лидеров по уровню развития человеческого потенциала (пятое место). Значения показателей Республики Беларусь в индексе развития ИКТ МСЭ в 2017 г. (URL: <http://handle.itu.int/11.1002/pub/80f52533-en>, <https://www.itu.int/net4/ITU-D/idi/2017/index.html#idi2017economycard-tab&BLR>) приведены в таблице.

Профиль Республики Беларусь в индексе развития ИКТ
Международного союза электросвязи, 2017 г.

<i>Субиндекс доступа к ИКТ</i>	7,87
Количество абонентов стационарных телефонных линий на 100 чел.	49,01
Количество абонентов мобильной связи на 100 чел.	124,17
Пропускная способность сети Интернет (бит/с на одного пользователя)	168 517,63
Процент домохозяйств, имеющих компьютер	66,96
Процент домохозяйств, имеющих доступ к сети Интернет	62,46
<i>Субиндекс использования ИКТ</i>	6,54
Процент пользователей сети Интернет	71,11
Абоненты широкополосного Интернета на 100 чел.	33,30
Абоненты беспроводных сетей на 100 чел.	69,49
<i>Субиндекс практических навыков использования ИКТ</i>	8,93
Средняя продолжительность обучения	12,00
Совокупная доля учащихся средней школы	107,12
Совокупная доля учащихся высшей школы	87,94

Уровни сбора медицинских данных. Достижения в области ИКТ в здравоохранении обусловили появление новых способов накопления, хранения, управления информацией о пациентах посредством оцифровки информации, связанной со здоровьем. Развитие ИКТ в здравоохранении также привело к созданию огромного количества информации, касающейся диагностики, тестирования, мониторинга, лечения, выставления счетов за медицинские услуги и управления ресурсами здравоохранения [2]. Можно выделить как минимум три вида цифровых медицинских карт, которые по наполнению очень похожи, но в их использовании есть некоторые различия:

EMR (electronic medical record) – цифровая версия карты пациента, созданной некоторыми поставщиками для отдельных медицинских учреждений. Данные, включенные в EMR, в основ-

ном используются для диагностики и лечения пациента. В отличие от бумажных карт EMR может помочь врачам отслеживать данные с течением времени и выявлять пациентов с особыми потребностями. Самый большой недостаток EMR заключается в том, что она ограничена одним поставщиком (в реалиях Республике Беларусь – локальными медицинскими информационными системами). EMR всего лишь на один шаг выше бумажной карты пациента.

EHR (electronic health record) – электронная медицинская карта, которая также является цифровой версией бумажной карты пациента. Она включает в себя данные всех врачей и учреждений, занимающихся лечением пациента. Авторизованные поставщики и персонал могут получить доступ к этой информации для лечения пациента. EHR следует за пациентом и не зависит от места его нахождения. Она по иерархии выше EMR и объединяет записи о наблюдении пациента во всех медицинских учреждениях.

PHR (personal health record) – личная медицинская карта. Первое упоминание о PHR встречается в раннем докладе Института медицины США «Компьютерная карта пациента: важная технология для здравоохранения» (URL: <https://population.un.org/wpp/Methodogy/>). Данные, хранящиеся в PHR, могут поступать из EHR, но здесь пациент отвечает за управление и доступ к PHR. Это помогает ему собирать и управлять собственной информацией о здоровье в частной и конфиденциальной среде. Пациент может вводить в свою PHR такие данные, как сведения об образе жизни или лекарствах, отпускаемых без рецепта, а также добавлять информацию из других источников, включая устройства для наблюдения за домом и дополнительных поставщиков услуг по уходу. PHR представляет собой наиболее полную медицинскую карту пациента и предоставляет ему все функции управления доступом к своим медицинским данным. Между тем доверять только PHR нельзя, так как пациент не является квалифицированным специалистом в области здравоохранения. Одни лишь записи PHR не могут служить достаточным основанием для постановки диагноза. Так же, как EMR и EHR, PHR может храниться в электронном виде и наполняться каким-либо ПО.

Таким образом, все три вида медицинских карт являются коллекциями данных, которые содержат информацию, связанную со здоровьем. Все они могут также включать информацию личного характера.

Применение мировых практик и стандартов в информатизации здравоохранения Республики Беларусь. Информатизация здравоохранения привела к созданию большого количества медицинских информационных систем (МИС). МИС – это совокупность информационных, организационных, программных и технических средств, предназначенных для автоматизации медицинских процессов и (или) организаций.

Как правило, обмен данными между различными МИС сильно затруднен вследствие несовместимости протоколов и семантических моделей данных. Проблема интеграции МИС остро ощущается как внутри медицинских учреждений, так и между ними. В такой ситуации целесообразно воспользоваться стандартами в сфере ИТ в здравоохранении для унификации обмена медицинскими данными между различными медицинскими системами (рис. 1).



Рис. 1. Комплексное решение для мониторинга здоровья

Ресурсы для обмена медицинской информацией HL7. В 1987 г. была создана некоммерческая организация Health Level Seven International, которая занималась разработкой, внедрением и накоплением медицинских знаний. Со временем HL7 стала лидером в определении стандартов для обмена данными и взаимодействия. С момента основания HL7 представлено несколько поколений стандартов обмена медицинскими данными (рис. 2). Значительная часть стандартов HL7 утверждена Американским национальным институтом стандартов и принята в качестве международных (ISO).

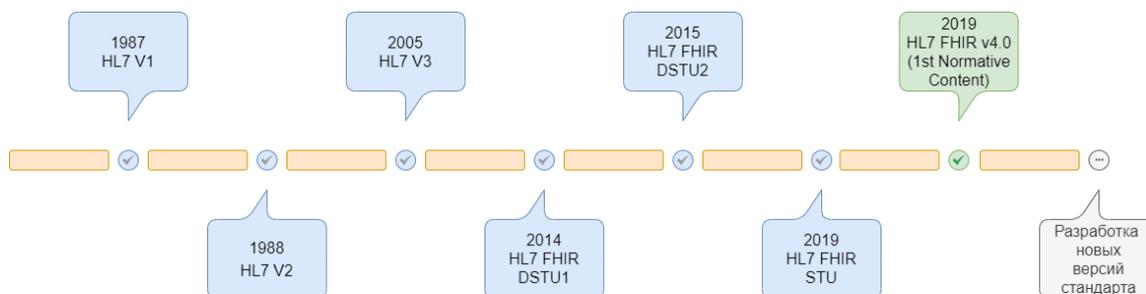


Рис. 2. Развитие стандартов HL7

Первоначальные форматы (например, HL7 v2 и v3), которые были отличным стартом для обмена данными, оказывались либо слишком неструктурированными, либо слишком ограниченными по объему, чтобы быть полезными для всех клинических приложений. Поэтому многие поставщики создавали свои собственные форматы данных и стандарты интерфейса прикладного программирования (Application Programming Interface, API) для обеспечения совместного использования данных. Каждая реализация требовала времени и денежных вложений. Это означало, что API-интерфейсы, предназначенные для обеспечения переносимости данных, становились препятствием для более крупных стратегий интеграции.

Новейшим стандартом HL7 является FHIR. Согласно HL7 цель модели данных FHIR состоит в том, чтобы помочь облегчить проблемы совместного использования данных и взаимодействия за счет упрощения реализации без ущерба для целостности информации, использования существующих логических и теоретических моделей для обеспечения последовательного, простого в реализации и строгого механизма обмена данными между приложениями здравоохранения (URL: <http://www.hl7.org>, https://www.researchgate.net/publication/323172606_Comparison_and_analysis_of_ISOIEEE_11073_IHE_PCD-01_and_HL7_FHIR_messages_for_personal_health_devices).

На данный момент стандарт FHIR имеет версию 4.0.1, выпущенную в релиз в октябре 2019 г., и эта версия начинает активно применяться в США, Европе и России.

Основные принципы FHIR. Модель данных FHIR – это стандарт для определения ресурсов и связанных метаданных, таких как клинический контент и другая соответствующая системная информация (например, возможности EHR) в согласованном, структурированном, но гибком модульном формате. Данные HL7 FHIR предназначены для взаимодействия между медицинскими платформами, однако структурированы таким образом, чтобы их можно было прочитать человеку.

Стандарт HL7 FHIR базируется на следующих принципах:

- следования за развивающимися веб-технологиями;
- практической реализуемости (удобен для разработчиков);
- включения в спецификацию только самых важных (общих) концептов, чтобы сохранить спецификацию небольшой и описать 20 % концептов, участвующих в 80 % информатизируемых сценариев;
- открытости стандарта и развития в ключе Open Source.

Основой FHIR является ресурс – независимая структурированная единица информации, используемая при обмене медицинскими данными. Большинство ресурсов представляют собой отображение реального мира в цифровых данных. Вот несколько примеров ресурсов: пациент (patient), визит (encounter), результат исследования (diagnosticReport). В рамках спецификации

описываются клинические, административные, финансовые и технические ресурсы. Каждый ресурс представлен набором стандартизованных атрибутов (элементов), в дополнение к этому в каждый ресурс встроены механизмы расширений. При разработке стандарта в рамках ресурса описываются только общие атрибуты, независимые от специализированного контекста (страна, специализация и т. д.), для всего остального встроены механизмы расширений (extensions).

Ресурсы могут быть связаны между собой и, соответственно, отражать взаимосвязи в реальном мире. Например, пациент (patient) пришел на обследование (encounter) в клинику (organization), ему поставили диагноз (condition) и назначили исследования (diagnosticOrder). При описании такого случая для обмена данными между системами потребуются взаимосвязанные ресурсы, в которых связь указывается через ссылку на другой ресурс в соответствующем атрибуте. Связи направлены от ресурса-источника к ресурсу-цели. Соответствующая обратная связь от цели к источнику существует в логическом смысле, но не представлена явно в целевом ресурсе.

Помимо модели данных и типов данных спецификация HL7 FHIR описывает несколько поддерживаемых подходов к организации обмена данными между системами. Системы могут как обмениваться ресурсами по отдельности, так и собирать связанные ресурсы в группы (bundles) и обмениваться этими группами ресурсов.

Варианты организации обмена данными:

RESTful API (HTTP) – взаимодействие между системами происходит посредством совершения операций над ресурсами с использованием REST-запросов (найти, получить, обновить ресурс и т. д.);

Messaging – взаимодействие между системами организовано в виде пересылки сообщений. Каждое сообщение содержит в себе информацию о произошедшем событии, которое отражено в информационной системе и о котором одна система хочет сообщить другой. Сообщение – это группа связанных между собой ресурсов определенного типа (message);

Documents – взаимодействие между системами происходит на уровне документов, т. е. одна система запрашивает документы у другой системы и получает их. Документ – это группа ресурсов, объединенных на момент формирования и подписания посредством специального ресурса Composition.

В зависимости от архитектуры системы и решаемых ею задач можно использовать подходящий вариант обмена.

Преимущества и недостатки стандарта HL7 FHIR. К преимуществам HL7 FHIR можно отнести следующее:

- это открытый стандарт с коммерческой экосистемой серверов, справочников и других инструментов для быстрой разработки;
- использование стандартных ИКТ XML, JSON, Web 2.0, HTTP и REST;
- расширяемость ресурсов (extensions);
- использование своих словарей (справочников) и единый подход к их ведению;
- профилирование ресурсов (один и тот же ресурс можно применять в одной системе с разным профилем);
- профилирование операций (расширение стандартного набора операций);
- хранение истории изменений для каждого ресурса, по которой можно выполнять поиск;
- применение параметров для поиска, фильтрации, сортировки результата запроса;
- использование только необходимых ресурсов и полей в рамках бизнес-логики ПО;
- связь между ресурсами посредством элементов (полей) ссылочного типа;
- возможность обмениваться одним ресурсом либо наборами ресурсов (bundles);
- поддержку различных подходов к организации обмена данными между системами: RESTful API (HTTP), Messaging, Documents, Services;
- низкий порог вхождения, хорошую документацию, наличие примеров;
- мощную общественную, правительственную (в США, Австралии) поддержку, а также поддержку лидеров ИТ-индустрии, таких как Apple, Google, Microsoft;
- наличие открытого сообщества экспертов по ИТ в медицине, регулярное проведение практических семинаров и конференций (онлайн и офлайн).

Стандарт FHIR ориентирован в первую очередь на простоту разработки и создание богатой экосистемы инструментов и библиотек. В отличие от предыдущих версий стандартов HL7, FHIR распространяется по открытой лицензии бесплатно и без ограничений. Разработка ведется публично, любой желающий и заинтересованный может внести свою лепту.

К недостаткам применения стандарта можно отнести следующее:

- минимальный обязательный набор элементов (полей) в рамках ресурса;
- обязательные к использованию словари (valueSets) в рамках ресурса;
- отсутствие обратной совместимости для версий младше FHIR 4.0.

Особенности применения стандарта HL7 FHIR в Республике Беларусь. В настоящее время в Беларуси наблюдается постепенный переход из EMR в EHR, т. е. локальные МИС начинают медленно делиться накапливаемыми данными с другими заинтересованными сторонами. Первым шагом к интеграции МИС стало создание полномасштабной системы обращения электронных рецептов в Республике Беларусь на базе стандарта HL7 FHIR.

Перечислим часть проблем, с которыми пришлось столкнуться на этапе разработки и внедрения первого ПО на основе стандарта FHIR в Республике Беларусь:

- отсутствие нормативно-правовых актов в сфере ИТ в здравоохранении, что существенно замедляет процесс внедрения ПО в эксплуатацию;
- более высокий приоритет бумажного носителя. Хранимые электронные документы и записи за редким исключением являются вторичными по отношению к документам на бумажном носителе и не имеют юридической значимости;
- отсутствие утвержденных справочников, единого хранилища справочной информации и корректного версионирования. Состав справочников не утвержден, не определен механизм их ведения, отсутствует своевременное уведомление об изменениях в составе справочников, что приводит к ошибкам при формировании ресурсов и обмену такими ресурсами;
- устаревшие технологии реализации МИС, которые не всегда имеют возможность беспрепятственной доработки для организации взаимодействия. Не закладывается бюджет на реализацию перехода конкретной МИС на новые технологии. Следовательно, значительно увеличиваются время и стоимость доработки с возможной потерей производительности и простоты использования;
- отсутствие глобального уникального идентификатора пациента. Сейчас роль такого идентификатора отведена карте медицинского обслуживания пациента, которая содержит уникальный номер, ассоциированный с пациентом (физическим лицом). Утерянная карта восстановлению не подлежит, номер карты пациента может быть изменен. Планируется, что карта медицинского обслуживания пациента будет заменена на ID-карту;
- слабая заинтересованность ключевых бенефициаров (медицинских работников, менеджеров здравоохранения всех уровней, пациентов) на этапах разработки, внедрения и развития ИС;
- низкий уровень доверия к ИТ на ранней стадии их внедрения, вследствие чего наблюдается высокая сопротивляемость новым технологиям;
- разный уровень информатизации регионов. Наиболее последовательно, комплексно и динамично информатизация здравоохранения осуществляется в Минске.

Некоторые из указанных проблем находятся на стадии решения. Скорость перехода напрямую связана с нормативно-правовой базой, а также с финансированием сферы здравоохранения.

Заключение. В целом в Республике Беларусь достигнуты неплохие успехи в развитии ИТ, создается национальная система ЭМК пациента и условия для постоянного развития электронного здравоохранения на основе структурированных технологических международных стандартов. Некоторые ИС, информационные ресурсы и ПО в сфере здравоохранения Беларуси соответствуют спецификации международного стандарта HL7 FHIR, что позволяет быстро организовать информационное взаимодействие медицинских систем между собой в рамках единого информационного пространства. Такое взаимодействие медицинских систем дает возможность повысить качество обслуживания населения и экономическую эффективность информатизации здравоохранения. Применение единого стандарта обмена медицинскими данными обеспечивает информационную совместимость и согласованность между различными МИС, а также позволяет повысить качество и ценность накапливаемых медицинских данных.

Список использованных источников

1. Щербакова, Е. М. Население мира по оценкам ООН пересмотра 2019 года [Электронный ресурс] // Демоскоп Weekly. – 2019. – № 821–822. – Режим доступа: <http://demoscope.ru/weekly/2019/0821/barom01.php>. – Дата доступа: 30.04.2020.
2. Healthcare Strategic Focus Area: Clinical Informatics / C. Bock [et al.] / National Institute of Standards and Technology, Technology Administration, Department of Commerce. – Gaithersburg, 2005. – P. 1–33.

References

1. Shcherbakova E. M. Naselenie mira po otsenkam OON peresmotra 2019 goda [World population according to UN estimates of the 2019 revision]. Demoskop Weekly [*Demoscope Weekly*], 2019, no. 821–822. Available at: <http://demoscope.ru/weekly/2019/0821/barom01.php> (accessed 30.04.2020) (in Russian).
2. Bock C. E., Carnahan L. J., Fenves S. J., Gruninger M., Kashyap V., ..., Sriram R. D. *Healthcare Strategic Focus Area: Clinical Informatics*. National Institute of Standards and Technology, Technology Administration, Department of Commerce, Gaithersburg, 2005, pp. 1–33.

Информация об авторах

Костюк Каролина Ивановна, инженер-программист, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.

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

Браницкий Александр Викентьевич, инженер-программист, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.

E-mail: sachabran@gmail.com

Роубо Виталий Вильгельмович, заведующий сектором разработки программных систем для здравоохранения, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.

E-mail: vitalikr@gmail.com

Нестерович Ирина Максимовна, инженер-программист, унитарное предприятие «Геоинформационные системы», Минск, Беларусь.

E-mail: nesterovichirina20@gmail.com

Information about the authors

Karolina I. Kostyuk, Software Engineer, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.

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

Aliaksandr V. Branitski, Software Engineer, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.

E-mail: sachabran@gmail.com

Vitaly V. Roubo, Head of the Sector for Developing Software Systems for Health Care, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.

E-mail: vitalikr@gmail.com

Irina M. Nesterovich, Software Engineer, Unitary Enterprise "Geoinformation Systems", Minsk, Belarus.

E-mail: nesterovichirina20@gmail.com

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

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

INTELLIGENT SYSTEMS

UDC 004.33.054
<https://doi.org/10.37661/1816-0301-2020-17-4-92-103>

Received 06.07.2020
Поступила в редакцию 06.07.2020

Accepted 22.10.2020
Принята к публикации 22.10.2020

Intelligent robot chair with communication aid using TEP responses and higher order spectra band features

Sathees Kumar Nataraj¹✉, Paulraj Murugesu Pandiyan², Sazali Bin Yaacob³,
Abdul Hamid Adom⁴

¹Department of Mechatronics Engineering,
AMA International University Bahrain, Salmabad 00973, Bahrain
✉E-mail: satheesjuly4@gmail.com

²Sri Ramakrishna Institute of Technology, Coimbatore, India

³Electrical, Electronics and Automation Section,
UniKL Malaysian Spanish Institute (UniKL MSI), Kulim, Malaysia

⁴School of Mechatronics Engineering, University of Malaysia Perlis,
Kangar, Perlis, Malaysia 02600

Abstract. In recent years, electroencephalography-based navigation and communication systems for differentially enabled communities have been progressively receiving more attention. To provide a navigation system with a communication aid, a customized protocol using thought evoked potentials has been proposed in this research work to aid the differentially enabled communities. This study presents the higher order spectra based features to categorize seven basic tasks that include Forward, Left, Right, Yes, NO, Help and Relax; that can be used for navigating a robot chair and also for communications using an oddball paradigm. The proposed system records the eight-channel wireless electroencephalography signal from ten subjects while the subject was perceiving seven different tasks. The recorded brain wave signals are pre-processed to remove the interference waveforms and segmented into six frequency band signals, i. e. Delta, Theta, Alpha, Beta, Gamma 1-1 and Gamma 2. The frequency band signals are segmented into frame samples of equal length and are used to extract the features using bispectrum estimation. Further, statistical features such as the average value of bispectral magnitude and entropy using the bispectrum field are extracted and formed as a feature set. The extracted feature sets are tenfold cross validated using multilayer neural network classifier. From the results, it is observed that the entropy of bispectral magnitude feature based classifier model has the maximum classification accuracy of 84.71 % and the value of the bispectral magnitude feature based classifier model has the minimum classification accuracy of 68.52 %.

Keywords: intelligent robot chair with communication aid, thought evoked potentials, bispectrum estimation ($B(f_1, f_2)$), multilayer neural network

For citation. Nataraj S. K., Paulraj M. P., Yaacob S. Bin, Adom A. H. Intelligent robot chair with communication aid using TEP responses and higher order spectra band features. *Informatics*, 2020, vol. 17, no. 4, pp. 92–103. <https://doi.org/10.37661/1816-0301-2020-17-4-92-103>

Интеллектуальное кресло-робот со вспомогательными средствами связи с использованием откликов ТЕР и характеристик диапазона спектра более высокого порядка

Сатис Кумар Натарадж^{1✉}, Паулраджд Муругеса Пандиян², Сазали бин Яacob³,
Абдул Хамид Адам⁴

¹Кафедра мехатроники, Международный университет АМА,
Бахрейн, Салмабад 00973, Бахрейн

✉E-mail: satheesjuly4@gmail.com

²Технологический институт Шри Рамакришны, Коимбатур, Индия

³Секция электротехники, электроники и автоматизации,
Малазийский испанский институт UniKL (UniKL MSI), Кулим, Малайзия

⁴Школа мехатроники, Университет Малайзии Перлис, Кангар, Перлис, Малайзия, 02600

Аннотация. В последние годы все больше внимания уделяется навигационным и коммуникационным системам на основе электроэнцефалограммы головного мозга для сообществ с разными возможностями. Для предоставления навигационной системе вспомогательных средств связи в работе предложен настраиваемый протокол, использующий вызванные мыслительные потенциалы, чтобы помочь сообществам с разными возможностями. Представлены функции, основанные на спектрах более высокого порядка, для классификации семи основных задач, таких как Вперед, Влево, Вправо, Да, НЕТ, Помощь и Расслабление, которые можно использовать для управления креслом-роботом, а также для связи с использованием необычной парадигмы. Предлагаемая система записывает восьмиканальный беспроводной сигнал электроэнцефалографии от десяти субъектов, в то время как субъект воспринимал семь различных задач. Записанные сигналы мозговых волн предварительно обрабатываются для удаления интерференционных волн и сегментируются на сигналы шести частотных диапазонов: дельта, тета, альфа, бета, гамма 1-1 и гамма 2. Сигналы полосы частот сегментируются на выборки кадров равной длины и используются для извлечения признаков с использованием оценки биспектра. Кроме того, статистические характеристики, такие как среднее значение биспектральной величины и энтропия с использованием области биспектра, извлекаются и формируются как набор характеристик. Извлеченные наборы функций проходят десятикратную перекрестную проверку с использованием классификатора многослойной нейронной сети. Результаты показали, что энтропия модели классификатора на основе характеристик биспектральной величины имеет максимальную точность классификации 84,71 %, а среднее значение модели классификатора на основе характеристик биспектральной величины – минимальную точность классификации 68,52 %.

Ключевые слова: интеллектуальное кресло-робот с коммуникационными средствами, вызванные мыслительные потенциалы, оценка биспектра ($B(f_1, f_2)$), многослойная нейронная сеть

Для цитирования. Интеллектуальное кресло-робот со вспомогательными средствами связи с использованием откликов ТЕР и характеристик диапазона спектра более высокого порядка / С. К. Натарадж [и др.] // Информатика. – 2020. – Т. 17, № 4. – С. 92–103. <https://doi.org/10.37661/1816-0301-2020-17-4-92-103>

Introduction. Movement and communication are the basic needs of human beings in their daily life and to live a meaningful life with interpersonal interactions [1]. Neuromuscular Disorder patients, such as Amyotrophic Lateral Sclerosis, neurodegenerative disease, muscular dystrophy, high cervical injuries or loss of the ability to speak (due to an accident) and Brain Stem Stroke have their walking abnormalities due to postural Instability and difficulty in communication with others due to loss of muscle control and speech [2–4]. Over the last decade, there has been an increasing attention on these patients to provide a navigation system and communication aid to enable them to lead a normal life [5–9]. In recent years, variety of BMI applications have arisen, Encouraged by new understanding of the non-invasive acquisition of human perception using powerful EEG amplifiers [10, 11], e. g. for cursor movement, acupuncture in pain relief [12], Neuro-prosthetic arm [13] and whole body movement [14], driver sleepiness detection [15], smart-living environmental control [16]. Currently, this research has been directed towards wheelchair navigation control and recognition of unspoken speech utterances without voluntary muscle activity [17–20].

Recently, several studies have examined thought evoked potential (TEP) based design of robotic wheelchair control using human thoughts [21], and communication systems using P300 speller and oddball paradigms [22]. Yet, the data acquisition protocols have shown a vital role in redefining the claimed action to command a navigation system or a communication system. The research work proposed by Kaufmann et al [17], involves positioning of four tactile stimulators and delivers navigation by concentrating their considerations on the desired tactile stimulus in an oddball paradigm to control the wheelchair. The results were validated through the participants navigating a virtual wheelchair. Theresa M. Vaughan [23], Frank H. Guenther [19] and Anne Porbadnigk [20] have developed several alternative communication systems using the recent developments in personal computers and new prosthetic methods to provide communication and control channels to individuals with difficulties in communication. Despite, none of the systems have produced an expanded utilization of the BMI technology to facilitate both navigation and communication through a customized brain activity recording protocol. Thus, in this study it is proposed to develop a customized thought controlled intelligent robot chair with communication aid (IRCC), as an initial step towards the possibility of navigation and speech production using a simple thought response based protocol (fig. 1). Depicts the block diagram of the proposed customized classification system for robot chair control along with a communication aid.

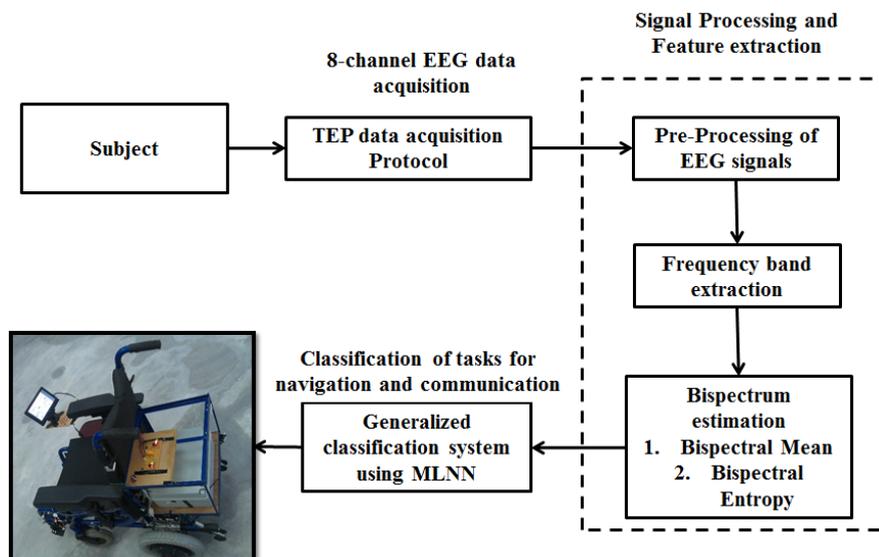


Fig. 1. Block diagram of the proposed thought controlled IRCC

The motivation towards this research is to establish a simple robot chair along with a communication aid, that can be used by an differentially enabled person, to control a wheel chair and to communicate their needs with others using TEP's. A simple data acquisition protocol has been proposed to develop the thought controlled IRCC; the tasks (Forward, Left, Right, Yes, No and Relax) were initially simulated and the subjects were requested to imagine during the data acquisition process. Further, the subjects were taught to pronounce the word loudly for the 'Help' task. The EEG signals are recorded for 12 sec. for each trial per task and are segmented into 10 sec. during the pre-processing stage for uniformity. In the pre-processing, the recorded brain wave signals were band-passed filtered in the frequency range of 0.5 to 100 Hz and segmented into six frequency bands Delta (δ), theta (θ), alpha (α), beta (β), Gamma 1 (γ_1) and Gamma 2 (γ_2). Thus, frequency band signals are segmented into frame segments (512 samples) and used to extract the features using higher order spectra (HOS) technique. The general motivation behind the use of bispectrum estimation is to detect and characterize the nonlinear properties of the TEP tasks, and they are potentially better to estimate the deviations from Gaussianness (normality) [24–26]. Thus in this study, the third order statistics bispectrum based feature extraction method has been implemented to extract the features from each frame of frequency band signals over each electrode position and the features such as the Mean of

bispectral magnitude (M) and the bispectral entropy features (E) were extracted. The non-linear features were extracted and associated with the corresponding TEP tasks. Then, the extracted feature sets were modeled using a supervised learning-multilayer neural network (MLNN) classifier and the classification performance was validated. The research methodology and the developed model results are explained in the subsequent sections of this paper.

Intelligent robot chair with communication database. The experimental setup and data acquisition procedures were implemented in the research lab at the School of Mechatronic Engineering, University Malaysia Perlis. The proposed study has been registered and approved from National Medical Research Registration (NMRR ID: NMRR-13-51-14570) and obtained Ethical approval from The Medical Research & Ethics Committee (MREC), Ministry of Health Malaysia. (Ref:(7)dIm.KKM/NIHSEC/800-2/2/2Jld2P13-179). This section elucidates some fundamental methods on the experimental setup which includes the wireless bio-amplifier setup and the placement of electrode channels for brain wave recording. Further, Suitable task selection, the data collection procedure and the formation of IRCC database were also presented. These processes are essential for the classification of thought evoked potentials to command an intelligent robot chair with communication aid.

Experimental setup and data acquisition tasks. In the experimental setup, a standard bio-signal acquisition system developed by ‘g-mobilab+’ 8-channel EEG data acquisition system was used to record the brain wave signals [27, 28]. The system consists of an electrode cap with nine differential screwable electrodes, bio-signal amplifier and wireless data acquisition using MATLAB® interactive programming environment. In this study, it is proposed to develop a BMI system which can be used to navigate the wheel chair and communicate with others using an oddball paradigm; through brain wave EEG signals when functional communications are disabled [29]. Thus, in the data acquisition protocol, three primary tasks that address the navigation of the robot chair and to select the isolated words in an oddball paradigm, such as Left, Forward and Right hand movement control are included. Further, three additional tasks have been included to use in emergency circumstances and to address the basic needs of a human being, they are Help, Yes, No tasks. Relax (normal) has been used as the reference signal in this experiment [30]. A semi-sound controlled room was used for the acquisition where the subjects were remained in a pleasant circumstance. The subject carried out seven different tasks. The EEG signals are recorded while the subject was settled comfortably and remained in totally static posture. No overt actions were made during the 12.0 sec. of the data acquisition process (fig. 2) depicts the tasks that were implemented using the TEP responses to command an IRCC.

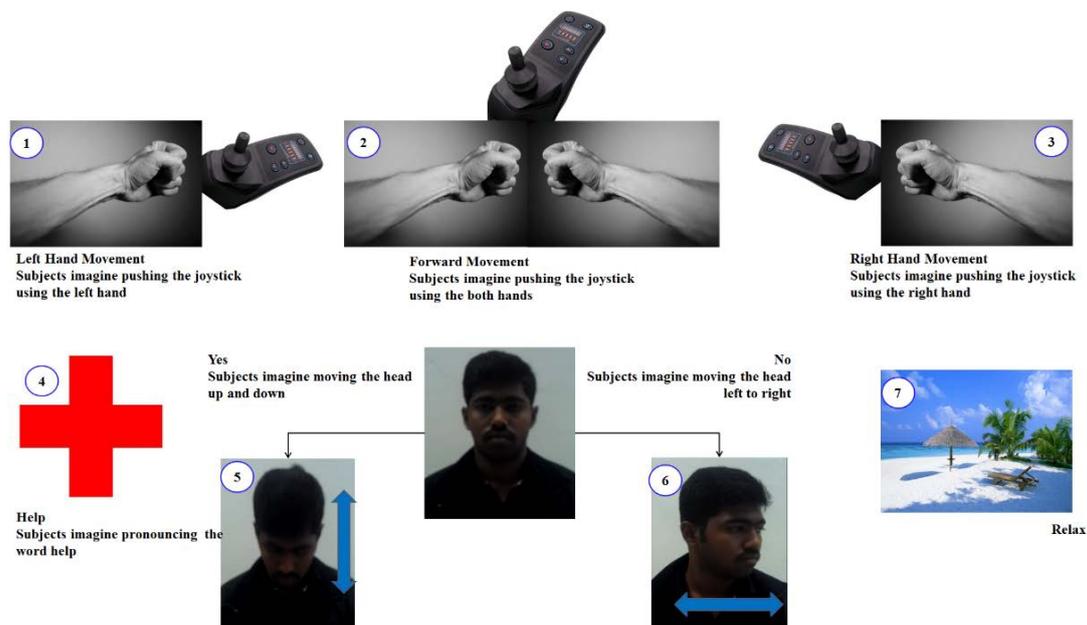


Fig. 2. Preliminary representation of the tasks (10.0 sec.) for subject to conduct the thought response data acquisition process

The system records the motor imaginary signal from the eight electrode positions such as Temporal (T3, T4), central (C3, C4), parietal (P3, P4), and occipital (O1, O2) while the subjects were performing the seven thought response tasks. In the electrode placement system, reference recording schemes were used. The electrodes are placed on T3, T4, C3, C4, P3, P4, O1 and O2 position with one common electrode on the left ear lobe of the body where potential remains fairly constant [31, 32]. The proposed IRCC system captures the brain wave patterns in order to identify the rhythmic activity for the seven different thoughts of an individual. Thus, during data collection, the EEG signals were recorded at a sampling rate of 256 Hz from a grid of 8 Ag/AgCl scalp electrodes which were placed on the scalp according to the international 10–20 lead system [31, 32]. The electrodes are placed on the scalp of the selected locations and tested for level of impedance using *g-tec* impedance checker. The impedance level was also tested after completing and maintained below 10 K Ω .

TEP data acquisition and IRCC database. In the data acquisition process, ten healthy BMI-naive volunteers (eight male, aged 21–30 years and two female, aged 24 years) were participated. During the data acquisition of each task, the subjects were requested to view the simulation of the specific task on the LCD monitor as depicted in fig. 2(1) to fig. 2(7) until recording all the trials. The simulation depicts the movement of a joystick moving left, forward and right movement for the left, forward and right directions respectively. For the additional tasks like ‘Yes’ and ‘No’, the simulation presents a volunteer performing head movements up-down and left-right movements and for ‘Help’ task, the subject was requested to pronounce imaginarily the word ‘help’ respectively. Then, the subject was requested to imagine the tasks as simulated on the monitor. When, the subject performs a specified task, the EEG signals emanated were recorded for 12.0 sec. from Parietal (P3 and P4), temporal (T3, T4), central (C3, C4), occipital (O1, O2), positions. Ground electrode and reference electrodes are placed in the Fpz position and left earlobe locations in order to make individual’s thought evoked tasks, recording comparable over time and to another individual’s record, International 10–20 system was used for the electrode placement [28, 31, 33].

The procedure of thought stimulus took the following format:

1. A simulation was presented on the LCD monitor (Left task) for 10 sec.
2. A 1 KHz tone (beep) sounded, the monitor displays the simulation of a moving joystick in left direction.
3. Then the monitor is turned off, the subject was given a time break of 5 sec. and requested to imagine the respective task.
4. The bio-signal recording was carried out for 12 sec. while the subjects performing the task.
5. The subject was given a time break of approximately two minutes after completing each trial.
6. The simulation continued until ten trials were performed.
7. Similarly, the next task simulation in the procedure was presented.

*Note that eyes remain open in all the mental tasks.

The recorded EEG signals are contaminated with unknown noise component lying within a 50–60 Hz frequency range, which are due to power-line noises. A simple first order IIR notch filter was designed for removing the Power line noise from the recorded EEG signals. The center frequency of the filter, F_0 was chosen to be at exactly 50 Hz and the bandwidth, $\Delta F = 4$ Hz. Then, the signals are converted into digital signals using a sampling frequency of 256 Hz simultaneously, the acquisition process was repeated 10 times for each task and the subject was requested to take a rest for ten minutes after each task. Similarly, this procedure was repeated for ten subjects and the recorded signals were combined and a database was formulated. The database was named as IRCC database. The IRCC database consists of data pertaining to 10 different subjects (for 7 tasks and each was performed for 10 trials). The collected database was validated using analysis of variance (ANNOVA) technique and the significance level of $4.29 \times 10^{-4} < p$ value was obtained when validated on a task basis.

Feature extraction using bispectrum estimation

Preprocessing. In the preprocessing stage, the 16-bit digitized signals with 256 Hz sampling frequency were trimmed to segregate the intermediate 10 sec. signals from 12 sec. signal. The trimmed raw signals are filtered to remove the artifacts and EMG’s below 0.5 Hz and above 100 Hz using 6th order band pass filters [34]. The segmented brain waves are categorized into six traditional bands: Delta (δ) 0.1–4 Hz, Theta (θ) 4–8 Hz, Alpha (α) 8–16 Hz, Beta (β) 16–32 Hz, Gamma 1, γ_1 (32–64 Hz) and Gamma 2, γ_2 (64–100 Hz).

Thus, each frequency band signals are segmented into frames such that a frame length of 2 sec. having 512 samples per frame along with an overlap of 1s ($m = 256$ samples). Thus, the first frame consists of $n = 512$ samples. The second frame was initiated after a lap of $m-1$ samples such that the second frame overlaps with the $n-m$ samples of the first frame. This procedure was repeated until all the frequency band signals were counted. Then, each frame is considered as an input to extract the higher order spectra also known as polyspectral representations of higher order statistics.

Bispectrum estimation. In various BMI applications, EEG signals have been analyzed using power spectra in several distinctive frequency bands. The power spectrum estimation provides the good statistical description of signals with an arbitrary distribution function. Moreover, power spectrum representation gives us the full canonical description in the case of stationary signals. In case of non-gaussianity or non-linear mechanisms, Higher order spectra can be used to determine the higher order moments or complaints which provide additional information on the phase characteristics and realistic information of the EEG signal [24]. In this paper, bispectrum $B(f_1, f_2)$ analysis has been employed to study the brain wave patterns of the visual stimuli. The bispectrum estimation is particularly the third-order statistics of a signal, which represents the Fourier transform of the third order correlation with highly interdependent frequency components [24, 35]. The mathematical representation of the bispectrum estimation is expressed in equation

$$B(f_1, f_2) = E[X(f_1)X(f_2)X^*(f_1 + f_2)], \quad (1)$$

where $X(f)$ is the DFT at frequency samples $x(nT)$, using the FFT algorithm. The frequency (f) may be normalized by the Nyquist frequency to be between 0 and 1. $X^*(f_1 + f_2)$ denotes complex conjugate and therefore the bispectrum obtained using equation (1) is a complex valued function which represents the product of three Fourier coefficients. In this feature extraction process, the non-redundant region or the positive bispectrum sequence (Ω) = $0 \leq f_2 \leq f_1 \leq (f_1 + f_2) \leq 1$ has been used to extract the Mean of bispectral magnitude features and the whole bispectrum region of computation has been used to extract the and grand mean of the bispectral magnitude features respectively.

To extract the $B(f_1, f_2)$ sequence in frequency domain, the EEG signal acquired from each channel were used to extract the six frequency band signals, namely Delta (δ), Theta (θ), Alpha (α), Beta (β), Gamma 1 (γ_1) and Gamma 2 (γ_2). Each frequency band signals were segmented into frames such that each frame has 512 samples. The positive fourier coefficients of $B(f_1, f_2)$ was estimated in the i^{th} frame of each channel. Thus the bispectrum sequence for $B(f_1, f_2)_{\delta i}^j$ was obtained from δ band, where i and j are the frame numbers and an electrode channel number respectively. Similarly, the bispectrum sequence for $B(f_1, f_2)_{\theta i}^j$, $B(f_1, f_2)_{\alpha i}^j$, $B(f_1, f_2)_{\beta i}^j$, $B(f_1, f_2)_{\gamma_1 i}^j$ and $B(f_1, f_2)_{\gamma_2 i}^j$, were also computed.

From the estimated sequences, two statistical features, namely, Mean of bispectral magnitude (M) and bispectral entropy (E) features are computed to characterize the distribution of bispectrum sequence as well as to minimize the dimension of the feature vector as shown in equation (2), (3). Therefore, for eight channels we have 48 (6×8) features per frame. The statistical features are extracted from all the trials and are used to form the feature set. Simultaneously, the features are derived from each task and the corresponding feature set consisting of 5600 samples are formulated and used to train and test the classifier models

$$M^{ij} = \left\{ M_{\delta}^{ij}(f_1, f_2), M_{\theta}^{ij}(f_1, f_2), M_{\alpha}^{ij}(f_1, f_2), M_{\beta}^{ij}(f_1, f_2), M_{\gamma_1}^{ij}(f_1, f_2) \text{ and } M_{\gamma_2}^{ij}(f_1, f_2) \right\}; \quad (2)$$

$$E^{ij} = \left\{ E_{\delta}^{ij}(f_1, f_2), E_{\theta}^{ij}(f_1, f_2), E_{\alpha}^{ij}(f_1, f_2), E_{\beta}^{ij}(f_1, f_2), E_{\gamma_1}^{ij}(f_1, f_2) \text{ and } E_{\gamma_2}^{ij}(f_1, f_2) \right\}, \quad (3)$$

where M and E represents the mean and entropy (bispectral magnitude of the Ω) in the i^{th} frame of the j^{th} electrode channel position.

$$E = -\sum_n pn \log(pn), \quad (4)$$

where

$$pn = \frac{|B(f_1, f_2)|}{\sum_{\Omega} |B(f_1, f_2)|}, \quad \Omega = \text{region of the bispectral magnitude}. \quad (5)$$

Classification of imaginary tasks using MLNN. MLNN are biologically inspired tools used for information processing and they are nonlinear in nature [36]. Classification of TEP responses to categorize the navigational tasks basically falls on pattern recognition problem. In this analysis, generalized IRCC system has been developed using MLNN for Multi-class pattern classification. The feature vectors derived from the mean of bispectral magnitude (5600×48 feature vectors) and bispectral entropy (5600×48 feature vectors) are processed subsequently and then associated with the seven different visual response tasks. Also, the feature vectors are normalized using binary normalization methods, where the dataset is recycled between 0.1 to 0.9 and partitioned into training and testing sets [37]. The training set has 4480×48 (80 % of master data set) and the testing set has the remaining 1120×48 (20 % of master data set) for the classification of the TEP tasks.

In this work, the MLNN models were organized with 48 input neurons, 25 hidden neurons and three neurons in the output layer. As the logistic sigmoid function scales any range of values between 0.1 and 0.9, in the designed MLNN models, logistic sigmoidal function was used to activate the neurons in the hidden and output layer. The Mean Squared Error (MSE) tolerance of 0.1 was used for training the neural network. In order to improve the performance rate, the learning rate, momentum factor and number of iterations were chosen based on the experimental observations in different trials. The learning rate and momentum factor for the models were chosen as 0.1 and 0.8 respectively. The generalization capability of the model was validated by performing 10 trials for training and testing method. The network models were trained using Levenburg Marquath Model. The MLNN model for spectral features were trained with ten trial weights for each subset. On the first subset, the network model was trained using 9/10 of the feature set and the classification rate was estimated using 1/10 subset of the remaining feature set. This process was repeated until all the 2/10 subset are used for the validation set [36, 37]. Further, the network training parameters, mean classification rate are shown in table 1, 2.

Table 1

The mean classification performance of the IRCC system using MLNN classifier and the mean of bispectral magnitude features

MLNN Classification results using entropy of bispectral magnitude features							
No. of training samples	4480	No. of hidden neurons	25	Output neurons	3	Training tolerance	0.03
No. of testing samples	1120	Input neurons	48			Testing tolerance	0.1
Trial	Training time (sec.)		Number of epochs		Classification accuracy (%)		
1	1272		159		79.52		
2	1096		137		68.52		
3	1150		144		71.85		
4	1289		161		80.56		
5	1270		159		79.40		
6	1221		153		76.34		
7	1366		161		80.37		
8	1250		156		78.10		
9	1327		166		82.95		
10	1154		144		72.12		
Minimum	1096		137		68.52		
Mean	1240		154		76.97		
Maximum	1366		166		82.95		

Table 2

The Mean classification performance of the IRCC system using MLNN classifier and bispectral entropy features

MLNN Classification results using entropy of bispectral magnitude features							
No. of training samples	4480	No. of hidden neurons	25	Output neurons	3	Training tolerance	0.03
No. of testing samples	1120	Input neurons	48			Testing tolerance	0.1
Trial	Training time (sec.)		Number of epochs		Classification accuracy (%)		
1	1183		148		73.91		
2	1192		149		74.50		
3	1157		145		72.34		
4	1237		155		77.29		
5	1301		163		81.30		
6	1321		165		82.54		
7	1419		167		83.47		
8	1355		169		84.71		
9	1519		211		84.40		
10	1140		142		71.24		
Minimum	1140		142		71.24		
Mean	1282		161		78.57		
Maximum	1519		211		84.71		

Results and Discussion. In this paper, the 16-bit digitized signals were filtered to remove the artifacts and are categorized into six traditional bands: Delta (δ) 0.1–4 Hz, Theta (θ) 4–8 Hz, Alpha (α) 8–16 Hz, Beta (β) 16–32 Hz, Gamma 1, γ_1 (32–64 Hz) and Gamma 2, γ_2 (64–100 Hz). The frequency band signals segmented into frames of equal samples and are used to extract the higher order spectra also known as polyspectral representations of higher order statistics. Further, to reduce the dimension of the bispectrum, Mean of bispectral magnitude and grand mean of the bispectral magnitude features are extracted and associated it with one of the TEP tasks. The extracted features are classified using MLNN algorithm. The classification performance of the developed models are summarized in table 1, 2 for statistical features of the $B(f_1, f_2)$ sequence. The comparison of mean training time, mean number of epochs and mean classification accuracy obtained during testing sessions using the statistical features are depicted in fig. 3–5.

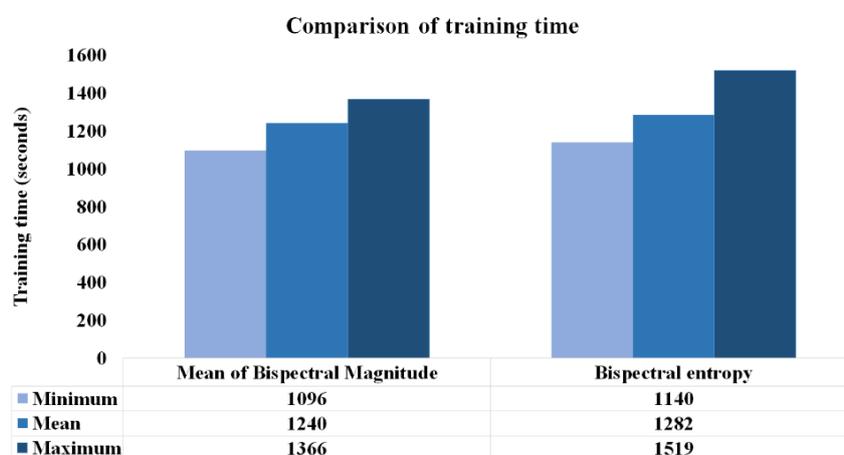


Fig. 3. Comparison of mean training time using statistical features of cross-correlation coefficients

From fig. 3, it is observed that the MLNN model based on the bispectral features has the mean training time in the range of 1096 to 1366 sec. using the mean of bispectral magnitude subset of testing set and the bispectral entropy feature set has the mean training time in the range of 1140 to 1519 sec. respectively. It is also observed that the mean training time of 1240 sec. has been obtained using the mean of bispectral magnitude features and 1519 sec. has been obtained using bispectral entropy features. The mean maximum training time of 1519 sec. has been obtained from bispectral entropy features and mean minimum training time of 1096 sec. has been obtained for the mean of the bispectral magnitude features.

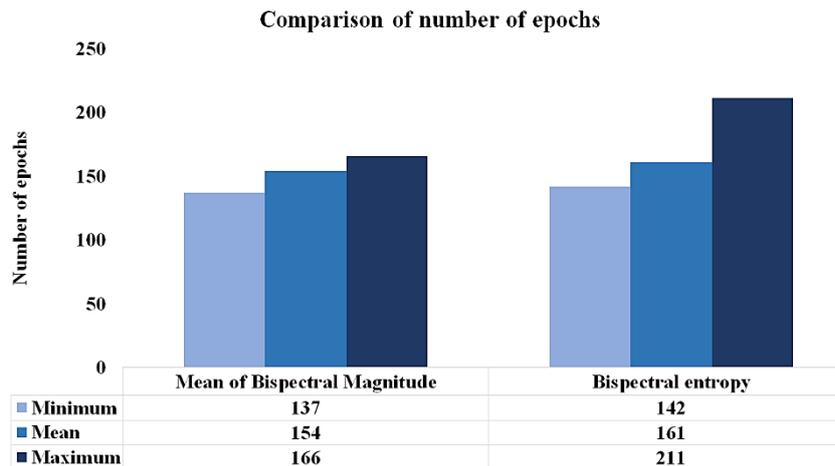


Fig. 4. Comparison of mean number of epochs using statistical features of cross-correlation coefficients

From fig. 4, it is observed that the MLNN model based on the bispectral features has the mean number of epochs in the range of 137 to 166 epochs using the mean of bispectral magnitude subset of testing set and the bispectral entropy feature set has the mean number of epochs in the range of 142 to 211 epochs respectively. It is also observed that the mean number of epochs of 154 epochs has been obtained using the mean of bispectral magnitude features and 161 epochs has been obtained using the bispectral entropy features. The mean maximum number of epochs of 211 epochs has been obtained from bispectral entropy features and the mean minimum number of epochs of 137 epochs has been obtained for the mean of bispectral magnitude features.

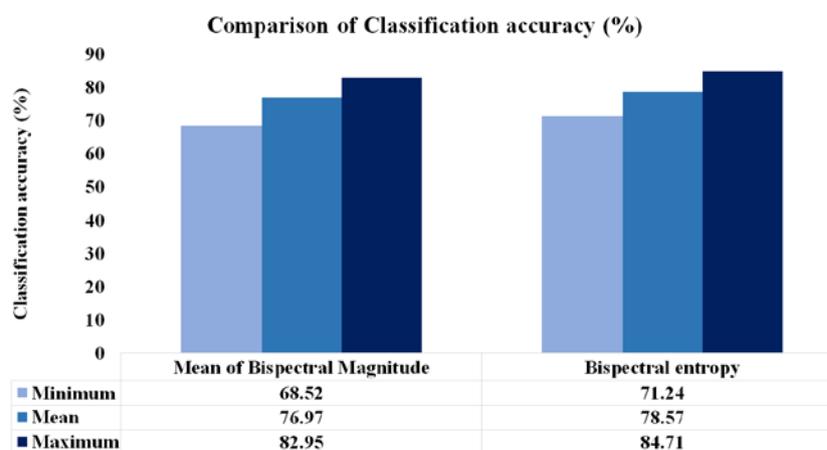


Fig. 5. Comparison of mean classification accuracy, using statistical features of cross-correlation coefficients

From fig. 5, it is observed that the MLNN model based on the bispectral features has the mean classification accuracy in the range of 68.52 to 82.95 % using the mean of bispectral magnitude subset of testing set and the bispectral entropy feature set has the mean classification accuracy in the range of 71.24 to 84.71 % respectively. It is also observed that the mean classification accuracy of 76.97 % has been obtained using the mean of the bispectral magnitude features and 78.57 % has been obtained using the bispectral entropy features. The mean max classification accuracy of 84.71 % has been obtained from the bispectral entropy features and the mean minimum classification accuracy of 68.52 % has been obtained for the mean of bispectral magnitude features.

Conclusion. The regards to the objective of this research work, a simple thought controlled intelligent robot chair with communication aid has been developed using statistical features of the bispectrum estimation and MLNN algorithm. The proposed system uses the TEP response task signals recorded from ten subjects and are segmented into six frequency band has been chosen to study the third order Fourier coefficient of the TEP tasks. Then, statistical features such as the mean of bispectral magnitude using the non-redundant region $(\Omega) = 0 \leq f_2 \leq f_1 \leq (f_1 + f_2) \leq 1$ and grand mean of the bispectral magnitude using the bispectrum region of computation are extracted. The extracted feature vectors based on third order higher order spectra features (mean and grand mean of bispectral magnitude) are distinguished easily for the different classes of TEP tasks. The feature vectors are associated with the corresponding output targets and are classified using MLNN classifiers.

The test results obtained from this analysis has a less misclassification error of 11.60 % (130/1120) samples during the testing stage. The obtained results open many possible areas of applications and improvements in thought controlled robot chair navigation and communication system for differentially enabled communities. In the future analysis, non-linear feature extraction algorithms, classification algorithms and online training sessions so as to be used to improve the recognition accuracy of the IRCC system. Further, it is propitious to explore useful characteristics of brain wave signals based on effective feature extraction and classification methods.

References

1. Whitlatch C. J., Orsulic-Jeras S. Meeting the informational, educational, and psychosocial support needs of persons living with dementia and their family caregivers. *Gerontologist*, 2018, vol. 58, no. suppl_1, pp. S58–S73.
2. Lees A. J., Blackburn N. A., Campbell V. L. The nighttime problems of Parkinson's disease. *Clinical Neuropharmacology*, Dec. 1988, vol. 11, no. 6, pp. 512–519.
3. Lacomis D., Terry Petrella J., Giuliani M. J. Causes of neuromuscular weakness in the intensive care unit: A study of ninety-two patients. *Muscle and Nerve*, May 1998, vol. 21, no. 5, pp. 610–617.
4. Jerath N. U., Simoens K., Mann D., Kollasch S., Grosland N., ..., Reddy C. G Survey of the functional priorities in patients with disability due to neuromuscular disorders. *Disability and Rehabilitation: Assistive Technology*, 2019, vol. 14, no. 2, pp. 133–137.
5. Millan J. R., Renkens F., Mouriño J., Gerstner W. Noninvasive brain-actuated control of a mobile robot by human EEG. *IEEE Transactions on Biomedical Engineering*, 2004, vol. 51, no. 6, pp. 1026–1033.
6. Philips J., Millan J. del R., Vanacker G., Lew E., Galan F., ..., Nuttin M. Adaptive shared control of a brain-actuated simulated wheelchair. *IEEE 10th International Conference on Rehabilitation Robotics*, Noordwijk, Netherlands, 13–15 June 2007, pp. 408–414.
7. Speier W., Arnold C., Lu J., Deshpande A., Pouratian N. Integrating language information with a hidden Markov model to improve communication rate in the P300 speller. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2014, vol. 22, no. 3, pp. 678–684.
8. Gürkök H., Nijholt A. Brain-computer interfaces for multimodal interaction: a survey and principles. *International Journal of Human-Computer Interaction*, Jun. 2011, vol. 28, no. 5, pp. 292–307.
9. Näätänen R., Kujala T., Light G. *Mismatch Negativity: A Window to the Brain*. Oxford University Press, 2019, 288 p.
10. Trejo L. J., Rosipal R., Matthews B. Brain-computer interfaces for 1-D and 2-D cursor control: designs using volitional control of the EEG spectrum or steady-state visual evoked potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2006, vol. 14, no. 2, pp. 225–229.
11. Wang P., Shen J., Shi J. Feature extraction of eeg for imagery left-right hands movement. *Chinese Journal of Sensors and Actuators*, 2010, vol. 9.

12. Loeser J. D., Black R. G., Christman A. Relief of pain by transcutaneous stimulation. *Journal of Neurosurgery*, 1975, vol. 42, no. 3, pp. 308–314.
13. Armiger R. S., Tenore F. V., Bishop W. E., Beaty J. D., Bridges M. M., ..., Harshbarger S. D. A real-time virtual integration environment for neuroprosthetics and rehabilitation. *The Johns Hopkins APL Technical Digest*, 2011, vol. 30, no. 3, pp. 198–206.
14. Machado S., Araújo F., Paes F., Velasques B., Cunha M., ..., Ribeiro P. EEG-based brain-computer interfaces: an overview of basic concepts and clinical applications in neurorehabilitation. *Reviews in the Neurosciences*, 2010, vol. 21, no. 6, pp. 451–468.
15. Yeo M. V. M., Li X., Shen K., Wilder-Smith E. P. V. Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science*, 2009, vol. 47, no. 1, pp. 115–124.
16. Lin C.-T., Lin F.-C., Chen S.-A., Lu S.-W., Chen T.-C., Ko L.-W. EEG-based brain-computer interface for smart living environmental auto-adjustment. *Journal of Medical and Biological Engineering*, 2010, vol. 30, no. 4, pp. 237–245.
17. Kaufmann T., Herweg A., Kübler A. Toward brain-computer interface based wheelchair control utilizing tactually-evoked event-related potentials. *Journal of NeuroEngineering and Rehabilitation*, 2014, vol. 11, no. 1, p. 7.
18. Nataraj S. K., Paulraj M. P., Yaacob S. Bin, Adom A. H. Performance comparison of TEP and VEP responses using bispectral estimation to command an intelligent robot chair with communication aid. *Indian Journal of Science and Technology*, 2015, vol. 8, no. 20, pp. 1–11.
19. Guenther F. H., Brumberg J. S., Wright E. J., Nieto-Castanon A., Tourville J. A., ..., Kennedy P. R. A wireless brain-machine interface for real-time speech synthesis. *PLoS One*, 2009, vol. 4, no. 12, p. e8218.
20. Porbadnigk A., Wester M., Calliess J., Schultz T. EEG-based speech recognition impact of temporal effects. *International Conference on Bio-Inspired Systems and Signal Processing*, Porto, Portugal, 14–17 January 2009, pp. 376–381.
21. Stamps K., Hamam Y. Towards inexpensive BCI control for wheelchair navigation in the enabled environment – a hardware survey. *International Conference on Brain Informatics*, Toronto, Canada, 28–30 August 2010, pp. 336–345.
22. Lopes A. C., Pires G., Vaz L., Nunes U. Wheelchair navigation assisted by Human-Machine shared-control and a P300-based Brain Computer Interface. *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, San Francisco, California, USA, 25–30 September 2011, pp. 2438–2444.
23. Vaughan T. M., Wolpaw J. R., Donchin E. EEG-Based communication: prospects and problems. *IEEE Transactions on Rehabilitation Engineering*, 1996, vol. 4, no. 4, pp. 425–430.
24. Nikias C. L., Raghuveer M. R. Bispectrum estimation: A digital signal processing framework. *Proceedings of the IEEE*, 1987, vol. 75, no. 7, pp. 869–891.
25. Sun L., Feng Z., Lu N., Wang B., Zhang W. An advanced bispectrum features for EEG-based motor imagery classification. *Expert Systems with Applications*, 2019, vol. 131, pp. 9–19.
26. Yuvaraj R., Acharya U. R., Hagiwara Y. A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals. *Neural Computing & Applications*, 2018, vol. 30, no. 4, pp. 1225–1235.
27. Guger C., Allison B., Edlinger G. *Brain-Computer Interface Research: A State-of-the-Art Summary*. Springer, 2013, 123 p.
28. Nataraj S. K., Yaacob S. Bin, Paulraj M. P., Adom A. H. EEG based Intelligent robot chair with communication aid using statistical cross correlation based features. *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Belfast, United Kingdom, 2–5 November 2014, pp. 12–18.
29. Kübler A., Kotchoubey B., Hinterberger T., Ghanayim N., Perelmouter J., ..., Birbaumer N. The thought translation device: A neurophysiological approach to communication in total motor paralysis. *Experimental Brain Research*, 1999, vol. 124, no. 2, pp. 223–232.
30. Nataraj S. K., Paulraj M. P., Yaacob S. Bin, Adom A. H. Statistical cross-correlation band features based thought controlled communication system. *AI Communications*, 2016, vol. 29, no. 4, pp. 497–511.
31. Teplan M. Fundamentals of EEG measurement. *Measurement Science Review*, 2002, vol. 2, no. 2, pp. 1–11.
32. Kaiser D. A. What is quantitative EEG? *Journal of Neurotherapy*, 2007, vol. 10, no. 4, pp. 37–52.
33. Tai K., Blain S., Chau T. A review of emerging access technologies for individuals with severe motor impairments. *Assistive Technology*, 2008, vol. 20, no. 4, pp. 204–219.
34. Ortner R., Grünbacher E., Guger C. *State of the Art in Sensors, Signals and Signal Processing*, 2013.
35. Raghuveer M. R., Nikias C. L. Bispectrum estimation: A parametric approach. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1985, vol. 33, no. 5, pp. 1213–1230.
36. Paulraj M. P., Sivanandam S. N. *Introduction to Artificial Neural Networks*. India, Vikas Publishing House, 2003, 236 p.

37. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence (IJCAI)*, Montréal, Canada, 20–25 August 1995, vol. 14, no. 2, pp. 1137–1145.

Information about the authors

Dr. Sathees Kumar Nataraj, Assistant Professor (Grade 3), the Department of Mechatronics Engineering, AMA International University, Bahrain. He received in Mechatronic Engineering Ph.D and Master of Science from University of Malaysia Perlis, Malaysia and Bachelor of Engineering from K. S. Rangaswamy College of Technology, India.

E-mail: satheesjuly4@gmail.com

Prof. Dr. Paulraj Murugesu Pandiyan, Principal at Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India. He holds a PhD in Computer Science and carries 32 years of Teaching Experience and more than 10 years of Research and Guiding Experience in the field of Neural Networks.

Prof. Dr. Sazali Bin Yaacob, Professor in the Department of Electrical Engineering, Universiti Kuala Lumpur Malaysian Spanish Institute, and also the head of Intelligent Automotive Systems Research Cluster focused on signal processing, driver behaviour, energy management. Received his BEng in Electrical Engineering from University of Malaysia Perlis and later pursued his MSc in System Engineering at University of Surrey and PhD in Control Engineering from University of Sheffield, United Kingdom. He received Chartered Engineer status by the Engineering Council, United Kingdom in 2005 and is also a member to the IET (UK).

Prof. Dr. Abdul Hamid Adom, Professor in Mechatronic Engineering Program (RK24), School of Mechatronic Engineering at University of Malaysia Perlis. He received his B.E, MSc and PhD from Liverpool John Moores University, UK.

Информация об авторах

Доктор Сатис Кумар Натарадж, доцент (класс 3), факультет мехатроники Международного университета АМА, Бахрейн. Получил в области мехатронной инженерии докторскую степень и звание магистра наук в Университете Малайзии Перлис, бакалавра технических наук в Технологическом колледже им. К. С. Рангасвами, Индия.

E-mail: satheesjuly4@gmail.com

Профессор, доктор Паулрадж Муругеса Пандиан, директор Технологического института Шри Рамакришны, Коимбатур, Тамилнаду, Индия. Имеет докторскую степень в области компьютерных наук, 32-летний опыт преподавания и более 10 лет исследовательского и руководящего опыта в области нейронных сетей.

Профессор, доктор Сазали бин Яакоб, профессор кафедры электротехники Малазийского испанского института Университета Куала-Лумпур, а также возглавляет исследовательский кластер интеллектуальных автомобильных систем, занимающийся обработкой сигналов, поведением водителей, управлением энергопотреблением. Получил степень бакалавра электротехники в Университете Малайзии Перлис, а затем степень магистра в области системной инженерии в Университете Суррея и доктора наук в области техники управления в Университете Шеффилда, Соединенное Королевство. Присужден статус дипломированного инженера Инженерным советом Соединенного Королевства в 2005 г., является членом IET (Великобритания).

Профессор, доктор Абдул Хамид Адом, профессор программы мехатронной инженерии (RK24) в Школе мехатронной инженерии в Университете Малайзии Перлис. Получил степень бакалавра, магистра и доктора в Ливерпульском Университете им. Джона Мурса, Великобритания.

УЧЕНЫЕ БЕЛАРУСИ
*SCIENTISTS OF BELARUS***КРОТ АЛЕКСАНДР МИХАЙЛОВИЧ**
(к 60-летию со дня рождения)

5 сентября 2020 г. исполнилось 60 лет доктору технических наук, профессору, заведующему лабораторией моделирования самоорганизующихся систем Объединенного института проблем информатики Национальной академии наук Беларуси А. М. Кроту.

Александр Михайлович родился в деревне Гольшаны Ошмянского района Гродненской области. В 1982 г. окончил Белорусский государственный университет (факультет радиофизики и электроники) и был направлен в очную аспирантуру Института технической кибернетики (ИТК) АН БССР. В 1985 г. защитил кандидатскую диссертацию, а в 1991 г. в возрасте 31 года – докторскую диссертацию. В 1997 г. ему было присвоено ученое звание профессора. С 1982 по 1993 г. работал в должностях младшего научного сотрудника, научного сотрудника, старшего научного сотрудника и ведущего научного сотрудника в ИТК АН БССР (ныне ОИПИ НАН Беларуси), где в 1993 г. создал лабораторию моделирования самоорганизующихся систем и был избран на должность заведующего. Указом Президента Республики Беларусь от 25.07.2013 г. № 328 назначен первым заместителем председателя Высшей аттестационной комиссии Республики Беларусь, в которой работал до 25 июля 2016 г., одновременно оставаясь заведующим лабораторией на общественных началах. С 1 августа 2016 г. по настоящее время – заведующий лабораторией моделирования самоорганизующихся систем ОИПИ НАН Беларуси.

Одновременно работал по совместительству в должности профессора в Белорусском государственном университете информатики и радиоэлектроники с 1993 по 1997 г. и в Белорусском государственном университете на факультете радиофизики и электроники с 1996 по 1998 г., на факультете прикладной математики и информатики с 2008 по 2016 г.

Исследовательская деятельность А. М. Крота посвящена теории и методам моделирования процессов самоорганизации в сложных системах, аналитическим моделям природных явлений самоорганизации и неустойчивых процессов формирования в активных средах (в частности, в движущейся вязкой жидкости и газопылевых распределенных космических средах), методам нелинейной динамики, алгоритмам цифрового спектрального анализа и обработке сигналов (изображений), теории вычислительной сложности и методам вычислительной аэрогидродинамики.

Александр Михайлович разработал теорию матричной декомпозиции хаотических аттракторов в пространстве состояний сложных нелинейных динамических систем, предложил новую модель квазистационарных процессов на основе обобщенного спектрального анализа в дискретных базисах собственных функций, обосновал принцип вычислительного дуализма между стационарными и квазистационарными системами и на его основе – метод собственных преобразований в алгебраических полях, синтезировал эффективные алгоритмы цифровой спектральной обработки сигналов и изображений.

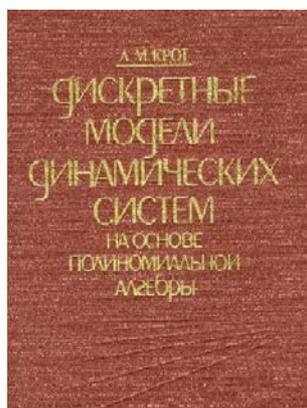
Им была создана статистическая теория формирования гравитирующих космогонических тел, в рамках которой предложены модели и эволюционные уравнения статистической механики формирования планетарных систем, выведено нелинейное времязависимое Шрёдингер-подобное волновое уравнение, открыт закон звездного мироздания для внесолнечных планетарных систем, связывающий температуру, размер и массу каждой звезды.

Опубликовал свыше 300 научных работ, среди которых три монографии:

Дискретные модели динамических систем на основе полиномиальной алгебры. – Минск: Наука і тэхніка, 1990. – 312 с.;

Быстрые алгоритмы и программы цифровой спектральной обработки сигналов и изображений (соавт. Е. Б. Минервина). – Минск: Наука і тэхніка, 1995. – 407 с.;

Статистическая теория формирования гравитирующих космогонических тел. – Минск: Беларус. навука, 2012. – 448 с.



Более 100 трудов и глав монографий Александра Михайловича опубликованы на английском языке в международных изданиях, в том числе 20 – в авторитетных зарубежных изданиях, имеющих импакт-фактор, таких как *Chaos, Solitons and Fractals*, *Planetary and Space Science*, *Electronics Letters*, *Advances in Space Research*, «Доклады АН СССР», «Журнал вычислительной математики и математической физики», «Радиотехника и электроника» и др.

Подготовил восемь кандидатов наук (канд. техн. наук Е. Б. Минервину, канд. техн. наук Б. Х. Абдурахманова, канд. техн. наук М. Н. Долгих, канд. физ.-мат. наук Н. А. Романовскую, канд. физ.-мат. наук В. Ф. Дайлюденко, канд. техн. наук В. М. Демко, канд. физ.-мат. наук В. А. Балдина, канд. техн. наук Г. А. Прокоповича) и в настоящее время является научным руководителем аспирантов и соискателей.

Активный участник и организатор, докладчик, член программных комитетов, председатель и руководитель сессий ряда международных симпозиумов и конференций, проводимых под эгидой международных научных обществ за рубежом: в США (1998, 1999, 2000, 2001), Греции (1997, 1998, 2002), Италии (1999, 2001), Германии (1998, 2009, 2016), Финляндии (2000), на Кипре (2000), в Дании (2002) и Австрии (2006, 2009).

Областью практического применения результатов А. М. Крота являются алгоритмы и программные средства автоматизированных систем управления динамическими испытаниями изделий машиностроения, а также автоматизированных систем цифровой обработки сигналов и изображений, внедренных на предприятиях России и Беларуси. В настоящее время практические результаты его работ направлены на развитие и применение компьютерных технологий численного моделирования и виртуального анализа аэрогидродинамических процессов в проточных частях реальных и проектируемых изделий машиностроения, разработку компьютерных систем анализа и распознавания динамических процессов на основе нейронных сетей.

А. М. Крот удостоен звания лауреата премии комсомола Минщины (1988) в области науки и техники за работу «Разработка и внедрение алгоритмических и программно-технических средств автоматизации испытаний изделий в машиностроении», премии Ленинского комсомола Беларуси в области науки и техники (1990) за работу «Создание и практическое применение средств цифровой обработки сигналов и изображений», премии Президиума центрального правления НТО радиотехники, электроники и связи имени А. С. Попова (1991) за работу в области цифровой спектральной обработки сигналов и изображений, премии Президиума НАН Беларуси за лучшую научную работу молодых ученых Беларуси (1993) «Методы и алгоритмы цифровой обработки сигналов для анализа и моделирования динамических систем», премии Министерства промышленности Республики Беларусь в области науки и техники (2010) за ра-

боту «Создание и внедрение в процесс проектирования новой техники методики компьютерного моделирования и нелинейного анализа воздушных потоков в тракте сельскохозяйственных машин и взаимодействия воздушных потоков с растительной массой» в номинации «Информатика, вычислительная техника, телекоммуникации, автоматизированные системы управления технологической подготовкой производства», а также персональной надбавки от Президента Республики Беларусь (2007) за выдающийся вклад в социально-экономическое развитие Беларуси и гранта Президента Республики Беларусь в науке (2019) на развитие теоретических основ возникновения хаотических волновых процессов в планетарных системах, аэрогидродинамических устройствах и электронных схемах и создание информационных технологий компьютерного моделирования аэрогидродинамических и электродинамических процессов для совершенствования продукции машиностроения и электронных устройств.

Приказом Государственного комитета по науке и технологиям Республики Беларусь от 20.09.2019 г. № 270 профессор А. М. Крот назначен руководителем научной секции «Программно-математические методы решения фундаментальных задач и технических проблем по созданию технологий будущего» (государственного экспертного совета № 6).

Александр Михайлович отличается искренностью и честностью суждений, добросовестностью и точностью в работе, пользуется заслуженным уважением коллектива. Известны его принципиальность, верность научному методу и объективному критическому суждению, готовность всегда поддержать коллектив, бороться за интересы каждого сотрудника лаборатории, энтузиазм в ведении научно-исследовательской и проектно-организационной работы, твердое созидательное мироощущение, базирующееся на гуманитарных ценностях и любви к нашему Отечеству, умение выслушать и принять к сведению различные точки зрения в коллективе, а также индивидуальный подход к каждому сотруднику и ученику.

Благодаря научной и образовательной деятельности А. М. Крота в лаборатории моделирования самоорганизующихся систем ОИПИ НАН Беларуси растет достойная смена молодых ученых, которые будут работать на будущее нашей страны, на ее развитие и процветание.

Александр Михайлович работает над теоретическими исследованиями в области фундаментальной науки, изучая основы физических процессов мироздания, готовит к выходу англоязычную версию гравитационной теории, внедряет и применяет свои разработки на практике при выполнении моделирования различных механизмов и явлений на заказ от государственных учреждений и заводов-подрядчиков. Благодаря внедрению новых моделей для исследования коллективу лаборатории удалось усовершенствовать ряд инженерных систем. Лаборатория под руководством А. М. Крота развивает международное сотрудничество в рамках обмена опытом в смежных сферах и продвижения учеников на престижные публичные научные площадки.

Коллеги и сотрудники ОИПИ НАН Беларуси поздравляют Александра Михайловича с высокими достижениями в научной деятельности и в связи с 60-летием. Желают юбиляру новых творческих идей и успехов во всех его начинаниях, а также крепкого здоровья и всегда хорошего настроения.

М. Я. Ковалев, Е. Н. Ефимов, В. М. Демко, А. А. Крот

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

ДУДКИН АЛЕКСАНДР АРСЕНТЬЕВИЧ
(к 70-летию со дня рождения)



13 октября 2020 г. исполнилось 70 лет известному ученому в области технической кибернетики и информатики Александру Арсентьевичу Дудкину.

Александр Арсентьевич родился в г. Городок Витебской области. В 1968 г. после окончания средней школы № 1 поступил на физико-математический факультет Витебского государственного педагогического института им. С. М. Кирова (в настоящее время Витебский государственный университет им. П. М. Машерова).

После окончания в 1972 г. с отличием института по специальности «Математика» четыре года учительствовал в Чашникском районе: преподавал математику, был директором Горской восьмилетней школы и заместителем директора Черейской средней школы. В 1976 г. поступил в аспирантуру Института технической кибернетики АН БССР. В 1977–1978 гг. проходил срочную службу в рядах Советской Армии.

Вся дальнейшая трудовая деятельность А. А. Дудкина связана с Институтом технической кибернетики (ИТК), позже переименованным в Объединенный институт проблем информатики (ОИПИ) НАН Беларуси. После окончания аспирантуры с 1980 г. работал в лаборатории системного программирования и логического синтеза в должности младшего научного сотрудника, с 1987 г. – научного сотрудника, а с 1988 г. – старшего научного сотрудника. В 1989 г. был зачислен старшим научным сотрудником в лабораторию проблем построения САПР СБИС Брестского отделения ИТК АН БССР, с 1994 г. стал ведущим научным сотрудником, заместителем заведующего лабораторией идентификации систем ОИПИ НАН Беларуси, а в 2015 г. – заведующим этой лабораторией. Параллельно с 1989 г. преподавал в БНТУ и БГУИР, в 2008–2014 гг. был заведующим филиалом кафедры ЭВМ БГУИР при ОИПИ НАН Беларуси, председателем государственной экзаменационной комиссии и государственной аттестационной комиссии для проведения итоговой аттестации аспирантов (докторантов, соискателей) БГУ, БрГТУ и БНТУ.

В 1987 г. в ИТК АН БССР защитил кандидатскую диссертацию «Синтез быстродействующих дискретных устройств в базисе ПЛМ на основе алгоритмов решения матричных задач группирования» по специальности «Системы автоматизации проектирования» (научный руководитель – член-корреспондент АН БССР А. Д. Закревский). В 2010 г. в БГУИР защитил докторскую диссертацию «Методы обработки и анализа цифровых изображений топологических слоев интегральных микросхем» по специальности «Системный анализ, обработка и управление информацией» (научный консультант – лауреат Государственной премии Республики Беларусь, профессор Р. Х. Садыхов). В 1991 г. Александру Арсентьевичу было присвоено ученое звание старшего научного сотрудника, в 2010 г. – звание доцента и в 2016 г. – звание профессора.

Научные работы А. А. Дудкина относятся к следующим основным направлениям исследований:

1. Автоматизация проектирования дискретных устройств и цифровых СБИС.
2. Цифровая обработка сигналов.
3. Модели и методы систем компьютерного зрения.

В рамках первого направления А. А. Дудкин разработал методы построения быстродействующих схем на базе матричных СБИС, с использованием которых создан ряд пакетов прикладных программ и систем проектирования устройств управления СБИС, внедренных на предприятиях электронной промышленности Беларуси и России: ЦКБ «Алмаз» и ЦНИИ «Агат», ПКБ «Альфа», МЧЗ «Луч» и др. Основные итоги этой работы описаны в статьях юбилея [1, 2]. Следует отметить также его вклад в создание в 1992–1994 гг. под научным руко-

водством профессора, лауреата Государственной премии Республики Беларусь В. А. Мищенко сквозной системы функционально-логического проектирования цифровых СБИС на базе библиотек. В качестве одного из языков описания поведения синтезируемых схем в системе использовался VHDL [3, 4].

В рамках второго направления под руководством А. А. Дудкина и при его активном участии разработана информационная технология обработки, анализа и идентификации изображений интегральных схем, которая реализована в автоматизированных системах восстановления топологии таких схем по изображениям отдельных слоев. Она внедрена в концерне «Планар» и НПО «Интеграл». К наиболее существенным результатам в этой области можно отнести следующие:

- комплекс алгоритмов сшивки цифровых изображений топологических слоев СБИС, основанных на использовании специальных эвристик, что способствует корректному восстановлению топологии СБИС;
- ряд алгоритмов сегментации изображений топологических слоев с возможностью проблемно-ориентированной настройки на предварительную и постобработку с удалением шумов на изображениях без искажений топологии СБИС, обладающих устойчивостью к изменениям условий съемки;
- методы идентификации объектов на изображениях топологических слоев СБИС и фотошаблонов, основанные на совместном использовании быстрых параметрических алгоритмов сегментации и модифицированном преобразовании Хафа, что позволило снизить вычислительную сложность алгоритмов и повысить достоверность идентификации за счет построения многомерных признаков идентификации объектов СБИС;
- нейросетевой подход к анализу объектов топологических слоев СБИС в виде многослойного персептрона, неокогнитрона, их ансамблей, а также ансамблей сверточных сетей для реализации многоуровневого процесса распознавания, что позволило повысить стабильность классификации дефектов топологии.

Существенным теоретическим результатом второго направления исследований является разработка специализированных методов обучения иерархических многослойных нейронных сетей, которые учитывают нечеткость геометрических и яркостных описаний объектов идентификации. Разработан метод выбора архитектур нейронных сетей, основанный на экспериментах с различными конфигурациями, что позволяет пропустить этап повышения информативности без ущерба качеству обучения. Использование набора классификаторов, в котором по результатам классификации на тестовой выборке производится выбор наилучшей модели классификатора для обработки изображений, обеспечивает эффективное машинное обучение. Обучение нейросетевых модулей идентификации основано на алгоритме инкрементного обучения ансамблей нейронных сетей и отличается от известных отсутствием этапа эвристического выбора размеров скрытых слоев отдельных нейронных сетей ансамбля за счет синтеза многослойного персептрона. Результаты второго направления исследований описаны в монографии [5].

В рамках третьего направления также получен ряд новых научных результатов. Предложена нейросетевая технология прогнозирования состояния технических объектов по телеметрии, заданной многомерными временными рядами. Разработанный ансамбль нейронных сетей позволяет осуществлять более точное прогнозирование по сравнению с одиночными моделями и обеспечивает дообучение нейронных сетей в процессе их функционирования. Методика находится в стадии опытной эксплуатации в Центре управления полетом белорусского космического аппарата для диагностики подсистем целевой аппаратуры, энергообеспечения и корректирующей двигательной установки космического аппарата [6–8].

Разработана методика идентификации объектов дистанционного зондирования Земли на основе нейронных сетей и их ансамблей. Выделение площадных объектов на мультиспектральных снимках основывается на использовании яркостных характеристик спектральных каналов исходного изображения совместно с фрактальными и текстурными характеристиками. Полученные результаты нашли применение в системах точного земледелия при разработке алгоритмов мониторинга состояния посевов сельскохозяйственных культур и лесных массивов [9–11].

Разработан комплекс алгоритмов сжатия гиперспектральных изображений на основе блочно-субполосного вложенного алгоритма вейвлет-разложения с построением дерева значимости битовых плоскостей и удалением наименее значимых. Особенности алгоритмов являются оптимизация работы с памятью, ускорение поиска значимых вейвлет-коэффициентов за счет внедрения пирамиды с аппроксимирующими коэффициентами и применение контекстного моделирования для адаптации сжатия к специфике исходных данных [12, 13].

Профессор А. А. Дудкин опубликовал свыше 250 научных работ, в том числе две монографии, три книги по математическому обеспечению ЭВМ и четыре учебника. Под его руководством защищены две диссертации.

А. А. Дудкин был секретарем координационного совета государственной программы научных исследований «Информационные и космические технологии» (2011–2016 гг.), является членом советов по защите диссертаций при БГУИР и ОИПИ НАН Беларуси, членом государственного экспертного совета № 12, членом ученого совета ОИПИ НАН Беларуси, членом научного совета «Информационные технологии и системы» ОИПИ НАН Беларуси. Он был председателем и членом ряда организационных и программных комитетов международных конференций.

Профессор А. А. Дудкин является членом Белорусского общества по исследованию операций, Белорусского отделения международного общества нейронных сетей и Белорусской ассоциации по анализу и распознаванию изображений.

Коллектив ОИПИ НАН Беларуси, коллеги и друзья сердечно поздравляют Александра Арсентьевича, счастливого мужа и отца троих детей, а также деда троих внуков, с юбилеем, желают ему крепкого здоровья, личного счастья, активной трудовой деятельности и новых научных достижений.

Литература

1. Закревский, А. Д. Пакет программ синтеза комбинационных схем в базисе ПЛМ / А. Д. Закревский, А. А. Дудкин, П. Н. Бибило // УСиМ. – 1984. – № 1. – С. 21–23.
2. Система ЛОГИКА-М синтеза управляющих устройств в базисе ПЛМ и микропроцессоров / А. Д. Закревский [и др.] // УСиМ. – 1987. – № 2. – С. 31–35.
3. Система автоматизированного проектирования СБИС. Процедуры проектирования / А. А. Дудкин [и др.] ; под науч. ред. В. А. Мищенко // Материалы по математическому обеспечению ЭВМ. – Минск : Ин-т техн. кибернетики АН Беларуси, 1992. – 91 с.
4. Алгоритмы и подсистемы автоматизированного логического проектирования цифровых СБИС / А. А. Дудкин [и др.] // Материалы по математическому обеспечению ЭВМ. – Минск : Ин-т техн. кибернетики АН Беларуси, 1994. – 126 с.
5. Дудкин, А. А. Обработка изображений в проектировании и производстве интегральных схем / А. А. Дудкин, Р. Х. Садыхов. – Минск : ОИПИ НАН Беларуси, 2008. – 270 с.
6. Прототип программной нейросетевой системы контроля телеметрической информации / А. А. Дудкин [и др.] ; под общ. ред. М. И. Макарова // Ракетно-космическая техника. Информационные системы и технологии : науч. тр. : в 2 т. – М. : НИИ КС им. А. А. Максимова, 2012. – Т. 1, главы 1-V. – С. 303–317.
7. Doudkin, A. Ensembles of neural network for telemetry multivariate time series forecasting / A. Doudkin, Y. Marushko ; ed.: V. V. Krasnoproshin, S. V. Ablameyko // Proc. of the 13th Intern. Conf. on Pattern Recognition and Information Processing (PRIP'2016), 3–5 Oct. 2016, Minsk. – Minsk : BSU, 2016. – P. 53– 62. https://doi.org/10.1007/978-3-319-54220-1_6
8. Дудкин, А. А. Нейросетевая технология обработки сигналов от средств контроля технических объектов / А. А. Дудкин, Е. Е. Марушко // Доклады БГУИР. – 2018. – № 5(115). – С. 85–91.
9. Special areas detection and recognition on agricultural fields images / V. Ganchenko [et al.] ; ed.: R. Duro, F. Pena // Digital Image and Signal Processing for Measurement Systems. – 2012. – Ch. 8. – P. 201–233.

10. Дудкин, А. А. Подсистема параллельной обработки изображений для мониторинга состояния лесных угодий / А. А. Дудкин // Известия ЮФУ. Технические науки. – 2014. – № 12(161). – С. 129–137.

11. Ganchenko, V. Agricultural vegetation monitoring based on aerial data using convolutional neural networks / V. Ganchenko, A. Doudkin // Optical Memory and Neural Networks. – 2019. – Vol. 28, iss. 2. – P. 129–134.

12. Дудкин, А. А. Алгоритм сжатия гиперспектральных данных дистанционного зондирования Земли / А. А. Дудкин, Д. Ю. Перцев // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2017. – № 1. – С. 120–126.

13. Marushko, E. E. Methods of using ensembles of heterogeneous models to identify remote sensing objects / E. E. Marushko, A. A. Doudkin // Pattern Recognition and Image Analysis. – 2020. – Vol. 30, no. 2. – P. 211–216.

В. В. Старовойтов

Правила для авторов

Редакция журнала «Информатика» просит авторов руководствоваться приведенными ниже правилами:

1. Статьи принимаются в редакцию через электронную систему подачи по адресу <http://inf.grid.by> в формате файлов текстовых редакторов Microsoft Word. Основной текст статьи не должен превышать 17 стр., включая рисунки, таблицы и достаточное количество наиболее актуальных ссылок; обзорной статьи – 10 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.

Изложенный в статье материал должен быть четко структурированным: введение, цели и задачи, методы, результаты, заключение (выводы).

2. 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).

3. Статья в обязательном порядке должна иметь следующую структуру: индекс по универсальной десятичной классификации (УДК); инициалы и фамилии всех авторов, название статьи, полное название учреждений, где работают авторы, с указанием города, страны, аннотацию (150–250 слов), подрисуночные надписи, названия таблиц и ключевые слова (7–10) на русском и английском языках, адрес электронной почты контактного лица.

4. Аннотация (авторское резюме) должна кратко представлять результаты работы и быть информативной, содержательной. Приветствуется структура аннотации, повторяющая структуру статьи и включающая введение, цели и задачи, методы, результаты, заключение.

5. 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. Подрисуночные подписи с расшифровкой всех позиций, представленных на рисунке, набираются шрифтом гарнитуры основного текста размером 9 пт. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала.

6. 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, CHCl_3); цифры (римские и арабские); векторы; индексы (верхние и нижние), являющиеся сокращениями слов. Курсивом набираются латинские буквы, символы физических величин (в том числе и в индексе).

7. Сокращения в тексте статьи (за исключением единиц измерения) могут быть использованы только после упоминания полного термина. Единицы измерения физических величин следует приводить в Международной системе единиц (СИ).

8. Цитируемые в статье фамилии авторов теорем, теорий, законов и т. д. следует приводить в скобках на языке оригинала после русского написания. Например, теорема Эйлера (Euler).

9. 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 приводится список цитированных источников в *романском* (латинском) алфавите со следующей структурой: авторы (транслитерация), название статьи в транслитерированном варианте [перевод названия статьи на английский язык в квадратных скобках], название русскоязычного источника (транслитерация) [перевод названия источника на английский язык – парафраз (для журналов можно не делать)], выходные данные с обозначениями на английском языке.

11. Поступившие в редакцию статьи направляются на рецензирование специалистам. Основными критериями целесообразности публикации являются новизна и информативность статьи. Если по рекомендациям рецензента статья возвращается автору на доработку, то переработанная рукопись вновь рассматривается редколлегией. Статьи не по профилю журнала возвращаются авторам после заключения редколлегии.

12. Статьи, направляемые на доработку, должны быть возвращены в исправленном виде с ответами на все замечания.

13. Редакция журнала предоставляет возможность первоочередного опубликования статей, представленных лицами, которые осуществляют послевузовское обучение (аспирантура, докторантура, соискательство) в год завершения обучения.

14. Авторы несут ответственность за направление в редакцию статей, уже опубликованных ранее или принятых к публикации другими изданиями.

15. Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи. Окончательное решение о публикации принимается редакционной коллегией.

Индексы

00827

для индивидуальных
подписчиков

008272

для предприятий и
организаций