

p-ISSN 1607-3274
e-ISSN 2313-688X

Радіоелектроніка
Інформатика
Управління



Radio Electronics
Computer Science
Control

Радиоэлектроника
Информатика
Управление



2021/2



Національний університет «Запорізька політехніка»

Радіоелектроніка, інформатика, управління

Науковий журнал

Виходить чотири рази на рік

№ 2(57) 2021

Заснований у січні 1999 року.

Засновник і видавець – Національний університет «Запорізька політехніка».

ISSN 1607-3274 (друкований), ISSN 2313-688X (електронний).

Запоріжжя

НУ «Запорізька політехніка»

2021

National University «Zaporizhzhia Polytechnic»

Radio Electronics, Computer Science, Control

The scientific journal

Published four times per year

№ 2(57) 2021

Founded in January 1999.

Founder and publisher – National University «Zaporizhzhia Polytechnic».

ISSN 1607-3274 (print), ISSN 2313-688X (on-line).

Zaporizhzhia

NU «Zaporizhzhia Polytechnic»

2021

Национальный университет «Запорожская политехника»

Радиоэлектроника, информатика, управление

Научный журнал

Выходит четыре раза в год

№ 2(57) 2021

Основан в январе 1999 года.

Основатель и издатель – Национальный университет «Запорожская политехника».

ISSN 1607-3274 (печатный), ISSN 2313-688X (электронный).

Запорожье

НУ «Запорожская политехника»

2021

Науковий журнал «Радіоелектроніка, інформатика, управління» (скорочена назва – РІУ) видається Національним університетом «Запорізька політехніка» (НУ «Запорізька політехніка») з 1999 р. періодичністю чотири номери на рік.

Зареєстровано у Міністерстві юстиції України 19.11.2019 р. (Свідчення про державну реєстрацію друкованого засобу масової інформації серія КВ № 24220-14060 ПР.)

ISSN 1607-3274 (друкований), ISSN 2313-688X (електронний).

Наказом Міністерства освіти і науки України № 409 від 17.03.2020 р. «Про затвердження рішень Атестаційної колегії Міністерства щодо діяльності спеціалізованих вчених рад від 06 березня 2020 року» журнал включений до переліку наукових фахових видань України в категорії «А» (найвищий рівень), в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора наук і доктора філософії (кандидата наук).

Журнал включений до польського Переліку наукових журналів та рецензованих матеріалів міжнародних конференцій з присвоєною кількістю балів (додаток до оголошення Міністра науки та вищої освіти Республіки Польща від 31 липня 2019 р.: № 16981).

В журналі безкоштовно публікуються наукові статті англійською, російською та українською мовами.

Правила оформлення статей подано на сайті: <http://ric.zntu.edu.ua/information/authors>.

Журнал забезпечує **безкоштовний відкритий он-лайн доступ** до повнотекстових публікацій.

Журнал дозволяє авторам мати авторські права і зберігати права на видання без обмежень. Журнал дозволяє користувачам читати, завантажувати, копіювати, поширювати, друкувати, шукати або посилається на повні тексти своїх статей. Журнал дозволяє повторне використання його вмісту у відповідності Creative Commons ліцензією CC BY-SA..

Опублікованим статтям присвоюється унікальний ідентифікатор цифрового об'єкта DOI.

Журнал входить до наукометричної бази Web of Science.

Журнал реферується та індексується у провідних міжнародних та національних реферативних журналах і наукометричних базах даних, а також розміщується у цифрових архівах та бібліотеках з безкоштовним доступом у режимі on-line, повний перелік яких подано на сайті: <http://ric.zntu.edu.ua/about/editorialPolicies#custom-0>.

Журнал розповсюджується за Каталогом періодичних видань України (передплатний індекс – 22914).

Тематика журналу: телекомунікації та радіоелектроніка, програмна інженерія (включаючи теорію алгоритмів і програмування), комп'ютерні науки (математичне і комп'ютерне моделювання, оптимізація і дослідження операцій, управління в технічних системах, міжмашинна і людино-машинна взаємодія, штучний інтелект, включаючи системи, засновані на знаннях, і експертні системи, інтелектуальний аналіз даних, розпізнавання образів, штучні нейронні і нейро-нечіткі мережі, нечітку логіку, колективний інтелект і мультиагентні системи, гібридні системи), комп'ютерна інженерія (апаратне забезпечення обчислювальної техніки, комп'ютерні мережі), інформаційні системи та технології (структури та бази даних, системи, засновані на знаннях та експертні системи, обробка даних і сигналів).

Усі статті, пропонувані до публікації, одержують **об'єктивний розгляд**, що оцінюється за суттю без урахування раси, статі, віросповідання, етнічного походження, громадянства або політичної філософії автора(ів).

Усі статті проходять двоступінчасте закриті (анонімне для автора) **рецензування** штатними редакторами і незалежними рецензентами – провідними вченими за профілем журналу.

РЕДАКЦІЙНА КОЛЕГІЯ

Головний редактор – Субботін Сергій Олександрович – доктор технічних наук, професор, завідувач кафедри програмних засобів, Національний університет «Запорізька політехніка», Україна.

Заступник головного редактора – Піза Дмитро Макарович – доктор технічних наук, професор, директор інституту інформатики та радіоелектроніки, професор кафедри радіотехніки та телекомунікацій, Національний університет «Запорізька політехніка», Україна.

Члени редколегії:

Андрюлідакіс Іосіф – доктор філософії, голова департаменту телефонії Центру обслуговування мереж, Університет Яніни, Греція;

Бодяньскій Євгеній Володимирович – доктор технічних наук, професор, професор кафедри штучного інтелекту, Харківський національний університет радіоелектроніки, Україна;

Веннекенс Юст – доктор філософії, доцент, доцент факультету інженерних технологій (кампус Де Наїр), Католицький університет Льовена, Бельгія;

Рекомендовано до видання Вченою радою НУ «Запорізька політехніка», протокол № 11/21 від 31.05.2021.

Журнал зверстаний редакційно-видавничим відділом НУ «Запорізька політехніка».

Веб-сайт журналу: <http://ric.zntu.edu.ua>.

Адреса редакції: Редакція журналу «РІУ», Національний університет «Запорізька політехніка», вул. Жуковського, 64, м. Запоріжжя, 69063, Україна.

Тел: (061) 769-82-96 – редакційно-видавничий відділ

E-mail: rvv@zntu.edu.ua

Вольф Карстен – доктор філософії, професор, професор кафедри технічної інформатики, Дортмундський університет прикладних наук та мистецтв, Німеччина;

Вуттке Ганс-Дітріх – доктор філософії, доцент, провідний науковий співробітник інституту технічної інформатики, Технічний університет Льменау, Німеччина;

Горбань Олександр Миколайович – доктор фізико-математичних наук, професор, професор факультету математики, Університет Лестера, Велика Британія;

Городничий Дмитро Олегович – доктор філософії, кандидат технічних наук, доцент, провідний науковий співробітник Дирекції науки та інженерії, Канадська агенція прикордонної служби, Канада;

Дробахін Олег Олегович – доктор фізико-математичних наук, професор, перший проректор, Дніпровський національний університет імені Олеса Гончара, Україна;

Зайцева Олена Миколаївна – кандидат фізико-математичних наук, професор, професор кафедри інформатики, Жилінський університет в Жиліні, Словаччина;

Камеяма Мічітака – доктор наук, професор, професор факультету науки та інженерії, Університет Ішіномакі Сеншу, Японія;

Карташов Володимир Михайлович – доктор технічних наук, професор, завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Харківський національний університет радіоелектроніки, Україна;

Левашенко Віталій Григорович – кандидат фізико-математичних наук, професор, завідувач кафедри інформатики, Жилінський університет в Жиліні, Словаччина;

Луенго Давид – доктор філософії, професор, завідувач кафедри теорії сигналів та комунікацій, Мадридський політехнічний університет, Іспанія;

Марковска-Качмар Урсула – доктор технічних наук, професор, професор кафедри обчислювального інтелекту, Вроцлавська політехніка, Польща;

Олійник Андрій Олександрович – кандидат технічних наук, доцент, доцент кафедри програмних засобів, Національний університет «Запорізька політехніка», Україна;

Павліков Володимир Володимирович – доктор технічних наук, старший науковий співробітник, проректор з наукової роботи, Національний аерокосмічний університет ім. Н.Е. Жуковського «ХАІ», Україна;

Папшицький Марцін – доктор наук, професор, професор відділу інтелектуальних систем, Дослідний інститут систем Польської академії наук, м. Варшава, Польща;

Скруський Степан Юрійович – кандидат технічних наук, доцент, доцент кафедри комп'ютерних систем і мереж, Національний університет «Запорізька політехніка», Україна;

Табунчик Галина Володимирівна – кандидат технічних наук, професор, професор кафедри програмних засобів, Національний університет «Запорізька політехніка», Україна;

Тригано Томас – доктор філософії, старший викладач кафедри електричної та електронної інженерії, Інженерний коледж ім. С. Шамон, м. Ашдод, Ізраїль;

Хенке Карстен – доктор технічних наук, професор, науковий співробітник факультету інформатики та автоматизації, Технічний університет Льменау, Німеччина;

Шарпанських Олексій Альбертович – доктор філософії, доцент, доцент факультету аерокосмічної інженерії, Делфтський технічний університет, Нідерланди.

РЕДАКЦІЙНО-КОНСУЛЬТАТИВНА РАДА

Аррас Пітер – доктор філософії, доцент, доцент факультету інженерних технологій (кампус Де Наїр), Католицький університет Льовена, Бельгія;

Ліснянський Анатолій – кандидат фізико-математичних наук, головний науковий експерт, Ізраїльська електрична корпорація, Хайфа, Ізраїль;

Мадрицх Христіан – доктор філософії, професор факультету інженерії та інформаційних технологій, Університет прикладних наук Каринфії, Австрія;

Маркосян Мгер Вардкесович – доктор технічних наук, професор, директор Єреванського науково-дослідного інституту засобів зв'язку, професор кафедри телекомунікацій, Російсько-вірменський університет, м. Єреван, Вірменія;

Рубель Олег Володимирович – кандидат технічних наук, доцент факультету інженерії, Університет МакМастера, Гамільтон, Канада;

Тавхелідзе Автанділ – кандидат фізико-математичних наук, професор, професор школи бізнесу, технології та освіти, Державний університет ім. Ілії Чавчавадзе, Тбілісі, Грузія;

Уреутью Дору – доктор фізико-математичних наук, професор, професор кафедри електроніки та обчислювальної техніки, Трансильванський університет в Брашові, Румунія;

Шульц Пітер – доктор технічних наук, професор, професор факультету інженерії та комп'ютерних наук, Гамбургський університет прикладних наук (HAW Hamburg), Гамбург, Німеччина.

The scientific journal «Radio Electronics, Computer Science, Control» is published by the National University «Zaporizhzhia Polytechnic» NU «Zaporizhzhia Polytechnic» since 1999 with periodicity four numbers per year.

The journal is registered by the Ministry of Justice of Ukraine in 19.11.2019. (State Registration Certificate of printed mass media series KB № 24220-14060 IIP).

ISSN 1607-3274 (print), ISSN 2313-688X (on-line).

By the Order of the Ministry of Education and Science of Ukraine from 17.03.2020 № 409 “On approval of the decision of the Certifying Collegium of the Ministry on the activities of the specialized scientific councils dated 06 March 2020” journal is included in the list of scientific specialized periodicals of Ukraine in category “A” (highest level), where the results of dissertations for Doctor of Science and Doctor of Philosophy may be published.

The journal is included to the Polish List of scientific journals and peer-reviewed materials from international conferences with assigned number of points (Annex to the announcement of the Minister of Science and Higher Education of Poland from July 31, 2019: Lp. 16981).

The journal publishes scientific articles in English, Russian, and Ukrainian free of charge.

The article formatting rules are presented on the site: <http://ric.zntu.edu.ua/information/authors>.

The journal provides policy of on-line open (free of charge) access for full-text publications. The journal allow the authors to hold the copyright without restrictions and to retain publishing rights without restrictions. The journal allow readers to read, download, copy, distribute, print, search, or link to the full texts of its articles. The journal allow reuse and remixing of its content, in accordance with Creative Commons license CC BY-SA.

Published articles have a unique digital object identifier (DOI).

The journal is included into Web of Science.

The journal is abstracted and indexed in leading international and national abstracting journals and scientometric databases, and also placed to the digital archives and libraries with a free on-line access, full list of which is presented at the site: <http://ric.zntu.edu.ua/about/editorialPolicies#custom-0>.

The journal is distributed by the Catalogue of Ukrainian periodicals (the catalog number is 22914).

The journal scope: telecommunications and radio electronics, software engineering (including algorithm and programming theory), computer science (mathematical modeling and computer simulation, optimization and operations research, control in technical systems, machine-machine and man-machine interfacing, artificial intelligence, including data mining, pattern recognition, artificial neural and neuro-fuzzy networks, fuzzy logic, swarm intelligence and multiagent systems, hybrid systems), computer engineering (computer hardware, computer networks), information systems and technologies (data structures and bases, knowledge-based and expert systems, data and signal processing methods).

All articles proposed for publication receive an objective review that evaluates substantially without regard to race, sex, religion, ethnic origin, nationality, or political philosophy of the author(s).

All articles undergo a two-stage blind peer review by the editorial staff and independent reviewers – the leading scientists on the profile of the journal.

EDITORIAL BOARD

Editor-in-Chief – **Sergey Subbotin** – Dr. Sc., Professor, Head of Software Tools Department, National University “Zaporizhzhia Polytechnic”, Ukraine.

Deputy Editor-in-Chief – **Dmytro Piza** – Dr. Sc., Professor, Director of the Institute of Informatics and Radio Electronics, Professor of the Department of Radio Engineering and Telecommunications, National University “Zaporizhzhia Polytechnic”, Ukraine.

Members of the Editorial Board:

Iosif Androulidakis – PhD, Head of Telephony Department, Network Operation Center, University of Ioannina, Greece;

Evgeniy Bodyanskiy – Dr. Sc., Professor, Professor of the Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Ukraine;

Oleg Drobakhin – Dr. Sc., Professor, First Vice-Rector, Oles Honchar Dnipro National University, Ukraine;

Alexander Gorban – PhD, Professor, Professor of the Faculty of Mathematics, University of Leicester, United Kingdom;

Dmitry Gorodnichy – PhD, Associate Professor, Leading Research Fellow at the Directorate of Science and Engineering, Canada Border Services Agency, Ottawa, Canada;

Karsten Henke – Dr. Sc., Professor, Research Fellow, Faculty of Informatics and Automation, Technical University of Ilmenau, Germany;

Michitaka Kameyama – Dr. Sc., Professor, Professor of the Faculty of Science and Engineering, Ishinomaki Senshu University, Japan;

Volodymyr Kartashov – Dr. Sc., Professor, Head of the Department of Media Engineering and Information Radio Electronic Systems, Kharkiv National University of Radio Electronics, Ukraine;

Vitaly Levashenko – PhD, Professor, Head of Department of Informatics, University of Žilina, Slovakia;

David Luengo – PhD, Professor, Head of the Department of Signal Theory and Communication, Madrid Polytechnic University, Spain;

Ursula Markowska-Kaczmar – Dr. Sc., Professor, Professor of the Department of Computational Intelligence, Wrocław University of Technology, Poland;

Andrii Oliinyk – PhD, Associate Professor, Associate Professor, Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Ukraine;

Marcin Paprzycki – Dr. Sc., Professor, Professor of the Department of Intelligent Systems, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland;

Volodymyr Pavlikov – Dr. Sc., Senior Researcher, Vice-Rector for Research, N. E. Zhukovskiy National Aerospace University “KhAI”, Ukraine;

Alexei Sharpanskykh – PhD, Associate Professor, Associate Professor of Aerospace Engineering Faculty, Delft University of Technology, Netherlands;

Stepan Skrupsky – PhD, Associate Professor, Associate Professor of the Department of Computer Systems and Networks, National University “Zaporizhzhia Polytechnic”, Ukraine;

Galyna Tabunshchyk – PhD, Professor, Professor of the Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Ukraine;

Thomas (Tom) Trigano – PhD, Senior Lecturer of the Department of Electrical and Electronic Engineering, Sami Shamoon College of Engineering, Ashdod, Israel;

Joost Vennekens – PhD, Associate Professor, Associate Professor, Faculty of Engineering (Campus de Nair), Katholieke Universiteit Leuven, Belgium;

Carsten Wolff – PhD, Professor, Professor of the Department of Technical Informatics, Dortmund University of Applied Sciences and Arts, Germany;

Heinz-Dietrich Wuttke – PhD, Associate Professor, Leading Researcher at the Institute of Technical Informatics, Technical University of Ilmenau, Germany;

Elena Zaitseva – PhD, Professor, Professor, Department of Informatics, University of Žilina, Slovakia.

EDITORIAL-ADVISORY COUNCIL

Peter Arras – PhD, Associate Professor, Associate Professor, Faculty of Engineering (Campus De Nair), Katholieke Universiteit Leuven, Belgium;

Anatoly Lisnianski – PhD, Chief Scientific Expert, Israel Electric Corporation Ltd., Haifa, Israel;

Christian Madritsch – PhD, Professor of the Faculty of Engineering and Information Technology, Carinthia University of Applied Sciences, Austria;

Mher Markosyan – Dr. Sc., Professor, Director of the Yerevan Research Institute of Communications, Professor of the Department of Telecommunications, Russian-Armenian University, Yerevan, Armenia;

Oleg Rubel – PhD, Associate Professor, Faculty of Engineering, McMaster University, Hamilton, Canada;

Peter Schulz – Dr. Sc., Professor, Professor, Faculty of Engineering and Computer Science, Hamburg University of Applied Sciences (HAW Hamburg), Hamburg, Germany;

Avtandil Tavkhelidze – PhD, Professor, Professor of the School of Business, Technology and Education, Ilia State University, Tbilisi, Georgia;

Doru Ursuțiu – Dr. Sc., Professor, Professor, Department of Electronics and Computer Engineering, University of Transylvania at Brasov, Romania.

Recommended for publication by the Academic Council of NU «Zaporizhzhia Polytechnic», protocol № 11/21 dated 31.05.2021.

The journal is imposed by the editorial-publishing department of NU «Zaporizhzhia Polytechnic».

The journal web-site is <http://ric.zntu.edu.ua>.

The address of the editorial office: Editorial office of the journal «Radio Electronics, Computer Science, Control», National University «Zaporizhzhia Polytechnic», Zhukovskiy street, 64, Zaporizhzhia, 69063, Ukraine.

Tel.: +38-061-769-82-96 – the editorial-publishing department.

E-mail: rvv@zntu.edu.ua

Fax: +38-061-764-46-62

© National University «Zaporizhzhia Polytechnic», 2021

Научный журнал «Радиоэлектроника, информатика, управление» (сокращенное название – РИУ) издается национальным университетом «Запорожская политехника» (НУ «Запорожская политехника») с 1999 г. периодичностью четыре номера в год.

Зарегистрирован в Министерстве юстиции Украины 19.11.2019 г. (Свидетельство о государственной регистрации печатного средства массовой информации серия KB № 24220-14060 ПР).

ISSN 1607-3274 (печатный), ISSN 2313-688X (электронный).

Приказом Министерства образования и науки Украины № 409 от 17.03.2020 г. «Об утверждении решений Аттестационной коллегии Министерства относительно деятельности специализированных ученых советов от 06 марта 2020 года» журнал включен в перечень научных профессиональных изданий Украины в категории «А» (наивысший уровень), в которых могут публиковаться результаты диссертационных работ на соискание ученых степеней доктора наук и доктора философии (кандидата наук).

Журнал включен в польский Перечень научных журналов и рецензируемых материалов международных конференций с присвоенным числом баллов (приложение к объявлению Министра науки и высшего образования Республики Польша от 31 июля 2019 г.: № 16981).

В журнале бесплатно публикуются научные статьи на английском, русском и украинском языках.

Правила оформления статей представлены на сайте: <http://ric.zntu.edu.ua/information/authors>.

Журнал обеспечивает **бесплатный открытый он-лайн доступ** к полнотекстовым публикациям. Журнал разрешает авторам иметь авторские права и сохранять права на издание без ограничений. Журнал разрешает пользователям читать, загружать, копировать, распространять, печатать, искать или ссылаться на полные тексты своих статей. Журнал разрешает повторное использование его содержания в соответствии Creative Commons лицензией CC BY-SA..

Опубликованным статьям присваивается уникальный идентификатор цифрового объекта DOI.

Журнал входит в наукометрическую базу Web of Science.

Журнал реферирован и индексируется в ведущих международных и национальных реферативных журналах и наукометрических базах данных, а также размещается в цифровых архивах и библиотеках с бесплатным доступом on-line, полный перечень которых представлен на сайте: <http://ric.zntu.edu.ua/about/editorialPolicies#custom-0>.

Журнал распространяется по Каталогу периодических изданий Украины (подписной индекс – 22914).

Тематика журнала: телекоммуникации и радиоэлектроника, программная инженерия (включая теорию алгоритмов и программирования), компьютерные науки (математическое и компьютерное моделирование, оптимизация и исследование операций, управление в технических системах, межмашинное и человеко-машинное взаимодействие, искусственный интеллект, включая интеллектуальный анализ данных, распознавание образов, искусственные нейронные и нейро-нечеткие сети, нечеткую логику, коллективный интеллект и мультиагентные системы, гибридные системы), компьютерная инженерия (аппаратное обеспечение вычислительной техники, компьютерные сети), информационные системы и технологии (структуры и базы данных, системы, основанные на знаниях, и экспертные системы, обработка данных и сигналов).

Все статьи, предлагаемые к публикации, получают **объективное рассмотрение**, которое оценивается по существу без учета расы, пола, вероисповедания, этнического происхождения, гражданства или политической философии автора(ов).

Все статьи проходят двухступенчатое закрытое (анонимное для автора) **рецензирование** штатными редакторами и независимыми рецензентами – ведущими учеными по профилю журнала.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Главный редактор – Субботин Сергей Александрович – доктор технических наук, профессор, заведующий кафедрой программных средств, Национальный университет «Запорожская политехника», Украина.

Заместитель главного редактора – Пиза Дмитрий Макарович – доктор технических наук, профессор, директор института информатики и радиоэлектроники, профессор кафедры радиотехники и телекоммуникаций, Национальный университет «Запорожская политехника», Украина.

Члены редколлегии:

Андрюлидакис Иосиф – доктор философии, глава департамента телефонии Центра обслуживания сетей, Университет Янины, Греция;

Бодянский Евгений Владимирович – доктор технических наук, профессор, профессор кафедры искусственного интеллекта, Харьковский национальный университет радиоэлектроники, Украина;

Венискенс Юст – доктор философии, доцент факультета инженерных технологий (кампус де Наир), Католический университет Левена, Бельгия;

Вольф Карстен – доктор философии, профессор, профессор кафедры технической информатики, Дортмундский университет прикладных наук и искусств, Германия;

Вуттке Ганс-Дитрих – доктор философии, доцент, ведущий научный сотрудник института технической информатики, Технический университет Ильменау, Германия;

Горбань Александр Николаевич – доктор физико-математических наук, профессор, профессор факультета математики, Университет Лестера, Великобритания;

Городничий Дмитрий Олегович – доктор философии, кандидат технических наук, доцент, ведущий научный сотрудник Дирекции науки и инженерии, Канадское агентство пограничной службы, Канада;

Дробахин Олег Олегович – доктор физико-математических наук, профессор, первый проректор, Днепровский национальный университет имени Олеся Гончара, Украина;

Зайцева Елена Николаевна – кандидат физико-математических наук, профессор, профессор кафедры информатики, Жилинский университет в Жилине, Словакия;

Камеяма Мичитака – доктор наук, профессор, профессор факультета науки и инженерии, Университет Ишиномаки Сеншу, Япония;

Карташов Владимир Михайлович – доктор технических наук, профессор, заведующий кафедрой медиаинженерии и информационных радиоэлектронных систем, Харьковский национальный университет радиоэлектроники, Украина;

Левашенко Виталий Григорьевич – кандидат физико-математических наук, профессор, заведующий кафедрой информатики, Жилинский университет в Жилине, Словакия;

Луэнго Давид – доктор философии, профессор, заведующий кафедрой теории сигналов и коммуникации, Мадридский политехнический университет, Испания;

Марковска-Качмар Урсла – доктор технических наук, профессор, профессор кафедры вычислительного интеллекта, Вроцлавская политехника, Польша;

Олейник Андрей Александрович – кандидат технических наук, доцент, доцент кафедры программных средств, Национальный университет «Запорожская политехника», Украина;

Павликов Владимир Владимирович – доктор технических наук, старший научный сотрудник, проректор по научной работе, Национальный аэрокосмический университет им. Н.Э. Жуковского «ХАИ», Украина;

Панищкий Марцин – доктор наук, профессор, профессор отдела интеллектуальных систем, Исследовательский институт систем Польской академии наук, г. Варшава, Польша;

Скрупский Степан Юрьевич – кандидат технических наук, доцент, доцент кафедры компьютерных систем и сетей, Национальный университет «Запорожская политехника», Украина;

Табунщик Галина Владимировна – кандидат технических наук, профессор, профессор кафедры программных средств, Национальный университет «Запорожская политехника», Украина;

Тригано Томас – доктор философии, старший преподаватель кафедры электрической и электронной инженерии, Инженерный колледж им. С. Шамон, г. Ашдод, Израиль;

Хенке Карстен – доктор технических наук, профессор, научный сотрудник факультета информатики и автоматизации Технический университет Ильменау, Германия;

Шарпаньских Алексей Альбертович – доктор философии, доцент факультета аэрокосмической инженерии, Делфтский технический университет, Нидерланды.

РЕДАКЦИОННО-КОНСУЛЬТАТИВНЫЙ СОВЕТ

Аррас Питер – доктор философии, доцент факультета инженерных технологий (кампус де Наир), Католический университет Левена, Бельгия;

Лисьянский Анатолий – кандидат физико-математических наук, главный научный эксперт, Израильская электрическая корпорация, Хайфа, Израиль;

Мадритц Христиан – доктор философии, профессор факультета инженерии и информационных технологий, Университет прикладных наук Каринфии, Австрия;

Маркосан Мгер Вардкесович – доктор технических наук, профессор, директор Ереванского научно-исследовательского института средств связи, профессор кафедры телекоммуникаций, Российско-армянский университет, Ереван, Армения;

Рубель Олег Владимирович – кандидат технических наук, доцент факультета инженерии, Университет МакМастер, Гамильтон, Канада;

Тавхелидзе Автандил – кандидат физико-математических наук, профессор, профессор школы бизнеса, технологии и образования, Государственный университет им. Ильи Чавчавадзе, Тбилиси, Грузия;

Урсутью Дору – доктор физико-математических наук, профессор, профессор кафедры электроники и вычислительной техники, Трансильванский университет в Брашове, Румыния;

Шульц Питер – доктор технических наук, профессор, профессор факультета инженерии и компьютерных наук, Гамбургский университет прикладных наук (HAW Hamburg), Гамбург, Германия.

Рекомендовано к изданию Ученым советом НУ «Запорожская политехника», протокол № 11/21 от 31.05.2021.

Журнал сверстан редакционно-издательским отделом НУ «Запорожская политехника».

Веб-сайт журнала: <http://ric.zntu.edu.ua>.

Адрес редакции: Редакция журнала «РИУ», Национальный университет «Запорожская политехника», ул. Жуковского, 64, г. Запорожье, 69063, Украина.

Тел.: +38-061-769-82-96 – редакционно-издательский отдел

Факс: +38-061-764-46-62

E-mail: rvv@zntu.edu.ua

© Национальный университет «Запорожская политехника», 2021

ЗМІСТ

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ.....	7
<i>Романов О. М., Котюбін В. Ю.</i> АЛГОРИТМИ ПОШУКУ ПЕРІОДИЧНОСТЕЙ У ЦИФРОВИХ ПОСЛІДОВНОСТЯХ З БЛОКОВИМ КОДУВАННЯМ ЗА ЇХ КОРЕЛЯЦІЙНИМИ ВЛАСТИВОСТЯМИ.....	7
<i>Solomentsev O. V., Zaliskyi M. Yu., Shcherbyna O. A., Asanov M. M.</i> OPTIMIZATION OF PREVENTIVE THRESHOLD FOR CONDITION-BASED MAINTENANCE OF RADIO ELECTRONIC EQUIPMENT.....	19
<i>Туровський О. Л., Дробик О. В.</i> ПОРЯДОК ОЦІНКИ НЕСУЧОЇ ЧАСТОТИ СИГНАЛУ, ЩО ПРИЙМАЄТЬСЯ СУПУТНИКОВОЮ СИСТЕМОЮ ЗВ'ЯЗКУ В БЕЗПЕРЕРВНОМУ РЕЖИМІ.....	28
МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ.....	39
<i>Gorev V. N., Gusev A. Yu., Korniienko V. I., Safarov A. A.</i> ON THE KOLMOGOROV-WIENER FILTER FOR RANDOM PROCESSES WITH A POWER-LAW STRUCTURE FUNCTION BASED ON THE WALSH FUNCTIONS.....	39
<i>Kodnyanko V. A., Grigorieva O. A., Strok L. V.</i> COMBINED NEWTON'S THIRD-ORDER CONVERGENCE METHOD FOR MINIMIZE ONE VARIABLE FUNCTIONS.....	48
<i>Tarasov V. N., Bakhareva N. F.</i> DELAY MODELS BASED ON SYSTEMS WITH USUAL AND SHIFTED HYPEREXPONENTIAL AND HYPERERLANGIAN INPUT DISTRIBUTIONS.....	56
НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ.....	65
<i>Бодяньський С. В., Шафроненко А. Ю., Клімова І. М.</i> ОНЛАЙН МЕТОД МОЖЛИВІСНОЇ КЛАСТЕРИЗАЦІЇ ДАНИХ НА ОСНОВІ ЕВОЛЮЦІЙНОЇ ОПТИМІЗАЦІЇ КОТЯЧИХ ЗГРАЙ.....	65
<i>Gulayeva N. M., Yaremko S. A.</i> EXPERIMENTAL ANALYSIS OF MULTINATIONAL GENETIC ALGORITHM AND ITS MODIFICATIONS.....	71
<i>Gurskiy A. A., Denisenko A. V., Dubna S. M.</i> THE AUTOMATIC SYNTHESIS OF PETRI NETS BASED ON THE FUNCTIONING OF ARTIFICIAL NEURAL NETWORK.....	84
<i>Леоценко С. Д., Олійник А. О., Субботін С. О., Гофман Є. О., Льяшенко М. Б.</i> СИНТЕЗ ТА ВИКОРИСТАННЯ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ З ЙМОВІРНІСНИМ КОДУВАННЯМ СТРУКТУРИ.....	93
<i>Litvin A. A., Velychko V. Yu., Kavarynskyi V. V.</i> TREE-BASED SEMANTIC ANALYSIS METHOD FOR NATURAL LANGUAGE PHRASE TO FORMAL QUERY CONVERSION.....	105
<i>Пелещак Р. М., Литвин В. В., Черняк О. І., Пелещак І. Р., Дорошенко М. В.</i> СТОХАСТИЧНА ПСЕВДОСПІНОВА НЕЙРОННА МЕРЕЖА З ТРИДІАГОНАЛЬНИМИ СИНАПТИЧНИМИ ЗВ'ЯЗКАМИ.....	114
ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ.....	123
<i>Hnatushenko V. V., Kashtan V. Yu.</i> AUTOMATED PANSHARPENING INFORMATION TECHNOLOGY OF SATELLITE IMAGES.....	123
<i>Polska O. V., Kudermetov R. K., Shkarupylo V. V.</i> AN APPROACH WEB SERVICE SELECTION BY QUALITY CRITERIA BASED ON SENSITIVITY ANALYSIS OF MCDM METHODS.....	133
<i>Riznyk V. V.</i> FORMALIZATION CODING METHODS OF INFORMATION UNDER TOROIDAL COORDINATE SYSTEMS.....	144
<i>Talakh M. V., Holub S. V., Turkin I. B.</i> INFORMATION TECHNOLOGY OF CLIMATE MONITORING.....	154
<i>Чопоров С. В., Ігнатченко М. С., Кудин А. В., Кривохата А. Г., Гоменюк С. І.</i> РЕАЛІЗАЦІЯ КОНЕЧНО-ЕЛЕМЕНТНОЇ БІБЛІОТЕКИ КЛАССОВ С ИСПОЛЬЗОВАНИЕМ ОБОБЩЕННОГО ПРОГРАММИРОВАНИЯ.....	164
УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ.....	174
<i>Avramenko S. E., Zheldak T. A., Koriashkina L. S.</i> GUIDED HYBRID GENETIC ALGORITHM FOR SOLVING GLOBAL OPTIMIZATION PROBLEMS.....	174
<i>Vakaliuk T. A., Pilkevych I. A., Tokar A. M., Loboda R. I.</i> CRITERIA FOR ESTIMATING THE SENSORIMOTOR REACTION TIME BY THE SMALL UAV OPERATOR.....	189

CONTENTS

RADIO ELECTRONICS AND TELECOMMUNICATIONS	7
<i>Romanov O. M., Kotiubin V. Yu.</i> PERIODICITY SEARCH ALGORITHMS IN DIGITAL SEQUENCES WITH BLOCK CODING BY THEIR CORRELATION PROPERTIES.....	7
<i>Solomentsev O. V., Zaliskyi M. Yu., Shcherbyna O. A., Asanov M. M.</i> OPTIMIZATION OF PREVENTIVE THRESHOLD FOR CONDITION-BASED MAINTENANCE OF RADIO ELECTRONIC EQUIPMENT.....	19
<i>Turovsky A. L., Drobik O. V.</i> PROCEDURE FOR EVALUATION OF THE SUPPORTING FREQUENCY SIGNAL OF THE SATELLITE COMMUNICATION SYSTEM IN CONTINUOUS MODE.....	28
MATHEMATICAL AND COMPUTER MODELING	39
<i>Gorev V. N., Gusev A. Yu., Korniienko V. I., Safarov A. A.</i> ON THE KOLMOGOROV-WIENER FILTER FOR RANDOM PROCESSES WITH A POWER-LAW STRUCTURE FUNCTION BASED ON THE WALSH FUNCTIONS.....	39
<i>Kodnyanko V. A., Grigorieva O. A., Strok L. V.</i> COMBINED NEWTON'S THIRD-ORDER CONVERGENCE METHOD FOR MINIMIZE ONE VARIABLE FUNCTIONS.....	48
<i>Tarasov V. N., Bakhareva N. F.</i> DELAY MODELS BASED ON SYSTEMS WITH USUAL AND SHIFTED HYPEREXPONENTIAL AND HYPERERLANGIAN INPUT DISTRIBUTIONS.....	56
NEUROINFORMATICS AND INTELLIGENT SYSTEMS	65
<i>Bodyanskiy Ye. V., Shafronenko A. Yu., Klymova I. N.</i> ONLINE PROBABILISTIC FUZZY CLUSTERING METHOD BASED ON EVOLUTIONARY OPTIMIZATION OF CAT SWARM.....	65
<i>Gulayeva N. M., Yaremko S. A.</i> EXPERIMENTAL ANALYSIS OF MULTINATIONAL GENETIC ALGORITHM AND ITS MODIFICATIONS.....	71
<i>Gurskiy A. A., Denisenko A. V., Dubna S. M.</i> THE AUTOMATIC SYNTHESIS OF PETRI NETS BASED ON THE FUNCTIONING OF ARTIFICIAL NEURAL NETWORK.....	84
<i>Leoshchenko S. D., Oliinyk A. O., Subbotin S. A., Gofman Ye. O., Ilyashenko M. B.</i> SYNTHESIS AND USAGE OF NEURAL NETWORK MODELS WITH PROBABILISTIC STRUCTURE CODING.....	93
<i>Litvin A. A., Velychko V. Yu., Kaverynskyi V. V.</i> TREE-BASED SEMANTIC ANALYSIS METHOD FOR NATURAL LANGUAGE PHRASE TO FORMAL QUERY CONVERSION.....	105
<i>Peleshchak R. M., Lytvyn V. V., Cherniak O. I., Peleshchak I. R., Doroshenko M. V.</i> STOCHASTIC PSEUDOSPIN NEURAL NETWORK WITH TRIDIAGONAL SYNAPTIC CONNECTIONS.....	114
PROGRESSIVE INFORMATION TECHNOLOGIES	123
<i>Hnatushenko V. V., Kashtan V. Yu.</i> AUTOMATED PANSHARPENING INFORMATION TECHNOLOGY OF SATELLITE IMAGES.....	123
<i>Polska O. V., Kudermetov R. K., Shkarupylo V. V.</i> AN APPROACH WEB SERVICE SELECTION BY QUALITY CRITERIA BASED ON SENSITIVITY ANALYSIS OF MCDM METHODS.....	133
<i>Riznyk V. V.</i> FORMALIZATION CODING METHODS OF INFORMATION UNDER TOROIDAL COORDINATE SYSTEMS.....	144
<i>Talakh M. V., Holub S. V., Turkin I. B.</i> INFORMATION TECHNOLOGY OF CLIMATE MONITORING.....	154
<i>Choporov S. V., Ihnatchenko M. S., Kudin O. V., Kryvokhata A. G., Homeniuk S. I.</i> IMPLEMENTATION OF A FINITE ELEMENT CLASS LIBRARY USING GENERALIZED PROGRAMMING.....	164
CONTROL IN TECHNICAL SYSTEMS	174
<i>Avramenko S. E., Zheldak T. A., Koriashkina L. S.</i> GUIDED HYBRID GENETIC ALGORITHM FOR SOLVING GLOBAL OPTIMIZATION PROBLEMS.....	174
<i>Vakaliuk T. A., Pilkevych I. A., Tokar A. M., Loboda R. I.</i> CRITERIA FOR ESTIMATING THE SENSORIMOTOR REACTION TIME BY THE SMALL UAV OPERATOR.....	189

РАДИОЕЛЕКТРОНИКА ТА ТЕЛЕКОМУНІКАЦІЇ

RADIO ELECTRONICS AND TELECOMMUNICATIONS

РАДИОЭЛЕКТРОНИКА И ТЕЛЕКОММУНИКАЦИИ

УДК 621.391

АЛГОРИТМИ ПОШУКУ ПЕРІОДИЧНОСТЕЙ У ЦИФРОВИХ ПОСЛІДОВНОСТЯХ З БЛОКОВИМ КОДУВАННЯМ ЗА ЇХ КОРЕЛЯЦІЙНИМИ ВЛАСТИВОСТЯМИ

Романов О. М. – канд. техн. наук, заступник начальника з наукової роботи, науково-дослідний інститут Міністерства оборони України, Київ, Україна.

Котюбін В. Ю. – науковий співробітник, науково-дослідний інститут Міністерства оборони України, Київ, Україна.

АНОТАЦІЯ

Актуальність Для підвищення завадостійкості систем зв'язку і передачі даних широко застосовується завадостійке кодування. Найбільш розповсюдженими через свою ефективність є методи блокового кодування. В умовах часткової апріорної невизначеності виду і параметрів кодування перед декодуванням цифрової послідовності проводиться попередній аналіз для їх визначення. При блоковому кодуванні для визначення періоду цифрової послідовності, обумовленого додаванням до неї синхропослідовності, і який може визначати вид і параметри кодування, розповсюдженим підходом є використання їх кореляційних властивостей.

Метою роботи є викладення алгоритмів пошуку періодичностей у цифрових послідовностях з блоковим завадостійким кодуванням в умовах часткової апріорної невизначеності виду і параметрів завадостійкого коду.

Метод. У статті наведено два алгоритми пошуку періодичностей у цифрових послідовностях з блоковим кодуванням та описано принцип їх роботи. В основі одного алгоритму лежить обчислення АКФ, в основі іншого – ВКФ. Показано, що довжина цифрової послідовності повинна бути щонайменше в два рази більше максимально можливого періоду. Роботу обох алгоритмів проілюстровано на прикладах.

Результати. На основі запропонованих алгоритмів розроблено СПЗ. Результати визначення періоду цифрових послідовностей з блоковим завадостійким кодуванням при різних значеннях періоду підтвердили працездатність запропонованих алгоритмів. Обидва запропоновані алгоритми дають приблизно однаковий результат. Встановлено експериментальні залежності часу обчислення АКФ і ВКФ від довжини цифрової послідовності і максимально можливого періоду. Алгоритм пошуку періоду цифрової послідовності за ВКФ її складових є більш ефективним через меншу кількість обчислень.

Висновки. У роботі вперше отримано два алгоритми пошуку періодичностей у цифрових послідовностях з блоковим завадостійким кодуванням, які засновані на визначенні їх кореляційних функцій. Застосування розроблених алгоритмів на практиці дає змогу в умовах часткової апріорної невизначеності виду і параметрів завадостійкого коду визначити період цифрових послідовностей у масштабі реального часу навіть при великих значеннях періоду, а на його основі – ідентифікувати вид і параметри блокових завадостійких кодів.

КЛЮЧОВІ СЛОВА: автокореляційна функція, взаємна кореляційна функція, цифрова послідовність, період, блоковий завадостійкий код, синхропослідовність, система зв'язку і передачі даних.

АБРЕВІАТУРИ

LDPC – low-density parity-check;
TPC – turbo product code;
АКФ – автокореляційна функція;
ВКФ – взаємна кореляційна функція;
ЕОМ – електронно-обчислювальна машина;
ЗК – завадостійкий код;
СЗПД – система зв'язку і передачі даних;

СП – синхропослідовність;
СПЗ – спеціальне програмне забезпечення;
ЦП – цифрова послідовність.

НОМЕНКЛАТУРА

e – кількість помилкових біт у СП ЦП $s_1(n)$ та $s_2(n)$;
 K – довжина СП у блоці;
 L – довжина прийнятої ЦП;

l – довжина штучно утворених ЦП;
 L_{\min} – мінімально необхідна довжина ЦП;
 M – довжина перевірочних даних у блоці;
 N – довжина інформації у блоці;
 q – кількість блоків;
 R – кількість одиниць у τ комірці масиву $R(\tau)$;
 $R(\tau)$ – АКФ ЦП;
 $R_{12}(l)$ – ВКФ ЦП;
 $s(n)$ – ЦП;
 $s_1(n)$ – перша ЦП, штучно утворена з $s(n)$;
 $s_2(n)$ – друга ЦП, штучно утворена з $s(n)$;
 $s_p(n)$ – результуюча ЦП;
 T – період ЦП;
 T_{\min} – мінімально можливий період ЦП;
 T_{\max} – максимально можливий період ЦП;
 $\Delta_{СП}$ – величина зсуву СП від початку ЦП;
 τ – затримка ЦП у часі;
 \equiv – оператор логічної операції «еквівалентність».

ВСТУП

У СЗПД, які набули широкого розповсюдження, використовуються різноманітні методи, протоколи, стандарти передачі даних. Для підвищення ймовірності правильного прийому дані піддаються попередньому ЗК [1]. Найбільш розповсюдженими через свою ефективність [2] видами ЗК у сучасних СЗПД є згорткове кодування, кодування Ріда-Соломона, складені коди, TRC та LDPC коди [3–5]. Більшість ЗК, що використовуються практично, є блоковими.

Для декодування прийнятих даних за існуючими методами, алгоритмами, стандартами повинні бути відомі вид і параметри коду [1]. В умовах часткової апріорної невизначеності виду і параметрів кодування необхідно попередньо провести аналіз ЦП для їх визначення [5]. Розвитку методичного апарату аналізу ЦП, який проводиться перед декодуванням, приділяється значно менша увага [6].

Для синхронізації декодера блокових кодів перед кодованим блоком даних вставляється СП, яка обумовлює періодичність ЦП [3, 7]. На цьому заснований принцип дії методів визначення виду і параметрів ЗК [5]. Недоліки, притаманні практичним реалізаціям останніх [8, 9], обмежують їх застосування. Отже, актуальним є дослідження алгоритмів визначення виду і параметрів блокових ЗК, зокрема, алгоритмів пошуку періодичностей.

Об'єктом дослідження є процес визначення періоду ЦП даних з блоковим ЗК.

Предметом дослідження є алгоритми пошуку періоду ЦП даних з блоковим ЗК.

Метою роботи є викладення алгоритмів пошуку періодичностей у ЦП з блоковим ЗК в умовах часткової апріорної невизначеності параметрів ЗК та дослідження їх ефективності.

1 ПОСТАНОВКА ЗАДАЧІ

Нехай перед передаванням ЦП $s(n)$ розбивається на q блоків однакової довжини N біт. До кожного

блока ЦП додаються перевірочні дані довжиною M біт [1]. Для синхронізації декодера з початком блоків перед кожним блоком вставляється СП довжиною K біт. Період T такої ЦП визначається: $T=N+M+K$ біт, а сама послідовність може бути представлена схематично, як показано на рис. 1.

У сформованій таким чином ЦП K біт СП постійні для кожного з q блоків. Решта $N+M$ біт є статистично незалежними.

Необхідно визначити період T прийнятої ЦП $s(n)$.

2 ОГЛЯД ЛІТЕРАТУРИ

Аналіз останніх досліджень і публікацій показує, що найбільш розповсюдженим методом виявлення періодичностей у сигналах є аналіз їх кореляційних властивостей. Найбільшого поширення кореляційна функція набула в системах радіолокації [10–12], а також знайшла своє застосування у додатках цифрової обробки сигналів [13, 14].

Аналізу ЦП даних методами кореляційної обробки присвячені роботи [8, 9]. До недоліків алгоритмів, отриманих авторами, слід віднести необхідність проведення великої кількості обчислень для розпізнавання блокових кодів довжиною більше 255 біт [8], а також проведення обчислень і розпізнавання блокових кодів тільки в масштабі нереального часу [9], що є неприйнятним для практичного застосування.

3 МАТЕРІАЛИ І МЕТОДИ

Для пошуку періоду ЦП розроблено два алгоритми. Розглянемо перший з них, який базується на основі обчислення АКФ ЦП.

Для отримання інформації про період T прийнятої детермінованої ЦП довжиною L біт як дійсної функції часу використаємо її АКФ [7, 8]:

$$R(\tau) = \frac{1}{L} \sum_{n=0}^{L-1} s(n)s(n+\tau). \quad (1)$$

Під операцією множення для бінарної ЦП будемо розуміти логічну операцію «еквівалентність», таблиця істинності для якої представлена в табл. 1.

Таблиця 1 – Таблиця істинності для операції «еквівалентність»

s_1	s_2	$s_1 \equiv s_2$
“1”	“1”	“1”
“1”	“0”	“0”
“0”	“1”	“0”
“0”	“0”	“1”

Операція дає істинний результат при однакових значеннях елементів, над якими виконується. Тоді (1) набуває виду:

$$R(\tau) = \frac{1}{L} \sum_{n=0}^{L-1} s(n) \equiv s(n+\tau). \quad (2)$$

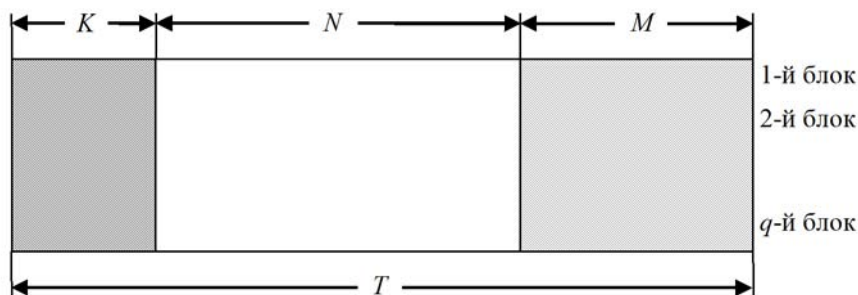


Рисунок 1 – Цифрова послідовність з періодом повторення T

Приклад обчислення АКФ двох фрагментів ЦП за формулою (2) представлений на рис. 2.

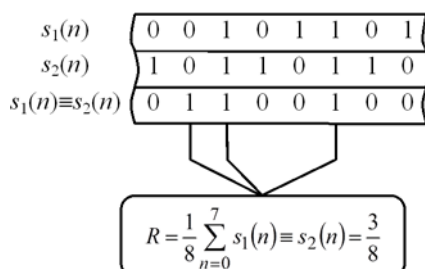


Рисунок 2 – Приклад обчислення АКФ

За відсутності зсуву ($\tau=0$) дві ЦП $s(n)$ будуть повністю ідентичними, що обумовить локальний максимум АКФ $R(0)$, значення якого дорівнює одиниці:

$$R(0) = \frac{1}{L} \sum_{n=0}^{L-1} s(n) \equiv s(n) = \frac{L}{L} = 1.$$

При $\tau \neq 0$ АКФ набуватиме максимуму при співпадіннях СП, що відповідає періоду T ЦП. Для його відшукування передбачимо можливість введення мінімально T_{\min} і максимально можливого періоду T_{\max} залежно від апріорно відомих даних. В більшості випадків доцільно покласти $T_{\min}=0$.

Визначимо мінімально необхідну довжину L_{\min} ЦП $s(n)$ для успішного відшукування періоду T .

Для цього спочатку розглянемо випадок синхронізації ЦП на початок СП (рис. 3). Як показано вище, за

відсутності зсуву ЦП (рис. 3а) спостерігатиметься локальний максимум АКФ.

При поступовому зсуванні ЦП ($\tau > 0$) через статистичну незалежність її елементів АКФ буде мати близьке до постійного значення, менше за максимальне (рис. 3б), і досягне локального максимуму при $\tau=T$ через ідентичність другої СП ЦП $s(n)$ і першої СП ЦП $s(n+\tau)$ (рис. 3в). Отже, в цьому випадку для відшукування періоду T ЦП достатньо її довжини $L_{\min}=T+K=2K+N+M$.

В умовах апріорної невизначеності виду і параметрів ЗК ЦП не синхронізована на початок СП (рис. 4), що обумовлює необхідність збільшення її мінімальної довжини на величину зсуву СП $\Delta_{СП}$ від початку: $L_{\min}=T+K+\Delta_{СП}=2K+N+M+\Delta_{СП}$.

Найбільш «жорстким» варіантом зсуву СП є її розрив із залишком на початку ЦП $K-1$ біт СП (рис. 5). У такому випадку для забезпечення співпадіння K біт СП ЦП, зсунутих між собою на $\tau=T$, їх довжина повинна бути не менше $L_{\min}=2T$.

Із порівняння отриманих значень мінімальної довжини ЦП для випадків синхронізації ЦП, наведених на рис. 3–5, видно, що в умовах апріорної невизначеності виду і параметрів ЗК мінімальна довжина ЦП повинна бути не менше подвоєного значення максимально можливого періоду ЦП $L_{\min}=2T_{\max}$.

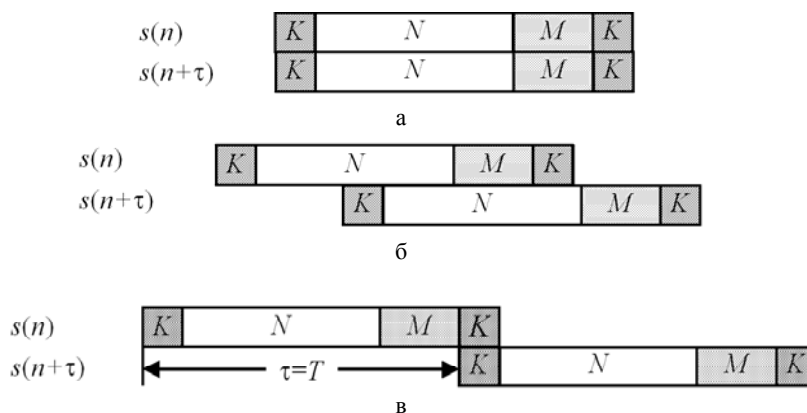


Рисунок 3 – Ілюстрація алгоритму пошуку за АКФ періоду ЦП, синхронізованої на початок СП:
 а – $\tau=0$; б – $0 < \tau < T$; в – $\tau=T$

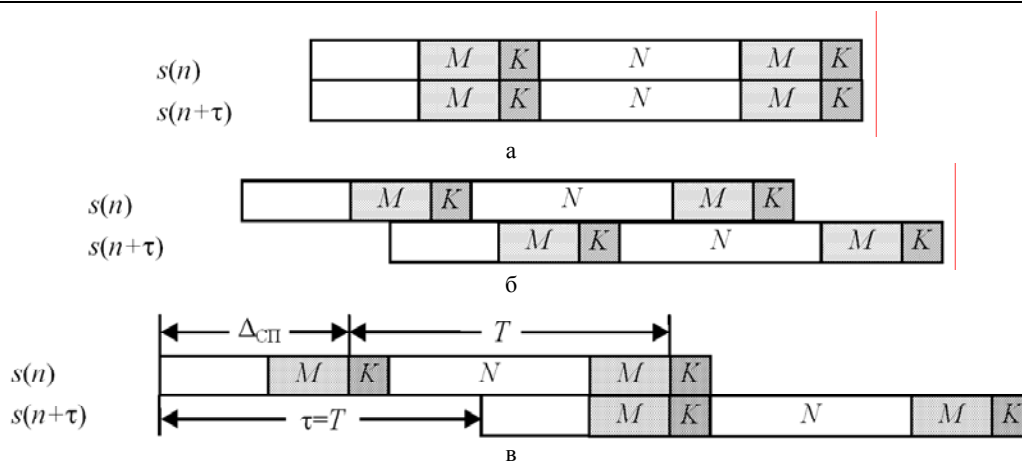


Рисунок 4 – Ілюстрація алгоритму пошуку за АКФ періоду ЦП, не синхронізованої на початок СП: а – $\tau=0$; б – $0<\tau<T$; в – $\tau=T$

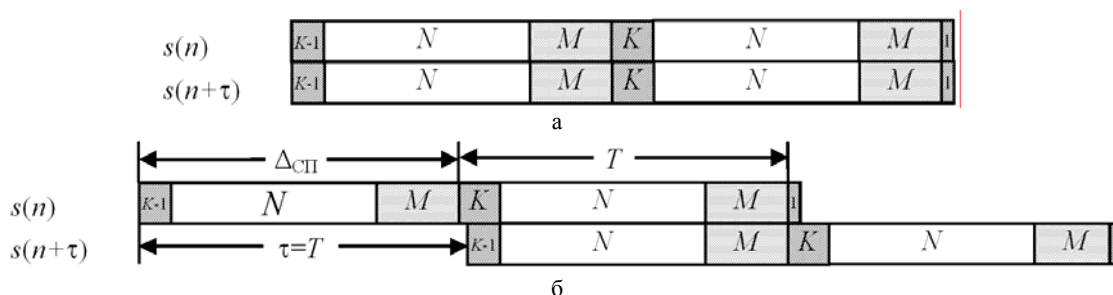


Рисунок 5 – Ілюстрація алгоритму пошуку за АКФ періоду ЦП із розривом СП: а – $\tau=0$; б – $\tau=T$

Суть алгоритму пошуку періоду полягає у зсуві другої ЦП відносно першої вправо з кроком один біт від $\tau=T_{\min}$ до $\tau=T_{\max}$ (рис. 3б, 4б). На кожному кроці зсуву над ними побітно виконується логічна операція «еквівалентність»: $s_p(n)=s(n)\equiv s(n+\tau)$. У τ комірці результуючого масиву $R(\tau)$ кількість одиниць R відповідає кількості біт, що співпали.

Очевидно, що при $\tau=0$ операція «еквівалентність» виконується над L біт ЦП. Із кожним кроком збільшення τ кількість пар елементів ЦП $s(n)$ і $s(n+\tau)$, над якими виконується операція «еквівалентність», зменшується на τ . Тому для корегування крайового ефекту, який при цьому виникає [13], модифікуємо формулу обчислення АКФ (2):

$$R(\tau) = \frac{1}{L-\tau} \sum_{n=\tau}^{L-1} s(n) \equiv s(n+\tau). \quad (3)$$

У формулі (3) операція «еквівалентність» і нормування здійснюються по кількості пар елементів ЦП, по яким реально обчислюється АКФ.

В отриманій описаним способом АКФ відшукуються максимальні значення $\max[R(\tau)]$. Найменше τ , $0<\tau\leq T_{\max}$, при якому функція $R(\tau)$ набуває максимуму, відповідає періоду T ЦП. Решта значень τ , при яких $R(\tau)$ набуває максимуму, відповідають кратним значенням періоду ЦП: $2T, 3T \dots$

Алгоритм пошуку періодичності ЦП за АКФ, що відображає описану послідовність дій, представлений на рис. 6. Результат пошуку періоду ЦП за наведеним алгоритмом представлений на рис. 9.

Для реалізації другого алгоритму пошуку періодичності у прийнятій ЦП $s(n)$ довжиною L за ВКФ штучно утворимо з неї дві ЦП $s_1(n)$ і $s_2(n)$, між якими шукатимемо кореляцію. Перша ЦП $s_1(n)$, довжиною $l=T_{\min}$, – від початку $s(n)$ до значення мінімально можливого періоду T_{\min} . Друга ЦП $s_2(n)$ – від значення $T_{\min}+1$ ЦП $s(n)$, також довжиною l (рис. 7): $s_1(n)=s(0 \dots T_{\min}-1)$; $s_2(n)=s(T_{\min} \dots 2T_{\min}-1)$.

Обчислимо значення ВКФ між утвореними ЦП [13]:

$$R_{12}(l) = \frac{1}{l} \sum_{n=0}^{l-1} s_1(n) \equiv s_2(n). \quad (4)$$

Для цього, як і в попередньому алгоритмі, побітно виконуватимемо логічну операцію «еквівалентність»: $s_p(n)=s_1(n)\equiv s_2(n)$. В результаті її виконання «одиниця» отримується за умови еквівалентності бітів. Отримана кількість одиниць підраховується із застосуванням оператора \sum і нормується до загальної кількості біт l у ЦП.

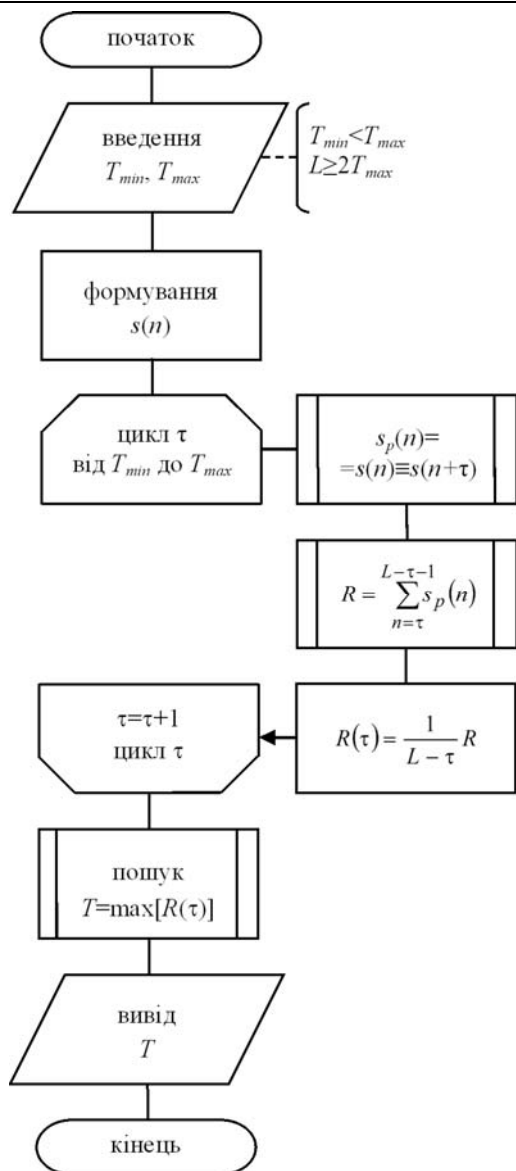


Рисунок 6 – Алгоритм пошуку періодичності ЦП за АКФ

Із кроком один біт збільшуватимемо l від значення T_{\min} до максимально можливого значення періоду T_{\max} , зсуваючи при цьому на відповідну кількість біт початок ЦП $s_2(n)$. На кожному кроці обчислюватимемо ВКФ ЦП. Через статистичну незалежність елементів ЦП значення ВКФ буде значно меншим порівняно із величинами локальних максимумів ВКФ (рис. 7б–ж). При досягненні довжини сформованих ЦП l значення істинного періоду T ЦП, що аналізується, зсув СП у $s_1(n)$ та $s_2(n)$ від їх початку буде однаковим (рис. 7з), що обумовить різке збільшення кількості біт, що співпадають між $s_1(n)$ та $s_2(n)$, і сплеск їх ВКФ. Подальше збільшення довжини сформованих ЦП призведе до циклічної появи сплесків ВКФ при співпадіннях другої та послідовуючих СП. Отже, істинне значення періоду ЦП T відповідає першому максимуму ВКФ і може бути підтверджене появою максимумів ВКФ на кратних періодах $2T, 3T \dots$. Максимальне значення періоду T_{\max} не може перевищувати цілої

частини від половини довжини прийнятої ЦП, що аналізується: $T_{\max} \leq L/2$. І навпаки: для відшукування періоду ЦП за ВКФ довжина ЦП L має бути такою, щоб в ній містилося не менше двох СП: $L \geq 2T_{\max}$.

Алгоритм пошуку періоду ЦП за ВКФ її складових, що відображає описану послідовність дій, представлений на рис. 8.

Результат пошуку періоду ЦП за наведеним алгоритмом представлений на рис. 10.

Принциповою відмінністю цього алгоритму від попереднього є пошук кореляційних властивостей різних ЦП, що слідують одна за одною у часі.

В обох розглянутих алгоритмах в загальному випадку крок зсуву може бути відмінним від одиниці.

Ймовірність появи помилкового біту в СП однієї з ЦП обумовлює зменшення обчисленої кореляційної функції при вірному періоді на e/l , де e – кількість помилкових біт у СП.

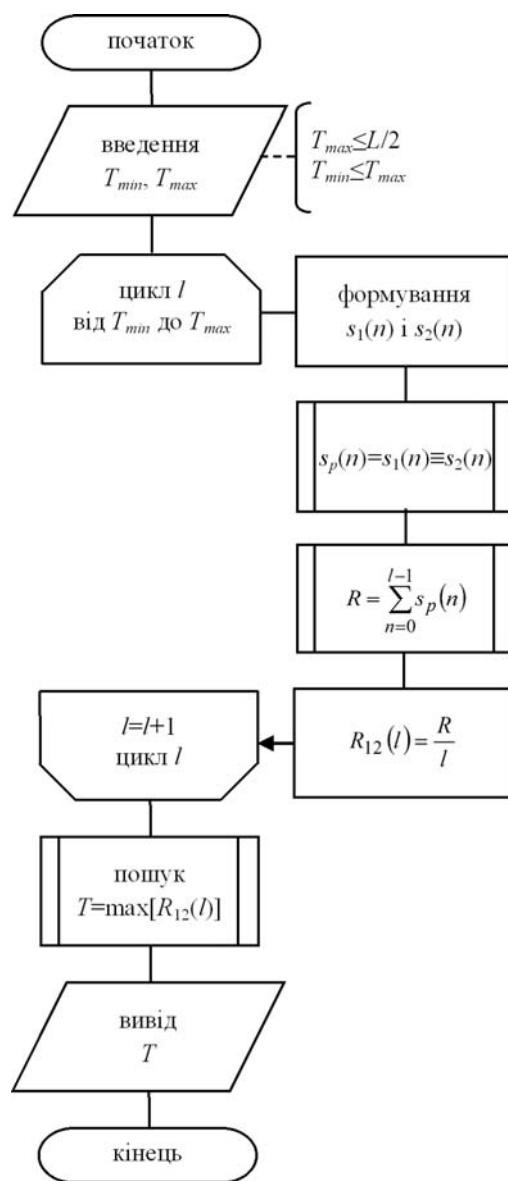


Рисунок 8 – Алгоритм пошуку періодичності ЦП за ВКФ її складових

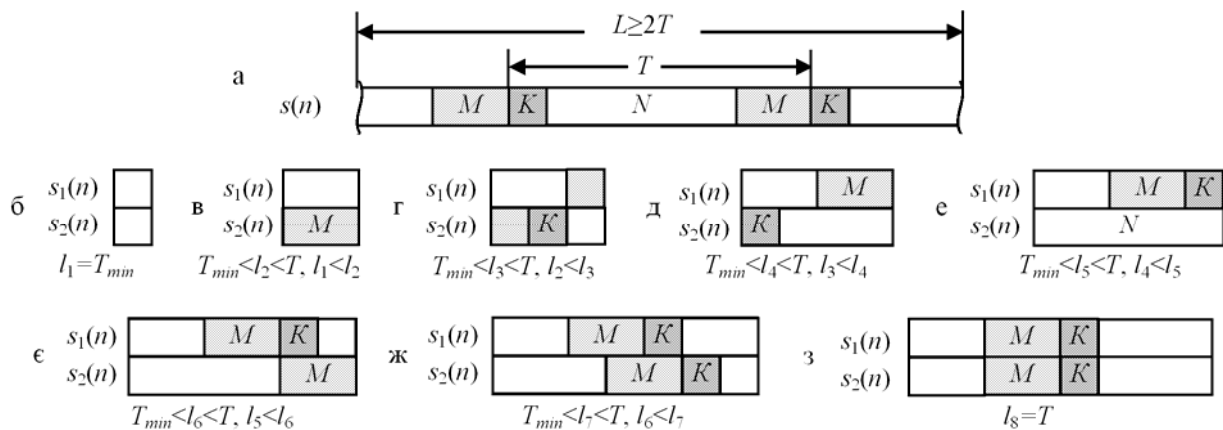


Рисунок 7 – Ілюстрація алгоритму пошуку періоду ЦП за ВКФ її складових:

а – вхідна ЦП; б-з – штучно утворені ЦП

4 ЕКСПЕРИМЕНТИ

Для визначення ефективності роботи запропонованих алгоритмів у Visual Studio 2019 мовою програмування C# розроблено СПЗ з підтримкою паралельних обчислень на багатоядерних центральних процесорах. Експерименти проводились на ЕОМ з 64-розрядною операційною системою Windows 10 і такими параметрами:

- двоядерний центральний процесор Intel Core 2 Duo E6700 з тактовою частотою 2,66 ГГц із 4 ГБ пам'яті;

- чотириядерний центральний процесор Intel Core i5-4460 з тактовою частотою 3,2 ГГц із 8 ГБ пам'яті;

- чотириядерний восьмипоточковий центральний процесор Intel Core i7-4790K з тактовою частотою 4 ГГц із 8 ГБ пам'яті.

У всіх експериментах задавалось: $T_{min}=0$, $\tau=1$ ($\Delta l=1$), $L = 2000, 4000, 8000, 16000, 32000, 48000, 64000, 80000, 96000, 112000, 128000$ біт. Значення T_{max} задавалось у межах від 1000 біт до $L/2$ (для ЦП з найбільшою довжиною $L=128000$ біт T_{max} становило 64000 біт), тобто від суттєво меншого за довжину ЦП до максимального.

Період ЦП визначався за обома описаними алгоритмами і вимірювався час його визначення без врахування введення і виведення даних та пошуку максимального значення у результуючому масиві $R(\tau)$ ($R(l)$).

При проведенні експериментів використовувалися реальні ЦП з телекомунікаційних мереж, ймовірність помилки у яких знаходилась у межах 10^{-5} – 10^{-6} . За такої ймовірності помилки кількість помилок у всій ЦП довжиною від 2000 до 128000 біт не перевищувала одну, що дає змогу вважати ЦП безпомилковими (ідеальними).

5 РЕЗУЛЬТАТИ

АКФ і ВКФ, обчислені за розробленими в роботі алгоритмами, для ЦП із періодом 1504 біт представлені на рис. 9, 10 відповідно.

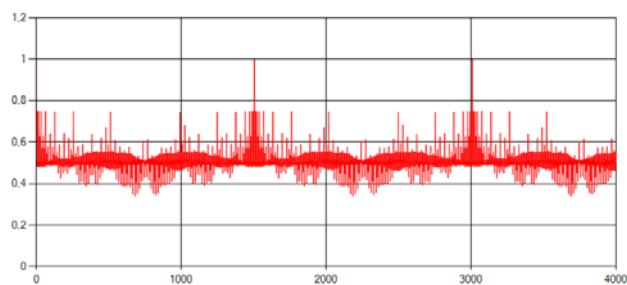


Рисунок 9 – Результати пошуку АКФ, $T=1504$ біти

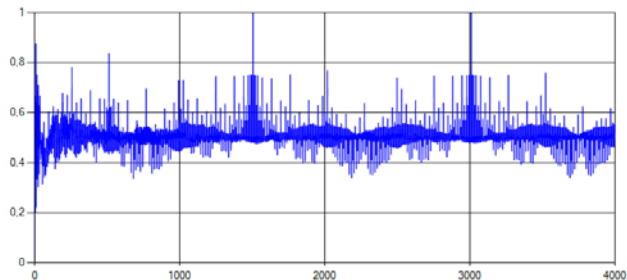


Рисунок 10 – Результати пошуку ВКФ, $T=1504$ біти

На рис. 11, 12 представлені результати обчислення АКФ ЦП довжиною $L=20000$ біт, істинний період якої становить $T=424$ біти. Рис. 11 відповідає пошуку АКФ при заданому максимальному для даної довжини ЦП $T_{max}=L/2=10000$ біт. Рис. 12 відповідає пошуку АКФ при $T_{max}=500$, який незначно перевищує T .

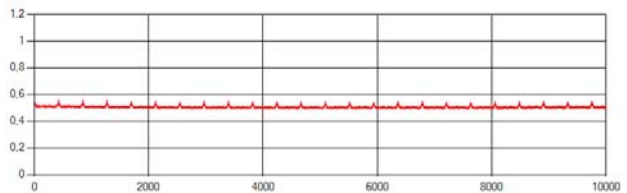


Рисунок 11 – Результати пошуку АКФ, $T_{max}=10000$ біт

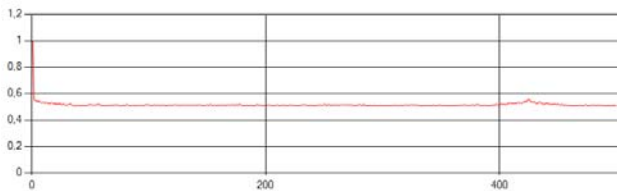


Рисунок 12 – Результати пошуку АКФ, $T_{\max}=500$ біт

Час обчислення за запропонованими алгоритмами АКФ, представлений у табл. 2, а ВКФ – у табл. 3, 4. Отримані результати також представлені у вигляді графіків. Графік залежності часу обчислення АКФ від довжини ЦП при різних значеннях T_{\max} представлений на рис. 13, а від T_{\max} при різних значеннях L – на рис. 14. Графік залежності часу обчислення ВКФ від максимально можливого періоду T_{\max} представлений на рис. 15.

Таблиця 2 – Час обчислення АКФ ЦП, с

T_{\max} , біт		L , біт										
		2000	4000	8000	16000	32000	48000	64000	80000	96000	112000	128000
1000	Core 2	0,002	0,004	0,008	0,016	0,032	0,045	0,062	0,076	0,091	0,106	0,123
	Core i5	0,001	0,003	0,007	0,014	0,027	0,034	0,044	0,047	0,055	0,058	0,062
	Core i7	0,000	0,000	0,001	0,003	0,006	0,010	0,013	0,017	0,020	0,023	0,027
2000	Core 2		0,007	0,014	0,029	0,060	0,089	0,121	0,150	0,179	0,211	0,242
	Core i5		0,005	0,013	0,026	0,038	0,053	0,066	0,075	0,078	0,088	0,095
	Core i7		0,001	0,003	0,006	0,013	0,020	0,027	0,035	0,039	0,046	0,053
4000	Core 2			0,025	0,055	0,116	0,174	0,238	0,294	0,354	0,412	0,480
	Core i5			0,021	0,038	0,059	0,074	0,094	0,107	0,122	0,138	0,152
	Core i7			0,005	0,012	0,025	0,038	0,052	0,064	0,077	0,091	0,102
8000	Core 2				0,093	0,214	0,332	0,457	0,571	0,691	0,811	0,941
	Core i5				0,059	0,088	0,115	0,148	0,178	0,209	0,237	0,268
	Core i7				0,020	0,047	0,072	0,096	0,123	0,148	0,175	0,199
16000	Core 2					0,369	0,602	0,850	1,080	1,322	1,560	1,820
	Core i5					0,127	0,185	0,248	0,304	0,367	0,428	0,488
	Core i7					0,078	0,130	0,182	0,223	0,282	0,335	0,384
24000	Core 2						0,812	1,172	1,532	1,890	2,253	2,610
	Core i5						0,238	0,330	0,412	0,504	0,597	0,688
	Core i7						0,174	0,250	0,331	0,404	0,485	0,560
32000	Core 2							1,454	1,921	2,403	2,882	3,387
	Core i5							0,396	0,517	0,632	0,754	0,882
	Core i7							0,309	0,410	0,510	0,615	0,716
40000	Core 2								2,252	2,849	3,448	4,048
	Core i5								0,599	0,744	0,898	1,052
	Core i7								0,481	0,612	0,738	0,875
48000	Core 2									3,237	3,961	4,680
	Core i5									0,850	1,031	1,216
	Core i7									0,687	0,851	0,994
56000	Core 2										4,414	5,253
	Core i5										1,144	1,352
	Core i7										0,944	1,125
64000	Core 2											5,865
	Core i5											1,493
	Core i7											1,233

Таблиця 3 – Час обчислення ВКФ ЦП (ЕОМ з процесором Intel Core 2 Duo), с

T_{\max} , біт	L , біт										
	2000	4000	8000	16000	32000	48000	64000	80000	96000	112000	128000
1000	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
2000		0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
4000			0,007	0,007	0,007	0,007	0,007	0,007	0,007	0,007	0,007
8000				0,026	0,026	0,026	0,026	0,026	0,026	0,026	0,026
16000					0,099	0,099	0,099	0,099	0,099	0,099	0,099
24000						0,222	0,222	0,222	0,222	0,222	0,222
32000							0,394	0,394	0,394	0,394	0,394
40000								0,613	0,613	0,613	0,613
48000									0,883	0,883	0,883
56000										1,195	1,195
64000											1,558

Таблиця 4 – Час обчислення ВКФ ЦП $L=128000$ біт

Час обчислення, с	T_{\max} , біт										
	1000	2000	4000	8000	16000	24000	32000	40000	48000	56000	64000
Core i5	0,001	0,002	0,008	0,028	0,061	0,102	0,158	0,224	0,309	0,402	0,522
Core i7	0,000	0,000	0,001	0,006	0,027	0,058	0,102	0,162	0,237	0,316	0,411

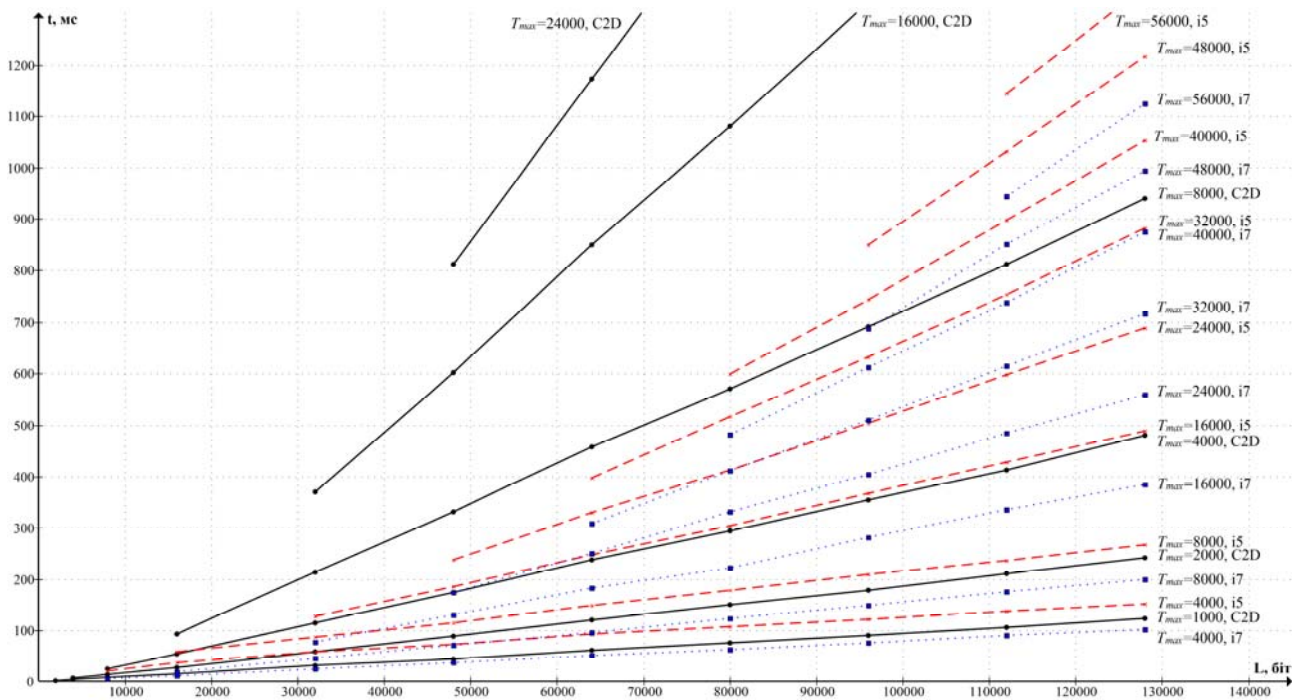


Рисунок 13 – Графік залежності часу обчислення АКФ від довжини ЦП L

6 ОБГОВОРЕННЯ

Для більшості ЦП, що досліджувались, за обома розробленими алгоритмами в результуючих кореляційних масивах $R(\tau)$ ($R(l)$) отримано максимальне значення, по якому вірно обчислено період ЦП. Винятком є ЦП з короткими СП, кореляція на довжині яких не забезпечує появу вираженого максимуму у результуючому масиві $R(\tau)$ ($R(l)$). Наприклад, в СЗПД, організованій за стандартом [15], довжина СП становить менше 1% від періоду, чого не достатньо для отримання вираженого максимуму в кореляційному масиві. Тому у подальшому доцільно дослідити вплив довжини СП у ЦП (відношення K/T) та її структури на величину максимумів кореляційних функцій. Також доцільно дослідити вплив шумів та завад на величину максимумів кореляційних функцій: із збільшенням ймовірності помилки у ЦП кількість біт СП, що співпали, зменшуватиметься, що, у свою чергу, призведе до зменшення максимумів АКФ і ВКФ.

На рис. 9, 10 видно максимуми кореляційних функцій, обумовлені співпадінням СП ЦП на періодах кореляції, кратних істинному. Також із порівняння рис. 9 і 10 видно, що результати обчислення АКФ і ВКФ однакові, за винятком початкової ділянки, на якій ВКФ обчислюється для дуже коротких ЦП, в той час як АКФ обчислюється для ЦП постійної довжини.

Із порівняння рис. 11, 12 видно, що алгоритм обчислення АКФ коректно працює як за умови $T < T_{\max} \ll L$, так і за умови $T \ll T_{\max} = L/2$. Другий випадок дозволяє отримати максимуми АКФ, які відпо-

відають кратним значенням періоду ЦП та дозволяють підтвердити визначений період.

Із порівняння табл. 2, 3 і 4 видно, що при однаковій довжині ЦП L і максимальному періоді T_{\max} час обчислення ВКФ в декілька разів (залежно від L і T_{\max}) менше часу обчислення АКФ, що обумовлено меншою кількістю операцій «еквівалентність», які виконуються у кожному циклі.

Із табл. 2 і рис. 13 видно, що час обчислення АКФ знаходиться в лінійній залежності від довжини ЦП. Коефіцієнт пропорційності визначається значенням T_{\max} , параметрами ЕОМ та ефективністю програмної реалізації алгоритмів.

Також з табл. 2 і рис. 14 видно, що залежність часу обчислень АКФ від максимального періоду ЦП T_{\max} може бути описана поліномом другого ступеню, коефіцієнти якого визначаються довжиною ЦП, параметрами ЕОМ та ефективністю програмної реалізації алгоритмів. Зменшення часу обчислення із збільшенням T_{\max} порівняно із лінійною залежністю обумовлене підвищенням ефективності паралельних обчислень із збільшенням кількості циклів.

Із табл. 3 і рис. 15 видно, що при $L \geq 2T_{\max}$ час обчислення ВКФ від довжини ЦП не залежить, оскільки у кожному циклі обчислення ВКФ використовується ЦП довжиною $2l$. Із табл. 3, 4 і рис. 15 видно, що час обчислення ВКФ знаходиться у поліноміальній залежності другого ступеню від T_{\max} . Коефіцієнти полінома визначаються параметрами ЕОМ та ефективністю програмної реалізації алгоритмів і, на відміну від АКФ, не залежать від довжини ЦП.

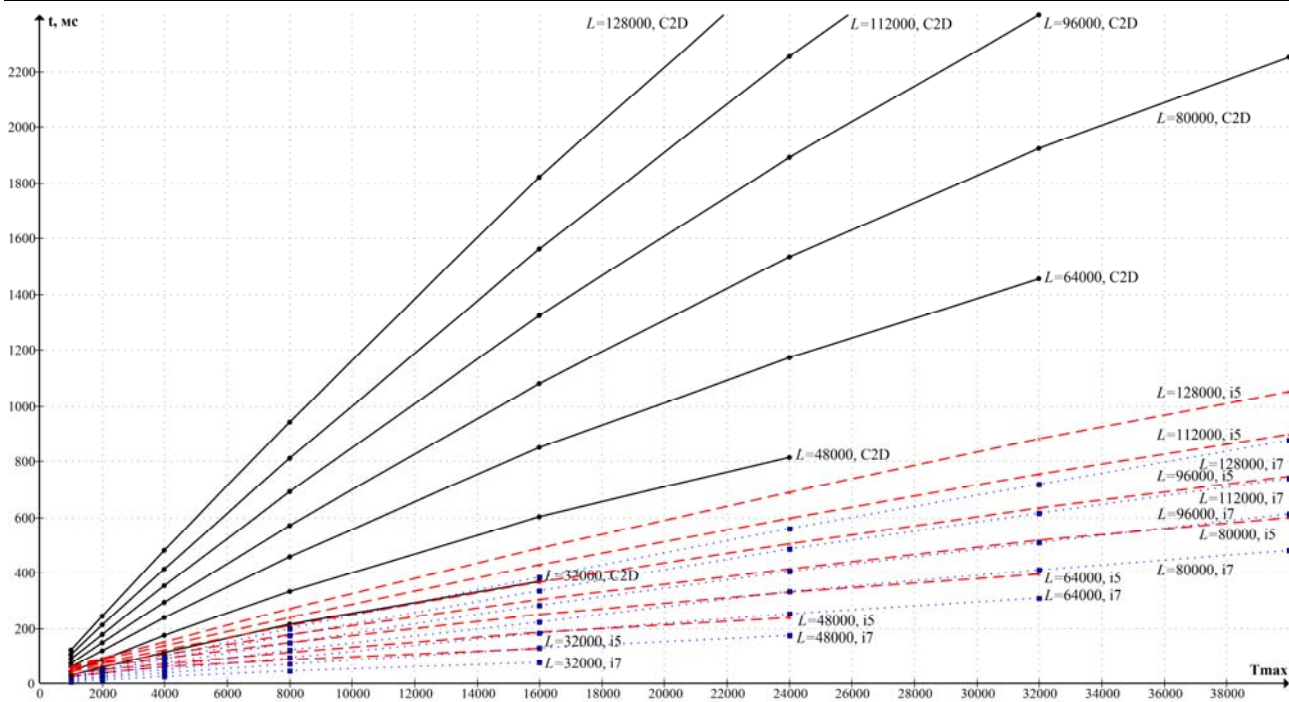


Рисунок 14 – Графік залежності часу обчислення АКФ від максимально можливого періоду T_{max}

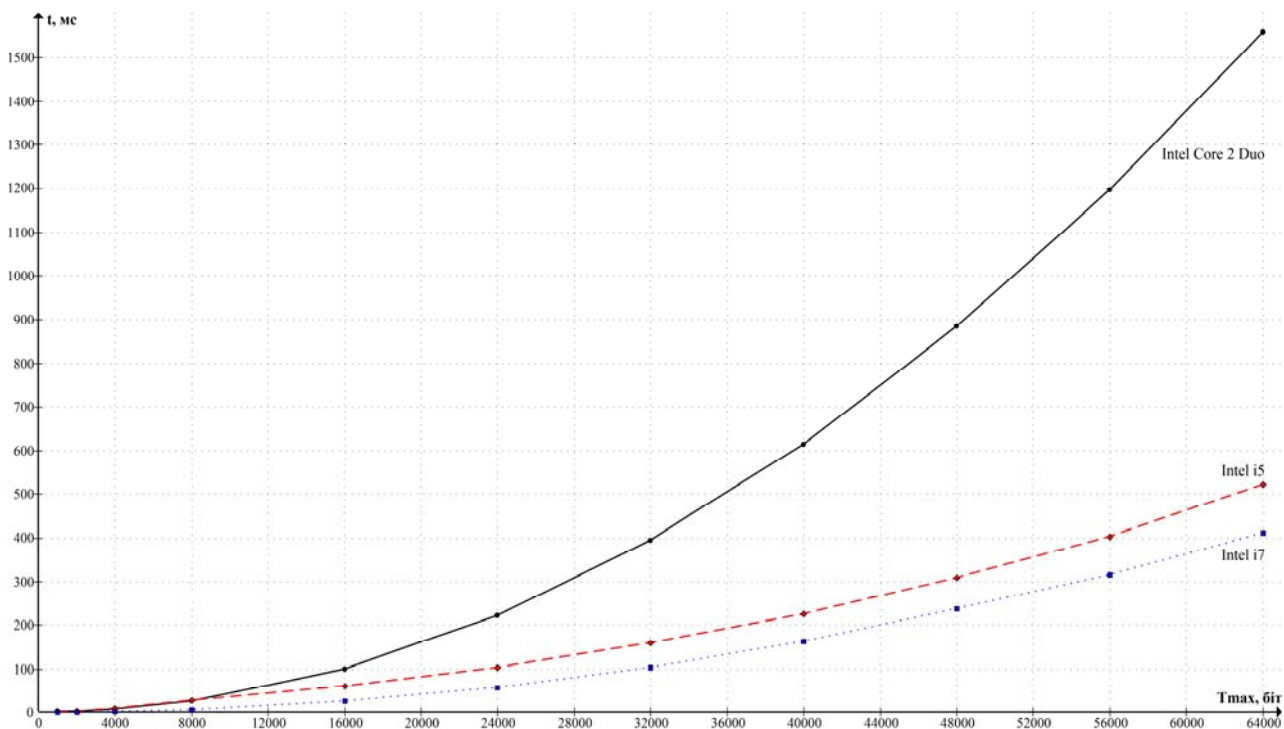


Рисунок 15 – Графік залежності часу обчислення ВКФ від максимально можливого періоду T_{max}

Час, витрачений на пошук АКФ і ВКФ, за обома алгоритмами від істинного значення періоду ЦП не залежить. Він також не залежить і від порядку заповнення ЦП, а визначається кількістю операцій у запропонованих алгоритмах для заданих L і T_{max} .

Той факт, що для обчислень кореляційних функцій достатньо ЦП довжиною лише $L_{min}=2T$, обумовлює порівняно невеликий обсяг обчислень. Період ЦП довжиною 100000 біт (T_{max} – не більше 50000 біт) за

запропонованими алгоритмами при програмній реалізації з підтримкою паралельних обчислень на багатоядерних центральних процесорах сучасних ЕОМ може бути обчислений у масштабі реального часу: при пошуку за АКФ менше 1 с, а за ВКФ – менше 0,5 с.

В умовах апріорної невизначеності періоду T довжина ЦП L і її максимально можливий період T_{max} можуть бути задані значно більше істинного періоду T .

ВИСНОВКИ

В роботі вперше отримано два алгоритми пошуку періодичностей у ЦП з блоковим ЗК в умовах часткової апіорної невизначеності виду і параметрів ЗК за АКФ і ВКФ ЦП та досліджено ефективність їх застосування.

Проведений експеримент підтвердив працездатність розроблених алгоритмів щодо визначення періоду ЦП. При застосуванні обох алгоритмів отримано приблизно однакові результати кореляційних характеристик ЦП.

На відміну від відомих раніше, обидва отриманих алгоритми дозволяють здійснювати пошук періоду ЦП у масштабі реального часу навіть при великих значеннях періоду ЦП. Час обчислення ВКФ в декілька разів менше часу обчислення АКФ, тому алгоритм пошуку періоду ЦП за ВКФ її складових є ефективнішим за алгоритм пошуку періоду ЦП за АКФ.

При проведенні експериментальних досліджень основна увага зосереджена на швидкодії алгоритмів. У подальшому доцільно дослідити вплив довжини і структури СП та ймовірності помилки у ЦП на величину максимумів кореляційних функцій.

До **недоліків** обох алгоритмів слід віднести слабку здатність пошуку періодичностей у ЦП з короткими СП. Робота над усуненням цього недоліку обумовлює напрямок **подальших наукових досліджень**.

Практичне значення отриманих результатів полягає у можливості визначення з допомогою розроблених алгоритмів у масштабі реального часу періоду ЦП для подальшого розпізнавання на його основі виду і параметрів ЗК у ЦП.

ЛІТЕРАТУРА / LITERATURA

1. Morelos-Zaragoza R. H. The Art of error correcting coding. 2nd ed. / Robert H. Morelos-Zaragoza. – Chichester : John Wiley & Sons, 2006. – 278 p.
2. Зубарев Ю. Б. Помехоустойчивое кодирование в цифровых системах передачи данных [Электрон. ресурс] / Ю. Б. Зубарев, Г. В. Овечкин // Электросвязь. – 2008. – № 12. – С. 58–61. – Режим доступа: http://mtdbest.ru/articles/obzor_dvoichnie_kodi_2.pdf.
3. TC Synchronization and Channel Coding. – Blue Book. Issue 3 [Electronic resource]. – Washington : CCSDS, 2017. – 50 p. – Access mode: <https://public.ccsds.org/Pubs/231x0b3.pdf>.
4. Сидоркина Ю. А. Анализ эффективности современных помехоустойчивых кодов [Электрон. ресурс] / Ю. А. Сидоркина, Б. И. Шахтарин, К. А. Балахонов // Вестник МГТУ им. Н. Э. Баумана. Сер. «Приборостроение». – 2014. – № 6. – С. 108–116. – Режим доступа: <https://cyberleninka.ru/article/n/analiz-effektivnosti-sovremennyh-pomeho-ustoychivyh-kodov>.
5. Котюбин В. Ю. Особенности визначення періодичності у інформаційній послідовності при проведенні технічного аналізу сигналів / В. Ю. Котюбин, О. М. Романов, Д. Ю. Бурлак // Теорія та практика створення, розвитку і застосування високотехнологічних систем спеціального призначення з урахуванням досвіду антитерористичної операції : XXII Всеукраїнська науково-практична конференція, Житомир, 26–27 квітня 2018 р. : тези доповідей. – Житомир : ЖВІ імені С. П. Корольова, 2018. – С. 153.
6. Романов О. М. Особенности разработки комплексов анализа цифровых последовательностей / О. М. Романов // Створення та модернізація озброєння і військової техніки в сучасних умовах : 17 науково-технічна конференція, Чернігів, 7–8 вересня 2017 р. : збірник тез доповідей. – Чернігів : ДНВЦ ЗС України, 2017. – С. 309–310.
7. Proximity-1 Space Link Protocol – Coding and Synchronization Sublayer. – Blue Book. Issue 1 [Electronic resource]. – Washington : CCSDS, 2003. – 24 p. – Access mode: <https://public.ccsds.org/Pubs/211x2b1s.pdf>.
8. Алгоритм распознавания типа и оценки параметров помехоустойчивых кодов / [С. Н. Кириллов, В. А. Ревуцкий, Д. С. Семин и др.] // Цифровая обработка сигналов и ее применение DSPA – 2011 : 13-я Международная конференция, Москва, 30 марта–1 апреля 2011 г. : доклады. – Москва : Российское научное общество радиотехники, электроники и связи им. А. С. Попова, 2011. – Выпуск XIII – 2. – С. 30–33.
9. Ревуцкий В. А. Устойчивые к мешающим факторам алгоритмы распознавания вида помехоустойчивых кодов в радиотехнических системах : автореф. дис. к-та техн. наук : 05.12.04 «Радиотехника, в том числе системы и устройства телевидения» / В. А. Ревуцкий. – Рязань : РГРТУ, 2013. – 19 с.
10. Теоретические основы радиолокации. Учебное пособие для вузов / Под ред. Ширмана Я. Д. – М. : «Советское радио», 1970. – 560 с.
11. Susek W. Noise Radar with Microwave Correlation Receiver / W. Susek, B. Stec, Cz. Rećko // Acta Physica Polonica A. – 2011. – Vol. 119, № 4. – P. 483–487.
12. Advanced Ultrawideband Radar : Signals, Targets, and Applications / ed. by James D. Taylor. – Boca Raton, FL : CRC Press, Taylor & Francis Group, 2016. – 476 p.
13. Ifeakor E. C. Digital Signal Processing : a practical approach. 2nd ed. / Emmanuel C. Ifeakor, Barrie W. Jervis. – Harlow, New York : Prentice Hall, 2002. – 933 p.
14. Sklar B. Digital communications. Fundamentals and applications. 2nd ed. / Bernard Sklar. – New Jersey : Prentice Hall, 2001. – 1104 p.
15. Digital Video Broadcasting (DVB). Framing structure, channel coding and modulation for 11/12 GHz satellite services : EN 300 421 V1.1.2 (1997-08). – Sophia Antipolis : European Telecommunications Standards Institute, 1997. – 24 p.

Статья поступила в редакцию 01.06.2020.
После доработки 01.04.2021.

УДК 621.391

АЛГОРИТМЫ ПОИСКА ПЕРИОДИЧНОСТЕЙ В ЦИФРОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЯХ С БЛОЧНЫМ КОДИРОВАНИЕМ ПО ИХ КОРРЕЛЯЦИОННЫМ СВОЙСТВАМ

Романов А. Н. – канд. техн. наук, заместитель начальника по научной работе, научно-исследовательский институт Министерства обороны Украины, Киев, Украина.

© Романов О. М., Котюбин В. Ю., 2021
DOI 10.15588/1607-3274-2021-2-1

Котюбин В. Ю. – научный сотрудник, научно-исследовательский институт Министерства обороны Украины, Киев, Украина.

АННОТАЦИЯ

Актуальность. Для повышения помехоустойчивости систем связи и передачи данных широко применяется помехоустойчивое кодирование. Наиболее распространенными в силу своей эффективности являются методы блочного кодирования. В условиях частичной априорной неопределенности вида и параметров кодирования перед декодированием цифровой последовательности проводится предварительный анализ для их определения. При блочном кодировании для определения периода цифровой последовательности, обусловленного добавлением к ней синхропоследовательности, и который может определять вид и параметры кодирования, распространенным подходом является использование их корреляционных свойств.

Целью работы является изложение алгоритмов поиска периодичностей в цифровых последовательностях с блочным помехоустойчивым кодированием в условиях частичной априорной неопределенности вида и параметров помехоустойчивого кода.

Метод. В статье приведены два алгоритма поиска периодичностей в цифровых последовательностях с блочным кодированием и описан принцип их работы. В основе одного алгоритма лежит вычисление АКФ, в основе другого – ВКФ. Показано, что длина цифровой последовательности должна быть в два раза больше максимально возможного периода. Работа обоих алгоритмов проиллюстрирована примерами.

Результаты. На основе предложенных алгоритмов разработано специальное программное обеспечение. Результаты определения периода цифровых последовательностей с блочным помехоустойчивым кодированием при различных значениях периода подтвердили работоспособность предложенных алгоритмов. Оба предложенных алгоритма дают примерно одинаковый результат. Установлены экспериментальные зависимости времени вычисления АКФ и ВКФ от длины цифровой последовательности и максимально возможного периода. Алгоритм поиска периода цифровой последовательности с помощью ВКФ ее составляющих является более эффективным из-за меньшего количества вычислений.

Выводы. В работе впервые получены два алгоритма поиска периодичностей в цифровых последовательностях с блочным помехоустойчивым кодированием, основанные на определении их корреляционных функций. Применение разработанных алгоритмов на практике позволяет в условиях частичной априорной неопределенности вида и параметров помехоустойчивого кода определять период цифровых последовательностей в масштабе реального времени даже при больших значениях периода, а на его основе – идентифицировать вид и параметры блочных помехоустойчивых кодов.

КЛЮЧЕВЫЕ СЛОВА: автокорреляционная функция, взаимная корреляционная функция, цифровая последовательность, период, блочный помехоустойчивый код, синхропоследовательность, система связи и передачи данных.

UDC 621.391

PERIODICITY SEARCH ALGORITHMS IN DIGITAL SEQUENCES WITH BLOCK CODING BY THEIR CORRELATION PROPERTIES

Romanov O. M. – PhD, Deputy Head for Research, Research Institute of the Ministry of Defense of Ukraine, Kiev, Ukraine.
Kotiubin V. Yu. – Researcher, Research Institute of the Ministry of Defense of Ukraine, Kiev, Ukraine.

ABSTRACT

Context. To improve the noise immunity of communication and data transmission systems, error-correcting coding is widely used. The most common because of their effectiveness are block coding methods. Under conditions of partial a priori uncertainty of the type and parameters of encoding, before decoding the digital sequence, a preliminary analysis is carried out to determine them. In block coding, to determine the period of a digital sequence caused by the addition of a sync sequence to it, and which can determine the type and parameters of coding, a common approach is to use their correlation properties.

Objective. The object of the research is the presentation of periodicity search algorithms in digital sequences with block error-correcting coding under conditions of partial a priori uncertainty of the type and parameters of the error-correcting code.

Method. The article presents two periodicity search algorithms in digital sequences with block coding and describes the principle of their operation. The basis of one algorithm is the calculation of the autocorrelation function, the basis of the other is calculation of the cross-correlation function. It is shown that the length of the digital sequence should be twice as long as the maximum possible period. The operation of both algorithms is illustrated by examples.

Results. Based on the proposed algorithms, special software has been developed. The results of determining the period of digital sequences with block error-correcting coding at different values of the period confirmed the efficiency of the proposed algorithms. Both proposed algorithms give approximately the same result. Experimental dependences of the calculation time of auto- and cross-correlation functions from the length of the digital sequence and the maximum possible period are established. The period search algorithm in a digital sequence, that use the cross-correlation function of its components, is more efficient due to fewer calculations.

Conclusions. For the first time, two periodicity search algorithms in digital sequences with block error-correcting based on the determination of their correlation functions are obtained. The application of the developed algorithms in practice allows, under partial a priori uncertainty of the type and parameters of the error-correcting code, to determine the period of digital sequences in real time even at large values of the period, and based on it, to identify the type and parameters of block error-correcting codes.

KEYWORDS: autocorrelation function, cross-correlation function, digital sequence, period, block error-correcting code, synchronization sequence, communication and data transmission system.

REFERENCES

1. Morelos-Zaragoza R. H. The Art of error correcting coding. 2nd ed. Chichester : John Wiley & Sons, 2006, 278 p.
2. Zubarev Ju. B., Ovechkin G. V. Pomehoustojchivoe kodirovanie v cifrovyyh sistemah peredachi dannyh [Electronic resource], *Jelektrosvjaz'*, 2008, No. 12, pp. 58–61. Access mode: http://mtdbest.ru/articles/obzor_dvoichnie_kodi_2.pdf.
3. TC Synchronization and Channel Coding. Blue Book. Issue 3 [Electronic resource]. Washington, CCSDS, 2017, 50 p. Access mode: <https://public.ccsds.org/Pubs/231x0b3.pdf>.
4. Sidorkina Ju. A., Shahtarin B. I., Balahonov K. A. Analiz jeffektivnosti sovremennyh pomehoustojchivyh kodov [Electronic resource], *Vestnik MGTU im. N. Je. Baumana. Ser. "Priborostroenie"*, 2014, No. 6, pp. 108–116. Access mode: <https://cyberleninka.ru/article/n/analiz-effektivnosti-sovremennyh-pomehoustojchivyh-kodov>.
5. Kotjubin V. Ju., Romanov O. M., Burlak D. Ju. Osoblyvosti vyznachennja periodychnosti u informacijnij poslidovnosti pry provedeni tehničnogo analizu sygnaliv. *Teorija ta praktyka stvorennja, rozvytku i zastosuvannja vysokotehnologichnyh system special'nogo pryznachennja z urahuvannjam dosvidu antyterrorystychnoi' operacii' : XXII Vseukrai'ns'ka naukovo-praktyčna konfere-ncija, Zhytomyr, 26–27 kvitnja 2018 r. : tezy dopovidej*. Zhytomyr ZhVI imeni S. P. Korol'ova, 2018, P. 153.
6. Romanov O. M. Osoblyvosti rozrobky kompleksiv analizu cyfrovyyh poslidovnostej, *Stvorennja ta modernizacija ozbrojennja i vijs'kovoï' tehniky v suchasnyh umovah : 17 naukovo-tehnična konferencija, Chernigiv, 7–8 veresnja 2017 r. : zbirnyk tez dopovidej*. Chernigiv, DNVC ZS Ukrai'ny, 2017, pp. 309–310.
7. Proximity-1 Space Link Protocol – Coding and Synchronization Sublayer. Blue Book. Issue 1 [Electronic resource]. Washington, CCSDS, 2003, 24 p. Access mode: <https://public.ccsds.org/Pubs/211x2b1s.pdf>.
8. Kirillov S. N., Revuckij V. A., Semin D. S. i dr. Algoritm raspoznavanija tipa i ocenki parametrov pomehoustojchivyh kodov, *Cifrovaja obrabotka signalov i ee primenenie DSPA – 2011 : 13-ja Mezhdunarodnaja konferencija, Moskva, 30 marta–1 aprlja 2011 g. : doklady*. Moscow : Rossijskoe nauchno-tehničeskoe obshhestvo radiotekhniki, jelektroniki i svjazi im. A. S. Popova, 2011, Vypusk XIII – 2, pp. 30–33.
9. Revuckij V. A. Ustojchivye k meshajushhim fakto-ram algoritmy raspoznavanija vida pomeho-ustojchivyh kodov v radiotekhnicheskijh sistemah : avtoref. dis. ... k-ta tehn. nauk : 05.12.04 "Radiotekhnika, v tom chisle sistemy i ustrojstva televidenija". Rjazan', RGRTU, 2013, 19 p.
10. Teoreticheskie osnovy radiolokacii. Uchebnoe posobie dlja vuzov / Pod red. Shirmana Ja. D. Moscow, "Sovetskoe radio", 1970, 560 p.
11. Susek W., Stec B., Rečko Cz. Noise Radar with Microwave Correlation Receiver, *Acta Physica Polonica A*, 2011, Vol. 119, No. 4, pp. 483–487.
12. Advanced Ultrawideband Radar : Signals, Targets, and Applications, ed. by James D. Taylor. Boca Raton, FL, CRC Press, Taylor & Francis Group, 2016, 476 p.
13. Ifeachor E. C., Jervis B. W. Digital Signal Processing : a practical approach. 2nd ed. Harlow. New York, Prentice Hall, 2002, 933 p.
14. Sklar B. Digital communications. Fundamentals and applications. 2nd ed. New Jersey, Prentice Hall, 2001, 1104 p.
15. Digital Video Broadcasting (DVB). Framing structure, channel coding and modulation for 11/12 GHz satellite services : EN 300 421 V1.1.2 (1997-08). Sophia Antipolis, European Telecommunications Standards Institute, 1997, 24 p.

OPTIMIZATION OF PREVENTIVE THRESHOLD FOR CONDITION-BASED MAINTENANCE OF RADIO ELECTRONIC EQUIPMENT

Solomentsev O. V. – Dr. Sc., Professor of the Department of Telecommunication and Radio Electronic Systems, National Aviation University, Kyiv, Ukraine.

Zaliskyi M. Yu. – PhD, Associate Professor of the Department of Telecommunication and Radio Electronic Systems, National Aviation University, Kyiv, Ukraine.

Shcherbyna O. A. – PhD, Associate Professor of the Department of Electronics, Robotics and Technology of Monitoring and Internet of Things, National Aviation University, Kyiv, Ukraine.

Asanov M. M. – PhD, Associate Professor, Simferopol, Ukraine.

ABSTRACT

Context. Operation costs throughout the life cycle of radio electronic equipment are very significant, which value far exceeds the initial cost of the equipment. Therefore, the up-to-date scientific and technical problem is to minimize operation costs. One of the ways to solve this problem is the introduction of statistical data processing technologies in the operation systems of radio electronic equipment.

Objective. The goal of the paper is to improve the efficiency of the condition-based maintenance with the determining parameters monitoring, which is widely used in civil aviation.

Method. The solution of this problem is based on finding the functional dependence of the efficiency indicator in the form of specific operation costs on the basic parameters of radio electronic equipment and its operation system. To determine this dependence, the probability-event model is used, as well as methods of probability theory and mathematical statistics, in particular methods of statistical classification of sample sets and functional transformations of random variables. To determine the optimal level of the preventive threshold by the criterion of minimizing operation costs, the method of statistical simulation of Monte-Carlo is used.

Results. Maintenance strategy with the determining parameters monitoring based on additional statistical data processing and technology of the optimal preventive threshold calculation are improved.

Conclusions. The obtained results can be used during the development and modernization of operation systems of radio electronic equipment in terms of application of statistical data processing procedures. A comparative analysis of the two maintenance strategies showed that the use of additional statistical data processing might reduce specific operation costs. The proposed technology for determining the optimal preventive threshold can be extended to use during the operation of complex technical systems, in particular for those whose technical condition is associated with the values of the determining parameters.

KEYWORDS: efficiency indicator, condition-based maintenance, statistical data processing, operation system, radioelectronic equipment.

ABBREVIATIONS

CBM is a condition-based maintenance;

OS is an operation system;

RCM is a reliability-centered maintenance;

REE is a radio electronic equipment.

NOMENCLATURE

a_0 is an initial determining parameter value in the $[P_-; P_+]$ interval;

a_1 is a tangent of the technical condition deterioration linear trend slope;

C_M and C_R are average costs of maintenance and repair;

C_Σ is a total operating cost;

C_P and C_{PM} are the costs of data processing and preventive maintenance, as a result of which failure does not occur;

$h(t)$ is a Heaviside function;

$m_1(P_F)$ and $m_1(P_S)$ are the mathematical expectations of the probability of radio electronic equipment failure and the failure-free operation;

N is a number of elementary sections of Δt duration during the observation interval;

P_+ and P_- are preventive thresholds;

P_S and P_F are statistical estimates of probabilities of failure-free operation and failure of radio electronic equipment;

O_+ and O_- are operational thresholds;

S_1 is a first strategy;

S_2 is a second strategy;

t_d is a moment the changepoint beginning;

T_Σ is an observation interval;

$X(t)$ is a monitoring parameter;

$x(t)$ is a fluctuation component due to the instability of the determining parameter;

$y(t)$ is a determining parameter;

$y_1(t)$ is a component that characterizes the trend in a time interval where there is no changepoint;

$y_2(t)$ is a component that characterizes the changepoint;

α is a probability of the first kind error;
 β is a probability of the second kind error;
 γ is a dimensionless coefficient;
 Δt is a time interval;
 σ is a standard deviation.

INTRODUCTION

Radioelectronic equipment in civil aviation plays a big role in ensuring the safety and regularity of aircraft flights. There are airborne and ground equipment.

The radioelectronic equipment operation efficiency is related to its reliability [1]. Maintaining the required reliability and efficiency level of equipment usage is the main task of the OS.

The main operation system elements are radioelectronic equipment, processes, personnel, documentation, expandable resources, equipment for measurement and monitoring, etc. The main operation processes are equipment intended use, maintenance, repair, operational data collection and processing, resource extension, flight inspections, storage and disposal [2].

The operation system can be considered as an object of design and improvement [3].

The operation system forms and implements control and corrective actions. Decision making is based on operational strategies, such as condition-based and reliability-based [4]. The condition-based maintenance (CBM) strategy includes control of the determining parameters or reliability level. Another strategy is RCM. CBM and RCM strategies include data processing procedures, including statistical ones [5]. So in the operation system it is advisable to use algorithms and procedures of statistical data processing in order to make correct and timely decisions during the radioelectronic equipment intended use [6].

The object of study is the process of statistical data processing in the operation system of radio electronic equipment.

The subject of study is the method of determining the optimal values of preventive threshold for condition-based maintenance of radio electronic equipment.

The purpose of the work is to improve the efficiency of the condition-based maintenance with the determining parameters monitoring, which is widely used in civil aviation.

1 PROBLEM STATEMENT

The generalized block diagram of the operation system is shown in Fig. 1.

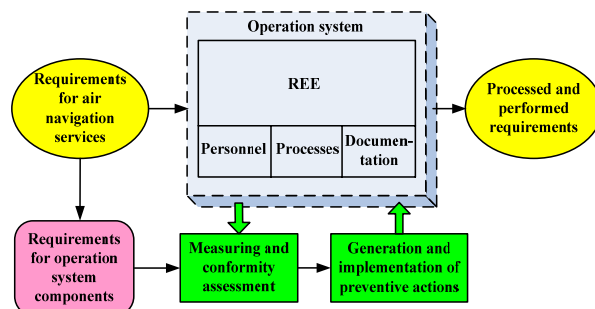


Figure 1 – Generalized block diagram of the OS

The circuit in Fig. 1 is built on the basis of the process approach (ISO 9001) and adaptability and management principles. The main OS element is REE. REEs use traffic control personnel during air traffic control [7]. REE signals in automatic mode are fed to the aircraft.

The task of the operation system elements and its control units is to provide the necessary levels of risk and safety [8]. At the same time, the OS operation costs should be minimal [9].

According to the National transport strategy of Ukraine [10], one of the main tasks is improvement of data collecting and processing. Data processing absence leads to additional costs and reduces efficiency [11].

In author's opinion, an important scientific area is the creation of a methodology for processing operational data and making decisions based on it.

This article discusses two tasks: the task of substantiating the need for statistical processing of operational data for one particular efficiency indicator type; the task of determining the preventive threshold during processing data for the condition-based maintenance strategy.

2 REVIEW OF THE LITERATURE

There are various concepts, approaches, regulatory documents in the aviation industry on the basis of which developers create equipment and its operation systems [12]. The modern approach is based on the entire equipment life cycle consideration: planning, design, manufacturing and operation [13].

Promising areas of research in the operation field are: remaining lifetime estimation, improvement of the procedures for diagnosing the technical condition, increasing the reliability of maintenance monitoring, cost optimization, statistical processing of operational data, processes analysis in deteriorating systems [14].

According to the second thermodynamics law, the entropy of any system tends to infinity. That's why, the integrity of the system is violated and the system breaks up into a large number of elements. It is the reason for the occurrence of degradation processes that are characteristic of physical, biological, and social systems [15]. Thus, in technical systems there are technical condition deterioration processes. The model of these processes is an unsteady, random process with two or

more quasistationary sections [16]. In the literature, such processes are called processes with changepoint.

The changepoint is time moment of transition from one stationary condition to another stationary or non-stationary condition. Mathematically the changepoint problem is associated with case when a series of observation contains values with one probability distribution before changepoint occurrence and values with another distribution after it.

Processing operational data of processes with changepoints will allow solving the problems of estimating the residual lifetime [17].

The literature analysis shows that sufficient attention is paid to the theoretical study of processes with changepoints. However, during the design and improvement there are insufficient engineering results in this area.

3 MATERIALS AND METHODS

The operation system is a system for managing the condition of REE and other OS elements [16]. The following content is needed for management: “when to do”, “what to do”, “whom to do”, “with what to do”. The main thing is the time to perform operational tasks, which is associated with the residual lifetime.

The following indicators are used as an OS efficiency indicator: availability factor, operational availability factor, operating efficiency factor, average risks [18].

In this article, we consider the efficiency indicator in the form of specific operation costs during the one REE operation. The REE technical condition will be characterized by one determining parameter.

In the operational management system for REE, we will consider two options for the maintenance and repair strategy.

The first strategy S_1 is based on the performing the periodic technical condition monitoring. The determining parameter value is compared with operational tolerances during the monitoring process. If tolerances are exceeded, failures occur and repairs are carried out. In S_1 , the trend analysis of the change in the determining parameter is not performed. Therefore, failures appear suddenly.

During the second S_2 strategy the statistical processing of operational data is made. This data is the trends of changes in the determining parameters. This strategy allows to evaluate the failure possibility and perform preventive and corrective actions. Repairs are performed in the event of a failure.

The analytical formula for the efficiency indicator can be obtained while using an event-probability model. The main components of this model are events, their probabilities and average costs.

Let us consider the S_1 strategy. We assume that the efficiency estimation is made on the observation interval T_Σ . The observation interval is divided into N elementary sections of Δt duration. During the Δt

interval, the REE may remain in an serviceable condition or a failure of the REE may occur with subsequent restoration. We assume that the events at different elementary intervals Δt are independent.

A random number of failures n may occur in the observation interval T_Σ . The number of events, when there are no failures, is equal to $N - n$.

So, two events are possible in scheduled maintenance S_1 strategy. Statistical estimates of these events probabilities are the following:

$$P_S = n/N, \quad P_F = (N - n)/N.$$

If the model of operating time between failures is known, then it is possible to determine the numerical characteristics of the estimates, at least the mathematical expectation and variance. In this article, as in many other works, we limited ourselves to considering only mathematical expectation.

We assume that in the absence of failure the average costs are equal to C_M . If on the Δt interval failure occurs, then the costs are equal to C_R . Table 1 presents the components of the event-probability model for S_1 strategy.

Table 1 – Components of an event-probability model for S_1 strategy

Events	Statistical estimates of the probability of an event	Average costs due to an event
Serviceable condition over time Δt	P_S	C_M
Failure and maintenance	P_F	C_R

Let us make an expression for the effectiveness indicator of the S_1 strategy based on Table 1:

$$E = m_1(C_\Sigma / S_1) = m_1(P_S)C_M + m_1(P_F)C_R,$$

where $m_1(P_F)$ and $m_1(P_S)$ are the mathematical expectations of the probability of REE failure and the failure-free operation in the Δt interval, respectively; C_Σ are the total operating costs.

Let us draw up a similar table of the event-probability model for S_2 strategy. Since the S_2 strategy involves making decisions based on data processing, the table will be more complex. It will contain the following components: event, event probability, possible decisions based on processing the operational data and their probabilities, as well as average costs.

Table 2 presents the components of the event-probability model for S_2 strategy. For the second strategy, an event of the possible failure occurrence is introduced. In this case, data processing algorithms can detect or skip it.

Table 2 – Components of the event-probability model for S_2 strategy

Events	Statistical estimates of the probability of an event	The decision on the technical condition of the REE	Probability of a decision	Average costs
Serviceable condition over time Δt	P_S	REE operable	$1-\alpha$	C_P
		REE inoperable	α	$C_P + C_M$
Failure and maintenance	P_F	REE operable	β	$C_P + C_R$
		REE inoperable (preventive maintenance)	$1-\beta$	$C_P + C_{PM}$

In Table 2, the values of C_P and C_{PM} are the costs of data processing and preventive maintenance, as a result of which failure does not occur; α is the conditional probability of making a decision about an inoperative condition in the absence of failure; β is the conditional probability of making a decision about an operable condition in the presence of failure.

Based on Table 2, let us make an expression for the effectiveness indicator of S_2 strategy:

$$E = m_1(C_\Sigma / S_2) = m_1(P_S)((1-\alpha)C_P + \alpha(C_P + C_M)) + m_1(P_F)(\beta(C_P + C_R) + (1-\beta)(C_P + C_{PM})).$$

Let us compare the effectiveness of S_1 and S_2 strategies by calculating the dimensionless coefficient γ of the following form:

$$\gamma = \frac{m_1(C_\Sigma / S_1)}{m_1(C_\Sigma / S_2)}.$$

After mathematical simplifications, we obtain the following expression:

$$\gamma = \frac{C_R + m_1(P_S)(C_M - C_R)}{C_P + C_F + m_1(P_S)(\alpha C_M - C_F)},$$

where $C_F = C_{PM} + \beta C_R - \beta C_{PM}$.

To analyze the coefficient γ , we calculate its limit ratios, on the assumption of $\alpha = \beta = 0$:

$$\gamma(\alpha = \beta = 0) = \frac{C_R + m_1(P_S)(C_M - C_R)}{C_P + C_{PM} - m_1(P_S)C_{PM}}.$$

If we assume that $C_P \ll C_R$, then after dividing the numerator and denominator by C_R , we get:

$$\begin{aligned} \gamma(\alpha = \beta = 0) &\approx \frac{m_1(P_S) \frac{C_M}{C_R} + (1 - m_1(P_S))}{\frac{C_{PM}}{C_R} (1 - m_1(P_S))} = \\ &= \frac{C_M}{C_{PM}} \frac{m_1(P_S)}{(1 - m_1(P_S))} + \frac{C_R}{C_{PM}}. \end{aligned}$$

From this ratio it follows that for more reliable equipment ($m_1(P_S) \rightarrow 1$), the value $\gamma \gg 1$. Therefore, the application of S_2 strategy is more appropriate.

Consider the option of calculating the limit ratio for the case of highly reliable equipment ($m_1(P_S) \rightarrow 1$) and the first and second kind errors presence. Then we write the coefficient γ in the following form:

$$\gamma(m_1(P_S) \rightarrow 1) \approx \frac{C_M}{C_P + \alpha C_M}.$$

Since $C_P \ll C_M$, we get:

$$\gamma(m_1(P_S) \rightarrow 1) \approx \frac{1}{\alpha}.$$

It follows that the application of S_2 strategy is more preferable.

Nomograms of the coefficient γ dependence on conditional probabilities α, β and the probabilities of the REE operational condition for the cost option $C_P = 5$, $C_{PM} = 50$, $C_M = 75$, $C_R = 500$ are shown in Fig. 2.

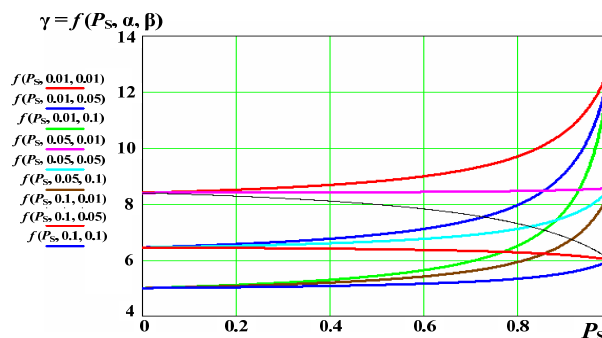


Figure 2 – Nomograms of the coefficient γ dependence

As can be seen from the graphs in Fig. 2, the coefficient $\gamma > 1$ for all initial data variants. This indicates that the use of the operational data processing by algorithm is reasonable and appropriate.

4 EXPERIMENTS

Preventive actions are possible in S_2 strategy. These actions can be carried out in various ways. In this paper,

we consider a variant connected with the preventive threshold presence [19]. Having a preventive threshold will eliminate potential failure [20].

Fig. 3 shows the trend of the monitoring parameter $X(t)$. Preventive thresholds are marked as P_+ and P_- , operational thresholds – as O_+ and O_- .

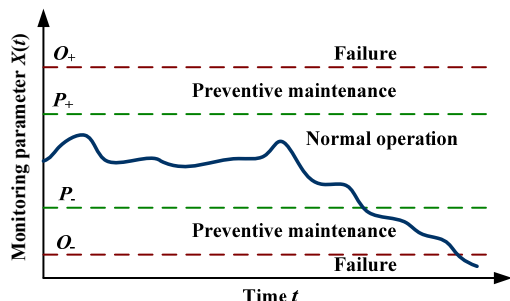


Figure 3 – The monitoring parameter trend

In accordance with S_2 strategy, the monitoring parameter values can be in one of three areas:

1) a normal operation area

$$P_- \leq X(t) \leq P_+;$$

2) an area of planning and performing the preventive maintenance

$$O_- \leq X(t) < P_-; P_+ < X(t) \leq O_+;$$

3) an area of inoperative state

$$X(t) < O_-; X(t) > O_+.$$

In the normal operation area, periodic monitoring is carried out using computer technology.

In the malfunctioning area, current repair is performed. After performing it, the REE is operational in the considered determining parameter.

In the case when $P_+ = O_+$ and $P_- = O_-$, the S_2 strategy coincides by nature with the S_1 strategy.

Let us perform statistical simulation of S_2 strategy based on the Monte-Carlo method. The main task of simulation will be to determine the optimal decision making threshold for preventive maintenance performing.

We will present the operational data in the statistical model form of the determining parameter trend of this kind:

$$y(t) = y_1(t) + y_2(t),$$

where $y_1(t)$ is a component that characterizes the trend in a time interval where there is no changepoint; $y_2(t)$ is a component that characterizes the changepoint.

In this study, we consider the following version of the trend components models of determining parameter:

$$y_1(t) = a_0 + x(t),$$

where a_0 is the initial determining parameter value in the $[P_-; P_+]$ interval; $x(t)$ is the fluctuation component due to the instability of the determining parameter, errors in the measuring equipment, etc. An automatic monitoring system of the REE condition is used in S_2 strategy. The measurement period is very short. The initial measurement results $y_1(t)$ averaging is performed to reduce fluctuation amplitudes.

During modeling, we assume that after maintenance and repair:

$$a_0 = \frac{P_- + P_+}{2}.$$

The fluctuation component has a Gaussian model with zero mathematical expectation and a given standard deviation σ .

The $y_2(t)$ component will be represented by the following model:

$$y_2(t) = a_1(t - t_d)h(t - t_d),$$

where a_1 is the tangent of the technical condition deterioration linear trend slope; t_d is the moment the changepoint begins; $h(t)$ is the Heaviside function.

In the general case a_1 and t_d values may be random. In this paper, we assume that they are described by a uniform distribution.

Let us describe the initial group of parameters that must be set to describe the OS:

1) the determined temporal characteristics of the model: the observation interval T_Σ , the duration of the scheduled maintenance preparation and performing, the repair duration t_R , the control operations frequency (system time) t_c ;

2) the determined parameters that characterize the average cost of processing C_p , preventive maintenance C_{PM} and repair C_R ;

3) the random variables that are characterized by probabilistic laws: $x(t)$, a_1 and t_d .

During simulation, we assume that the normal distribution law describes the fluctuation component. Therefore, as a result of averaging the $y_1(t)$, we also obtain a normal distribution with a standard deviation, depending on the number of samples during averaging.

When modeling the $y_2(t)$ component, assumptions were made about the uniform probability density function of a_1 and t_d values.

As a result of simulation, we will estimate the efficiency indicator in the form of total costs per unit:

$$E^* = \frac{m_1(C_\Sigma / S_2)}{T_\Sigma}.$$

The efficiency indicator in general case is a function of the preventive threshold value P_+ (or P_-). The simulation process solves the problem of finding the minimum effectiveness estimation E^* value and the preventive threshold value that corresponds to it.

5 RESULTS

The initial data for simulation are such values: $C_P = 0.01$, $C_{PM} = 100$, $C_R = 1000$, $t_c = 1$, $t_{PM} = 40$, $t_R = 300$, $\sigma = 10$, $a_0 = 200$, $O_+ = 300$, $O_- = 100$. The observation period is a random variable, but the maximum number of possible failures is $n = 40$. The changepoint beginning moment is described by a uniform probability density function in the $[50; 100]$ interval. The deterioration trend slope is also uniform in the $[-75; -30]$ interval.

An example of the determining parameter trend implementation for the given initial data set is presented in Fig. 4.

Figure 4 shows the determining parameter trend over three possible events in the OS. The first event is a case of prevented failure. At the same time, in OS managed to perform preventative maintenance. The second and third events correspond to the case when the failure occurs. At the same time, in OS there is no time to perform preventive maintenance.

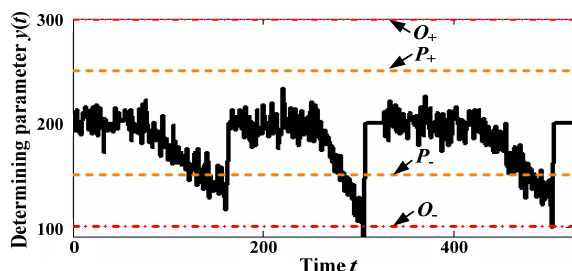


Figure 4 – An example of the determining parameter trend implementation

An example of a single procedure for evaluating the efficiency indicator for different values of the preventive threshold for the given initial data set is shown in Fig. 5.

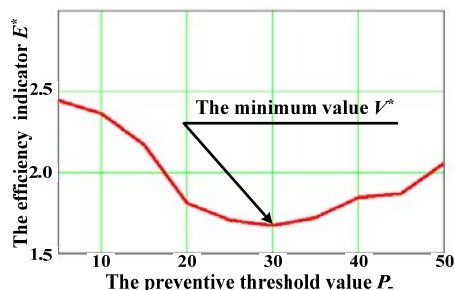


Figure 5 – The efficiency indicator estimates dependence on the preventive threshold value

The number of simulation procedures is 100. In each simulation procedure, the minimum value estimation is a

random variable. Therefore, it is advisable to calculate the mathematical expectation of the optimal value.

For a more accurate estimate of the V^* minimum value, the least-squares method approximation using a polynomial of the third degree is used in each simulation procedure. Such actions correspond to the proposed simulation-analytical-calculation method.

As a result, the mathematical expectation of optimal preventive threshold estimation is obtained $m_1^*(V^*) = 30.845$.

6 DISCUSSION

From the obtained results, the following conclusions can be made.

Operation is the longest period of the life cycle of REE and other technical systems. The analysis shows that operation costs are a significant value that exceeds the initial cost of the equipment. Therefore, the task of reducing operation costs deserves the attention of scientists.

The most common equipment maintenance strategy is the condition-based maintenance with the determining parameters monitoring. Authors proposed to improve this strategy by using additional statistical processing of determining parameters.

Analysis of scientific results has shown that not enough attention is paid to optimizing the preventive threshold, which is an integral part of the maintenance strategy. Therefore, the article defines a new efficiency indicator based on the use of probabilistic-event model. This model reflects the application of the method of statistical classification of radio electronic equipment technical condition and is used to find the analytical formula for the efficiency indicator. The efficiency indicator also takes into account the model of the determining parameter, which is the basis for solving the problem of finding the preventive threshold optimal value.

The simulation results analysis shows that the dependence $E^* = f(V^*)$ always has a minimum.

CONCLUSIONS

The article considers the task of substantiating the appropriateness of using statistical data processing in the REE operation system. An efficiency indicator in the form of average total costs is proposed. For this indicator, a comparative analysis of two strategies for planning and performing maintenance and repair work was carried out: S_1 strategy of performing periodic maintenance without trend analysis of the change in the determining parameter; S_2 strategy of performing preventive maintenance based on the statistical data analysis. For the performance indicator nomograms were built, which showed that S_2 strategy is more effective than S_1 .

Performance indicator is a function of the possible failures probability. Such failures can be eliminated using preventative maintenance. Therefore, a maintenance strategy with a preventive threshold was considered. This strategy helps to prevent possible failures and therefore reduces operating costs. The effectiveness analysis of the maintenance strategy with a preventive threshold was carried out based on the Monte-Carlo method. The problem of optimizing the preventive threshold value is solved based on the simulation results. This result was obtained using the analytical-calculation-modeling method.

The obtained results can be used during the design of new statistical data processing subsystem in the operation systems of REE.

The scientific novelty. The article proposes the method for determining the optimal preventive threshold for the condition-based maintenance strategy, this method is based on statistical processing of operational data represented by trends of the determining parameters.

The practical significance. The application of the method allows reducing the specific operation costs during radioelectronic equipment intended use.

Prospects for further research. Further research studies could be associated with determining analytical formulas for the optimal values of the preventive threshold, the use of different models of probability density functions of determining parameters trends, the use of more complex statistical processing, in particular the application of procedures for technical condition forecasting.

REFERENCES

1. Zhukov I.A. Implementation of integral telecommunication environment for harmonized air traffic control with scalable flight display systems, *Aviation*, 2010, Vol. 14, No. 3, pp. 117–122.
2. Zaliskyi M., Petrova Yu., Asanov M., Bekirov E. Statistical data processing during wind generators operation, *International Journal of Electrical and Electronic Engineering & Telecommunications*, 2019, Vol. 8, No. 1, pp. 33–38. DOI:10.18178/ijeetc.8.1.33-38
3. Solomentsev O., Zaliskyi M., Herasymenko T., Kozhokhina O., Petrova Yu. Data Processing in Case of Radio Equipment Reliability Parameters Monitoring, *Advances in Wireless and Optical Communications (RTUWO 2018). Latvia, 15–16 November, 2018: proceedings*. Riga, 2018, pp. 219–222.
4. Gertsbakh I. Reliability theory: with applications to preventive maintenance. New York, Springer, 2005, 219 p.
5. Zaliskyi M., Solomentsev O. Method of sequential estimation of statistical distribution parameters, *IEEE 3rd International Conference Methods and Systems of Navigation and Motion Control (MSNMC), Ukraine, 14–17 October, 2014: proceedings*. Kyiv, 2014, pp. 135–138.
6. Prokopenko I. G., Migel S. V., Prokopenko K. I. Signal modeling for the efficient target detection tasks, *International Radar Symposium, Germany, 19–21 June, 2013: proceedings*. Dresden, 2013, pp. 976–982.
7. Kharchenko V. P., Kuzmenko N. S., Ostroumov I. V. Identification of unmanned aerial vehicle flight situation, *IEEE 4th International Conference on Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), Ukraine, October, 2017: proceedings*. Kyiv, 2017, pp. 116–120.
8. Kuzmenko N. S., Ostroumov I. V., Marais K. An accuracy and availability estimation of aircraft positioning by navigational aids, *IEEE International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), Ukraine, 16–18 October, 2018: proceedings*. Kyiv, 2018, pp. 36–40.
9. Jardine A. K. S., Tsang A. H. C. Maintenance, replacement, and reliability: theory and applications, Second edition. BocaRaton, CRC Press, 2017, 364 p.
10. Natsional'na transportna stratehiya Ukrainy na period do 2030 roku [Electronic resource]. Access mode: <https://zakon.rada.gov.ua/laws/show/430-2018-p>.
11. Grall A., Dieulle L., Bérenguer C., Roussignol M. Continuous-time predictive-maintenance scheduling for a deteriorating system, *IEEE Transactions on reliability*, 2002, Vol. 51, No. 2, pp. 141–150. DOI: 10.1109/TR.2002.1011518
12. Nakagawa T. Maintenance theory of reliability. London, Springer-Verlag, 2005, 270 p.
13. Galar D., Sandborn P., Kumar U. Maintenance costs and life cycle cost analysis. BocaRaton, CRC Press, 2017, 492 p.
14. Solomentsev O., Zaliskyi M., Zuiiev O. Estimation of quality parameters in the radio flight support operational system, *Aviation*, 2016, Vol. 20, No. 3, pp. 123–128. DOI: 10.3846/16487788.2016.1227541.
15. Goncharenko A. Aircraft operation depending upon the uncertainty of maintenance alternatives, *Aviation*, 2017, Vol. 21, No. 4, pp. 126–131. DOI: 10.3846/16487788.2017.1415227.
16. Solomentsev O., Zaliskyi M., Herasymenko T., Kozhokhina O., Petrova Yu. Efficiency of Operational Data Processing for Radio Electronic Equipment, *Aviation*, 2019, Vol. 23, No. 3, pp. 71–77.
17. Hryshchenko Y. V. Reliability problem of ergatic control systems in aviation, *IEEE 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), Ukraine, 18–20 October, 2016: proceedings*. Kyiv, 2016, pp. 126–129.
18. Volianskyi R., Sadovoi O., Volianska N., Sinkevych O. Construction of Parallel Piecewise-Linear Interval Models for Nonlinear Dynamical Objects, *9th International Conference on Advanced Computer Information Technologies (ACIT), Czech Republic, 5–7 June, 2019: proceedings*. Ceske Bidejovice, 2019, pp. 97–100.
19. Huynh K. T., Grall A., Bérenguer C. A parametric predictive maintenance decision-making framework considering improved system health prognosis precision, *IEEE Transactions on Reliability, Institute of Electrical and Electronics Engineers*, 2018, Vol. 68, No. 1, pp. 375–396. DOI:10.1109/TR.2018.2829771.
20. Sun Q., Ye Z.-S., Peng W. Scheduling preventive maintenance considering the saturation effect, *IEEE Transactions on reliability, Institute of Electrical and Electronics Engineers*, 2019, Vol. 68, No. 2, pp. 741–752. DOI: 10.1109/TR.2018.2874265.

Received 26.10.2020.
Accepted 21.03.2021.

УДК 621.396.6.004.15

ОПТИМІЗАЦІЯ ПРЕВЕНТИВНОГО ПОРОГУ ПІД ЧАС ТЕХНІЧНОГО ОБСЛУГОВУВАННЯ ЗА СТАНОМ РАДІОЕЛЕКТРОННОГО ОБЛАДНАННЯ

Соломенцев О. В. – д-р техн. наук, професор кафедри телекомунікаційних та радіоелектронних систем, Національний авіаційний університет, Київ, Україна.

Заліський М. Ю. – канд. техн. наук, доцент кафедри телекомунікаційних та радіоелектронних систем, Національний авіаційний університет, Київ, Україна.

Щербина О. А. – канд. техн. наук, доцент кафедри електроніки, робототехніки і технологій моніторингу та інтернету речей, Національний авіаційний університет, Київ, Україна.

Асанов М. М. – канд. техн. наук, Сімферополь, Україна.

АНОТАЦІЯ

Актуальність. Експлуатаційні витрати протягом всього життєвого циклу радіоелектронного обладнання складають значну величину, що набагато перевищує власне вартість обладнання. Тому актуальною науково-технічною проблемою є мінімізація експлуатаційних витрат. Одним із шляхів вирішення цієї проблеми є впровадження технологій оброблення статистичних даних у системах експлуатації радіоелектронного обладнання.

Мета роботи – підвищення ефективності стратегії технічного обслуговування за станом з контролем визначальних параметрів, яка широко застосовується в цивільній авіації.

Метод. Вирішення поставленої проблеми базується на знаходженні функціональної залежності показника ефективності у вигляді питомих експлуатаційних витрат від основних параметрів радіоелектронного обладнання та системи його експлуатації. Для визначення цієї залежності використовується імовірно-подієва модель, а також методи теорії ймовірності та математичної статистики, зокрема методи статистичної класифікації вибірових сукупностей та функціональних перетворень випадкових величин. Для визначення оптимального рівня превентивного порога за критерієм мінімізації експлуатаційних витрат використовується метод статистичного моделювання Монте-Карло.

Результати. Удосконалена стратегія технічного обслуговування за станом з контролем визначальних параметрів, що базується на додатковій статистичній обробці експлуатаційних даних та технології знаходження оптимального превентивного порогу.

Висновки. Отримані результати можуть бути використані під час розробки та модернізації систем експлуатації радіоелектронного обладнання в частині застосування процедур оброблення статистичних даних. Порівняльний аналіз двох стратегій технічного обслуговування показав, що використання додаткової статистичної обробки даних дозволяє зменшити питомі експлуатаційні витрати. Запропонована технологія визначення оптимального превентивного порогу може бути поширена для використання під час експлуатації складних технічних систем, зокрема для таких, чий технічний стан пов'язаний зі значеннями визначальних параметрів.

КЛЮЧОВІ СЛОВА: показник ефективності, технічне обслуговування за станом, статистична обробка даних, система експлуатації, радіоелектронне обладнання.

УДК 621.396.6.004.15

ОПТИМИЗАЦИЯ ПРЕВЕНТИВНОГО ПОРОГА ВО ВРЕМЯ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ ПО СОСТОЯНИЮ РАДИОЭЛЕКТРОННОГО ОБОРУДОВАНИЯ

Соломенцев А. В. – д-р техн. наук, профессор кафедры телекоммуникационных та радиоэлектронных систем, Национальный авиационный университет, Киев, Украина.

Залиский М. Ю. – канд. техн. наук, доцент кафедры телекоммуникационных та радиоэлектронных систем, Национальный авиационный университет, Киев, Украина.

Щербина О. А. – канд. техн. наук, доцент кафедры электроники, робототехники и технологий мониторинга и интернета вещей, Национальный авиационный университет, Киев, Украина.

Асанов М. М. – канд. техн. наук, Симферополь, Украина.

АННОТАЦИЯ

Актуальность. Эксплуатационные расходы в течение всего жизненного цикла радиоэлектронного оборудования составляют значительную величину, намного превышающую собственно стоимость оборудования. Поэтому актуальной научно-технической проблемой является минимизация эксплуатационных затрат. Одним из путей решения этой проблемы является внедрение технологий обработки статистических данных в системах эксплуатации радиоэлектронного оборудования.

Цель работы – повышение эффективности стратегии технического обслуживания по состоянию с контролем определяющих параметров, которая широко применяется в гражданской авиации.

Метод. Решение поставленной проблемы базируется на нахождении функциональной зависимости показателя эффективности в виде удельных эксплуатационных расходов от основных параметров радиоэлектронного оборудования и системы его эксплуатации. Для определения этой зависимости используется вероятностно-событийная модель, а также методы теории вероятности и математической статистики, в частности методы статистической классификации выборочных совокупностей и функциональных преобразований случайных величин. Для определения оптимального уровня превентивного порога по критерию минимизации эксплуатационных расходов используется метод статистического моделирования Монте-Карло.

© Solomentsev O. V., Zaliskyi M. Yu., Shcherbyna O. A., Asanov M. M., 2021
DOI 10.15588/1607-3274-2021-2-2

Результаты. Усовершенствована стратегия технического обслуживания по состоянию с контролем определяющих параметров, основанная на дополнительной статистической обработке эксплуатационных данных и технологии нахождения оптимального превентивного порога.

Выводы. Полученные результаты могут быть использованы во время разработки и модернизации систем эксплуатации радиоэлектронного оборудования в части применения процедур обработки статистических данных. Сравнительный анализ двух стратегий технического обслуживания показал, что использование дополнительной статистической обработки данных позволяет уменьшить удельные эксплуатационные затраты. Предложенная технология определения оптимального превентивного порога может быть распространена для использования при эксплуатации сложных технических систем, в частности для таких, чье техническое состояние связано со значениями определяющих параметров.

КЛЮЧЕВЫЕ СЛОВА: показатель эффективности, техническое обслуживание по состоянию, статистическая обработка данных, система эксплуатации, радиоэлектронное оборудование.

ЛИТЕРАТУРА / LITERATURA

1. Zhukov I. A. Implementation of integral telecommunication environment for harmonized air traffic control with scalable flight display systems // *Aviation*. – 2010. – Vol. 14, № 3. – P. 117–122.
2. Statistical data processing during wind generators operation / [M. Zaliskyi, Yu. Petrova, M. Asanov, E. Bekirov] // *International Journal of Electrical and Electronic Engineering & Telecommunications*. – 2019. – Vol. 8, № 1. – P. 33–38. DOI:10.18178/ijeetc.8.1.33-38
3. Data Processing in Case of Radio Equipment Reliability Parameters Monitoring / [O. Solomentsev, M. Zaliskyi, T. Herasymenko et al.] // *Advances in Wireless and Optical Communications (RTUWO 2018)*, Latvia, 15–16 November, 2018: proceedings. – Riga, 2018. – P. 219–222.
4. Gertsbakh I. Reliability theory: with applications to preventive maintenance / I. Gertsbakh. – New York : Springer, 2005. – 219 p.
5. Zaliskyi M. Method of sequential estimation of statistical distribution parameters / M. Zaliskyi, O. Solomentsev // *IEEE 3rd International Conference Methods and Systems of Navigation and Motion Control (MSNMC)*, Ukraine, 14–17 October, 2014 : proceedings. – Kyiv, 2014. – P. 135–138.
6. Prokopenko I. G. Signal modeling for the efficient target detection tasks / I. G. Prokopenko, S. V. Migel, K. I. Prokopenko // *International Radar Symposium, Germany, 19–21 June, 2013: proceedings*. – Dresden, 2013. – P. 976–982.
7. Kharchenko V. P. Identification of unmanned aerial vehicle flight situation / V. P. Kharchenko, N. S. Kuzmenko, I. V. Ostroumov // *IEEE 4th International Conference on Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD)*, Ukraine, October, 2017: proceedings. – Kyiv, 2017. – P. 116–120.
8. Kuzmenko N. S. An accuracy and availability estimation of aircraft positioning by navigational aids / N. S. Kuzmenko, I. V. Ostroumov, K. Marais // *IEEE International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*, Ukraine, 16–18 October, 2018: proceedings. – Kyiv, 2018. – P. 36–40.
9. Jardine A. K. S. Maintenance, replacement, and reliability: theory and applications, Second edition / A. K. S. Jardine, A. H. C. Tsang. – BocaRaton : CRC Press, 2017. – 364 p.
10. Національна транспортна стратегія України на період до 2030 року [Електронний ресурс]. – Режим доступу: <https://zakon.rada.gov.ua/laws/show/430-2018-p>.
11. Continuous-time predictive-maintenance scheduling for a deteriorating system / [A. Grall, L. Dieulle, C. Bérenguer, M. Roussignol] // *IEEE Transactions on reliability*. – 2002. – Vol. 51, № 2. – P. 141–150. DOI: 10.1109/TR.2002.1011518
12. Nakagawa T. Maintenance theory of reliability / T. Nakagawa. – London : Springer-Verlag, 2005. – 270 p.
13. Galar D. Maintenance costs and life cycle cost analysis / D. Galar, P. Sandborn, U. Kumar. – BocaRaton : CRC Press, 2017. – 492 p.
14. Solomentsev O. Estimation of quality parameters in the radio flight support operational system / O. Solomentsev, M. Zaliskyi, O. Zuiev // *Aviation*. – 2016. – Vol. 20, № 3. – P. 123–128. DOI: 10.3846/16487788.2016.1227541.
15. Goncharenko A. Aircraft operation depending upon the uncertainty of maintenance alternatives / A. Goncharenko // *Aviation*. – 2017. – Vol. 21, № 4. – P. 126–131. DOI: 10.3846/16487788.2017.1415227.
16. Efficiency of Operational Data Processing for Radio Electronic Equipment / [O. Solomentsev, M. Zaliskyi, T. Herasymenko et al.] // *Aviation*. – 2019. – Vol. 23, № 3. – P. 71–77.
17. Hryshchenko Y. V. Reliability problem of ergatic control systems in aviation / Y. V. Hryshchenko // *IEEE 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*, Ukraine, 18–20 October, 2016 : proceedings. – Kyiv, 2016. – P. 126–129.
18. Construction of Parallel Piecewise-Linear Interval Models for Nonlinear Dynamical Objects / [R. Volianskyi, O. Sadovoi, N. Volianska, O. Sinkevych] // *9th International Conference on Advanced Computer Information Technologies (ACIT)*, Czech Republic, 5–7 June, 2019: proceedings. – Ceske Bidejovice, 2019. – P. 97–100.
19. Huynh K. T. A parametric predictive maintenance decision-making framework considering improved system health prognosis precision / K. T. Huynh, A. Grall, C. Bérenguer // *IEEE Transactions on Reliability, Institute of Electrical and Electronics Engineers*. – 2018. – Vol. 68, № 1. – P. 375–396. DOI: 10.1109/TR.2018.2829771.
20. Sun Q. Scheduling preventive maintenance considering the saturation effect / Q. Sun, Z.-S. Ye, W. Peng // *IEEE Transactions on reliability, Institute of Electrical and Electronics Engineers*. – 2019. – Vol. 68, № 2. – P. 741–752. DOI: 10.1109/TR.2018.2874265.

ПОРЯДОК ОЦІНКИ НЕСУЧОЇ ЧАСТОТИ СИГНАЛУ, ЩО ПРИЙМАЄТЬСЯ СУПУТНИКОВОЮ СИСТЕМОЮ ЗВ'ЯЗКУ В БЕЗПЕРЕРВНОМУ РЕЖИМІ

Туровський О. Л. – канд. техн. наук, доцент, професор кафедри засобів захисту інформації Національного авіаційного університету Київ, Україна.

Дробик О. В. – канд. техн. наук, професор, директор Наукового центру, Державний університет телекомунікацій, Київ, Україна.

АНОТАЦІЯ

Актуальність. Однією з особливостей систем супутникового зв'язку є переважне використання в них під час прийому сигналу в безперервному режимі фазової модуляції сигналів, що призначені для передачі корисної інформації. Використання вказаного виду модуляції потребує вирішення завдання щодо оцінки несучої частоти сигналу. А сама оцінка зводиться до задачі оцінки частоти максимуму в спектрі фрагменту синусоїдального сигналу на фоні адитивного гаусівського шуму. В статті розглянуто процес оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі по правилу максимальної правдоподібності.

Мета. Розробка порядку оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі по правилу максимальної правдоподібності.

Метод. Запропонований в роботі порядок та розроблений на його основі алгоритм дає змогу здійснити оцінку несучої частоти по правилу максимальної правдоподібності з врахуванням умов невизначеності всіх параметрів сигналу, що приймається супутниковою системою зв'язку в безперервному режимі.

Результати. З метою практичного впровадження вказаного алгоритму в діючі схеми супутникового зв'язку, в роботі запропоновані схеми його апаратної реалізації. Для ілюстрації співвідношення кордонів мінімально граничної дисперсії оцінки несучої частоти в роботі подані залежності, які дозволяють здійснити порівняння мінімально граничної дисперсії, визначеної нижнім кордоном Крамера-Рао та мінімально граничної дисперсії, визначеної з урахуванням всіх параметрів сигналу.

Висновки. Аналіз поданих залежностей показав, що в реальних умовах мінімальна дисперсія проведення несучої частоти сигналу по правилу максимальної правдоподібності, що приймається супутниковою системою зв'язку в безперервному режимі при невизначеності всіх параметрів сигналу може істотно відрізнятись від мінімальної дисперсії, отриманої на основі застосування нижнього кордону Крамера-Рао. Перспективи подальших досліджень слід спрямувати на розробку та створення алгоритмів та методик, спрямованих на МП-оцінку при умові максимального наближення мінімально граничної дисперсії оцінки несучої частоти реального сигналу до нижнього кордону Крамера-Рао, визначеного для оцінки несучої частоти в умовах визначеності інших параметрів сигналу.

КЛЮЧОВІ СЛОВА: сигнал, що приймається, оцінка несучої частоти сигналу, мінімально гранична дисперсія оцінки несучої частоти, нижній кордон Крамера-Рао.

АБРЕВІАТУРИ

МП-оцінка – оцінки несучої частоти сигналу по правилу максимальної подібності.

НОМЕНКЛАТУРА

ν – зміщення частоти вхідного сигналу від її номінального значення;

ϕ – початкова фаза сигналу;

τ – час затримки сигналу;

d – передана інформаційна послідовність сигналу;

T_s – інтервал слідування відліків комплексної огинаючої сигналу, що приймається;

T – тактовий інтервал спостереження за вхідним сигналом;

E_s – енергія одиничного імпульсу вхідного сигналу;

N_0 – інтервал спостережень, виражений в кількості інтервалів T_s ;

K – інтервал спостережень, виражений в тактових інтервалах;

Ω – приведені значення зміщення частоти вхідного сигналу від її номінального значення;

\approx

ϕ – випадкова величина фази сигналу;

\approx

τ – випадкова величина часу затримки сигналу;

\approx

Ω – випадкова величина приведеного значення зміщення частоти вхідного сигналу;

$z(t_n)$ – функція коливання вхідного сигналу;

I_n – дійсна частина функції коливання вхідного сигналу;

h – імпульсна характеристика фільтра, що формує вхідний сигнал;

\approx

a – вектор незалежних величин вхідного сигналу;

j – коефіцієнт комплексної складової величини, пов'язаної з вхідним сигналом;

Q_n – постійна складова дійсної частини вхідного сигналу по змінній коливання, при розрахунку щіль-

ності ймовірності по кількості відліків інтервалів спостереження N_0 за вхідним сигналом;

i_n – дійсна частина функції коливання вхідного сигналу по випадковій змінній комплексної огинаючої опорного сигналу;

q_n – постійна складова дійсної частини вхідного сигналу по випадковій змінній комплексної огинаючої опорного сигналу;

t_n – час оцінки n -го імпульсу вхідного сигналу;

σ^2 – значення дисперсії похибки оцінки частоти вхідного сигналу;

σ_v^2 – значення дисперсії похибки оцінки частоти вхідного сигналу при відсутності даних про v ;

$CRLB(v)$ – нижній кордон Крамера-Рао для дисперсії оцінки несучої частоти сигналу, що приймається, коли є данні про всі інші параметри, крім v ;

Λ – функція оцінки несучої частоти вхідного сигналу по ймовірним значенням його характеристик;

$E_{\varphi, \tau, d}^{\approx}$ – усереднене значення функції оцінки несучої частоти вхідного сигналу по ймовірним усередненим значенням φ і d ;

D_S – позитивне значення усередненої по інформаційним послідовностям функції оцінки несучої частоти вхідного сигналу;

$r^*(t_n)$ – функція комплексної огинаючої вхідного сигналу, відкорельована дійсними випадковими значеннями комплексних складових;

y – функція відклику узгодженого фільтра тракту прийому вхідного сигналу на вплив випадкового вхідного сигналу.

ВСТУП

Робота сучасних систем супутникового зв'язку вимагає обробки вхідних сигналів, що приймаються демодуляторами вхідних пристроїв як в безперервному, так і в пакетному режимі передачі даних. Під час прийому сигналу в безперервному режимі у вказаних демодуляторах переважно застосовується фазова модуляція сигналів, що призначені для передачі корисної інформації [1]. В свою чергу, при використанні такого виду модуляції в супутникових каналах зв'язку виникає проблема істотної невизначеності сигналу, що приймається по частоті (частотна невизначеність сигналу). Для її вирішення в фазокогерентних системах супутникового зв'язку для відновлення несучої і тактової частот та для когерентної демодуляції аналогових і цифрових сигналів з кутовою модуляцією здійснюється синхронізація по частоті несучого коливання сигналу, що приймається.

Об'єктом дослідження є процес синхронізації по частоті несучого коливання сигналу, що приймається супутниковою системою зв'язку.

© Туровський О. Л., Дробик О. В., 2021
DOI 10.15588/1607-3274-2021-2-3

Комплексна огинаюча сигналу містить невідомі величини, а саме частоту (v), фазу (φ), затримку сигналу (τ) та передану інформаційну послідовність (d). Завдання синхронізації фактично зводиться до оцінки істинних параметрів сигналу, що приймається – v, φ, τ, d , знання значень яких необхідно для демодуляції сигналу $z(t)$.

Тобто вказана синхронізація при фазовій модуляції сигналу потребує вирішення завдання щодо оцінки несучої частоти сигналу.

Для вирішення завдання оцінки несучої частоти ФМ сигналу в умовах невизначеності інформації φ, τ, d в ряді робіт пропонується застосувати правило максимальної правдоподібності. Відомо, що використання правила максимальної правдоподібності для оцінки несучої частоти (МП-оцінки) забезпечує асимптотично ефективну і асимптотично незміщену оцінку [2, 3,]. Сутність якої полягає в визначенні максимуму функції $\Lambda(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d})$ по випадковим значенням

величин вхідного сигналу. При наявності інформації про параметри φ, τ, d пошук $\Lambda(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d})$ по v може

забезпечити мінімально граничну дисперсію σ^2 , яка буде визначатися нижнім кордоном $CRLB(v)$ [3, 4].

При великих інтервалах спостереження ($K \gg 1$) значення дисперсії похибки оцінки частоти вхідного фазомодульованого сигналу (σ^2) може бути подане як функціональна залежність $\sigma^2 = f(E_S, T_s, N_0)$ [5, 6, 7]. Запропонований підхід дозволяє визначити нижній кордон $CRLB(v)$ при умові, що інші параметри сигналу відомі. Тобто оцінка проводиться в ідеальних умовах.

Зазвичай в умовах реального супутникового каналу частина інформації про параметри сигналу, що приймається, може бути відсутня. Виходячи з цього, мінімальна σ^2 оцінки несучої частоти реального супутникового сигналу може дещо відрізнятись в гіршу сторону від значення нижнього кордону $CRLB(v)$, отриманого при цих же умовах [6].

Предметом дослідження є процес оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі. Вказана оцінка зводиться до задачі оцінки частоти максимуму

$\Lambda(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d})$ в спектрі фрагмента синусоїдального сигналу, що приймається в безперервному режимі на фоні адитивного гаусівського шуму.

Підвищення якості оцінки несучою частотою фазомодульованого сигналу, що приймається супутниковою системою зв'язку в безперервному режимі потребує розробки відповідної нової математичної моделі

процесу оцінки та створення на її основі відповідного порядку оцінки частоти.

Мета роботи. Розробка порядку оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі по правилу максимальної правдоподібності.

1 ПОСТАНОВКА ЗАДАЧІ

Достовірна МП-оцінка несучої частоти сигналу, що приймається супутниковою системою зв'язку по функції $\Lambda(\tilde{\nu}, \tilde{\varphi}, \tilde{\tau}, \tilde{d})$ при прийнятті умови, що ν невідома, вимагає вирішення актуального наукового завдання щодо визначення мінімально граничної дисперсії оцінки сигналу $\sigma^2 = f(E_S, T_S, N_0)$ по нижньому кордону $CRB(\nu)$ та розробку алгоритму проведення вказаної оцінки з урахуванням випадкових значень $\tilde{\nu}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}$ сигналу, що приймається супутниковою системою зв'язку.

2 ОГЛЯД ЛІТЕРАТУРИ

Питання визначення оптимальних критеріїв для оцінки несучої частоти сигналу, що приймається системою зв'язку та розробки порядку її проведення присвячено ряд робіт.

Автори робіт [5, 8, 7] запропонували зважений байєсівський Крамерський Рао-кордон для спільного визначення часу синхронізації та зсуву несучої частоти, який враховує попередній розподіл параметрів оцінки та є точною нижньою межею для всіх розглянутих значень співвідношення сигнал/шум. Питання визначення мінімально граничної дисперсії оцінки саме несучої частоти в даній роботі не розглядалися.

В роботах [9, 10, 11] поданий кордон Крамера-Рао для дисперсії сумісної оцінки доплерівського зміщення несучої частоти та затримок сигналу з довільною шириною спектру. Подані результати запропоновано застосовувати для випадків, коли ширина спектру сигналу не дозволяє застосовувати інші методи оцінки в рамках вузькополосної моделі. Безпосередньо оцінка несучої частоти та визначення для неї мінімально граничної дисперсії в роботах не розглядаються.

Автори роботи [12] пропонують нижній кордон Крамера-Рао для на оцінки частоти когерентної послідовності імпульсів, пасивно перехоплених на рухомій антенні. Таку оцінку автора пропонують проводити з метою визначення місцезнаходження передаючого радару. Безпосередньо несуча частота вказаних імпульсів та її оцінка в роботі не розглядалися, хоча робота є показовою з точки зору застосування нижнього кордону Крамер-Рао для МП оцінки несучої частоти.

Робота [13] присвячена вирішенню завдання по оцінці затримки сигналу в часі з урахуванням зміни форми імпульса сигналу під час навігації в приміщен-

нях. Для оцінки часу затримки використовується критерій нижнього кордону Крамера-Рао. Робота достатньо показова з точки зору оцінки сигналу, але безпосередньо оцінка несучої частоти в ній не розглядається.

Автори роботи [14] пропонують вирази для дійсного нижнього кордону Крамера-Рао при сумісній оцінці доплерівського зміщення частоти та запізнення вхідного сигналу, а також оцінки синхронізації символів лінійно модульованого сигналу. В статті також обгрунтовані висновки про актуальність індивідуальної оцінки несучої частоти з відомою синхронізацією символів. Питання оцінки несучої частоти та визначення для неї мінімально граничної дисперсії в роботі не розглядаються.

В роботі [15] автори пропонують алгоритм на основі умовної максимізації очікування для сумісної оцінки каналу передачі даних в системах мультиплексування з ортогональним частотним розподілом каналів, фазового шуму генераторів та доплерівського зміщення несучої частоти. Для вказаної оцінки в роботі пропонується нижній гібридний кордон Крамер-Рао. Питання розрахунку нижнього кордону Крамер-Рао для несучої частоти та визначення для неї мінімально граничної дисперсії в роботі не розглядаються.

Питання визначення мінімально граничної дисперсії при оцінці якості роботи приймальної радіосистеми розглядалися в роботі [16]. В роботі безпосередньо запропоновані вирази для оцінки мінімальної дисперсії шуму та обгрунтовуються висновки про те, що така оцінка може вплинути на продуктивність системи на рівні оцінки зміщення несучої частоти та оцінки продуктивності каналу. Визначення значення нижнього кордону Крамера-Рао для запропонованої мінімально граничної дисперсії в роботі не проводилося.

В роботі [17] Розглянуті питання оцінки частоти для безперервних пакетів даних фіксованої довжини, що супроводжується шумом низького рівня. Мінімальна дисперсія оцінки частоти виводиться з точки зору спектральної щільності шуму на частотах, близьких до частоти сигналу. Але в роботі відсутній розгляд питання виводу нижнього кордону Крамер-Рао в якості мінімально граничної дисперсії для даного випадку.

3 МАТЕРІАЛИ ТА МЕТОДИ

Значне практичне значення для оцінки несучої частоти має визначення мінімально граничної дисперсії в умовах реального супутникового каналу, тобто в припущенні, що [1]:

- відсутня тактова синхронізація;
- відсутня інформація про інформаційну послідовність, що передається;
- відсутня інформація про початкову фазу несучого коливання.

Приймемо, що комплексна огинаюча опорного сигналу має вигляд [2]:

$$\tilde{r}(t) = \exp \left[j \left(2\pi \nu t + \tilde{\varphi} \right) \right] \sum_k \tilde{d}_k h \left(t - kT - \tilde{\tau} \right).$$

Прийmemo, що величина $\tilde{\nu}$ лежить в діапазоні від $-\nu_{\max}$ до ν_{\max} .

Інтервал $[-\nu_{\max}, \nu_{\max}]$ визначається початковою частотною невизначеністю сигналу, що приймається.

В подальшому прийmemo, що:

$$\tilde{d} = \left\{ \tilde{d}_0, \tilde{d}, \dots, \tilde{d}_{k-1} \right\} - \text{вектор незалежних випад-}$$

кових величин з нульовим середнім і з відповідними функціями кореляції [18].

Прийmemo, що значення величин $\tilde{\varphi}$ та $\tilde{\tau}$ і компоненти вектора \tilde{d} незалежні один від одного і всі інформаційні символи рівно ймовірні.

Таким чином, завдання полягає в тому, щоб отримати оцінку несучої частоти сигналу, що приймається по правилу максимальної правдоподібності (МП-оцінку) при відсутності інформації про передану інформаційну послідовність, початковій фазі несучої частоти сигналу, що приймається і тактовій синхронізації.

Запишемо сумісну функцію щільності ймовірності сигналу, що приймається в вигляді [19]:

$$p \left(z \uparrow \tilde{a} \right) = \left(\frac{1}{\sigma^2 2\pi} \right)^N \exp \left[-\frac{1}{2\sigma^2} \sum_n \left(I_n^2 + Q_n^2 - 2(I_n i_n + Q_n q_n) + i_n^2 + q_n^2 \right) \right]. \quad (1)$$

$$\text{де } \tilde{a} = \left\{ \tilde{\Omega}, \tilde{\varphi}, \tilde{\tau}, \tilde{d} \right\}; \quad I_n = \operatorname{Re} \left(z(t_n) \right); \quad Q_n = \operatorname{Im} \left(z(t_n) \right);$$

$$\tilde{r}(t_n) = \exp \left[j \left(\tilde{\Omega} t_n + \tilde{\varphi} \right) \right] \sum_k \tilde{d}_k h \left(t_n - kt - \tilde{\tau} \right).$$

комплексна огинаюча опорного сигналу при випадкових величинах його характеристик;

$$i_n = \operatorname{Re} \left(\tilde{r}(t_n) \right); \quad q_n = \operatorname{Im} \left(\tilde{r}(t_n) \right);$$

$$t_n = nT_s, n = -\frac{N-1}{2}, \dots, 0, 1, \dots, \frac{N-1}{2}.$$

Прийнявши, що $\sum_n I_n^2, \sum_n Q_n^2$ і $i_n^2 + q_n^2 = \left| \tilde{r}(t_n) \right|^2$ – по-

стійні, які не залежать від $\tilde{\nu}$, а $\left(\frac{1}{\sigma^2 2\pi} \right)^N$ – деяка по-

зитивна постійна величина, для рішення завдання по оцінці частоти несучого коливання по правилу максимальної подібності достатньо знайти максимум наступної функції [19]:

$$\Lambda \left(\tilde{\nu}, \tilde{\varphi}, \tilde{\tau}, \tilde{d} \right) = \exp \left[\frac{1}{\sigma^2} \sum_n \left(I_n i_n + Q_n q_n \right) \right]. \quad (2)$$

Відмітимо, що

$$I_n i_n + Q_n q_n = \operatorname{Re} \left((I_n + jQ_n)^* (i_n - jq_n) \right) = \operatorname{Re} \left(z(t_n)^* \tilde{r}^*(t_n) \right).$$

З урахуванням введених вище позначень, вираз (2) приймає наступний вигляд:

$$\Lambda \left(\tilde{\nu}, \tilde{\varphi}, \tilde{\tau}, \tilde{d} \right) = \exp \left[\frac{1}{\sigma^2} \operatorname{Re} \left[\sum_n z(t_n)^* \tilde{r}^*(t_n) \right] \right]. \quad (3)$$

Розглянемо суму $\sum_m \left(z(t_n)^* \tilde{r}^*(t_n) \right)$

$$\begin{aligned} \sum_m z(t_n)^* \tilde{r}^*(t_n) &= \sum_n z(t_n)^* \left(\exp \left[j \left(2\pi \tilde{\nu} t_n + \tilde{\varphi} \right) \right] \sum_k \tilde{d}_k h \left(t_n - kT - \tilde{\tau} \right) \right)^* = \\ &= \sum_n z(t_n)^* \exp \left(-j\tilde{\varphi} \right) \exp \left(-j2\pi \tilde{\nu} t_n \right) \sum_k \tilde{d}_k h \left(t_n - kT - \tilde{\tau} \right). \end{aligned}$$

Звідси

$$\begin{aligned} \sum_m z(t_n)^* \tilde{r}^*(t_n) &= \\ &= \exp \left(-j\tilde{\varphi} \right) \sum_k \tilde{d}_k \sum_n z(t_n) \exp \left(-j2\pi \tilde{\nu} t_n \right) h \left(t_n - kT - \tilde{\tau} \right). \end{aligned}$$

Прийmemo, що імпульсна характеристика фільтра h^* має кінцеву довжину, що дорівнює цілому числу символічних інтервалів сигналу, що приймається. Також прийmemo, що інтервал спостереження багато більше довжини імпульсної характеристики фільтра. При введених вище обмеженнях справедливе наступне співвідношення:

$$\sum_m \left(z(t_n)^* \tilde{r}^*(t_n) \right) = \exp \left(-j\tilde{\varphi} \right) \sum_k \tilde{d}_k y \left(\tilde{\nu}, kT + \tilde{\tau} \right),$$

$$\text{де } y \left(\tilde{\nu}, t \right) = \sum_n z(t_n) \exp \left(-j2\pi \tilde{\nu} t_n \right) h(t_n - t).$$

Фактично сигнал $y\left(\tilde{v}, t\right)$ можна трактувати як відклик узгодженого фільтра (УФ) на сигнал $z(t_n)^* \exp\left(-j2\pi\tilde{v}t\right)$.

Звідси, вираз (3) перепишемо в вигляді

$$\Lambda\left(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}\right) = \exp\left[\frac{1}{\sigma^2} \operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_n \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right]\right].$$

В подальшому ми повинні усереднити отриману функцію $\Lambda\left(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}\right)$ по так званим небажаним параметрам $\tilde{\varphi}$, $\tilde{\tau}$ та \tilde{d} . Процедура усереднення в загальному випадку є достатньо складною. Тому, для вирішення вказаної задачі усереднення розкладемо функцію $\Lambda\left(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}\right)$ в ряд Тейлора. Ми аналізуємо випадок малого відношення сигнал/шум, що дозволяє обмежитися лише трьома першими складовими розкладення [18]. Тому для $\Lambda\left(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}\right)$ запишемо:

$$\Lambda\left(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}\right) = 1 + \frac{1}{\sigma^2} \operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] + \frac{1}{2} \left[\frac{1}{\sigma^2} \operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right]\right]^2, \quad (4)$$

Усередним другу складову в (4) по інформаційній послідовності

$$E_d \left\{ \frac{1}{\sigma^2} \operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] \right\} = \frac{1}{\sigma^2} \operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k E_d^* \left\{ \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right) \right\} \right].$$

Всі інформаційні символи рівноймовірні.

Тому $E_d^* \left\{ \tilde{d}_k^* \right\} \rightarrow 0$ і

$$E_d \left\{ \frac{1}{\sigma^2} \operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] \right\}.$$

Розглянемо третю складову в (4). Відкинувши співмножники, що не залежать від \tilde{v} , запишемо

$$\left[\operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] \right]^2 = \frac{1}{4} \left[\sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right) \exp\left(-j\tilde{\varphi}\right) + \tilde{d}_k^* y^*\left(\tilde{v}, kT + \tilde{\tau}\right) \exp\left(j\tilde{\varphi}\right) \right]^2. \quad (5)$$

Звідси

$$\left[\operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] \right]^2 = \frac{1}{2} \sum_i \sum_k \tilde{d}_i^* \tilde{d}_k y\left(\tilde{v}, kT + \tilde{\tau}\right) y^*\left(\tilde{v}, kT + \tilde{\tau}\right) + \frac{1}{4} \exp\left(-j\tilde{\varphi}\right) \left(\sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right) \right)^2 + \frac{1}{4} \exp\left(j\tilde{\varphi}\right) \left(\sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right) \right)^2.$$

Відмітимо, що вираз (5) включає в себе складові з співмножниками виду $\exp\left(\pm 2j\tilde{\varphi}\right)$, усереднення яких по $\tilde{\varphi}$ дає 0.

Отже, усереднив вираз (5) по заважаючому параметру $\tilde{\varphi}$ і \tilde{d} , отримаємо

$$\Lambda\left(\tilde{v}, \tilde{\varphi}, \tilde{\tau}, \tilde{d}\right) = \left\{ \left[\operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] \right]^2 \right\} = E_d^* \left\{ \frac{1}{2} \sum_i \sum_k \tilde{d}_i^* \tilde{d}_k y\left(\tilde{v}, kT + \tilde{\tau}\right) y^*\left(\tilde{v}, kT + \tilde{\tau}\right) \right\}.$$

Звідси, з врахуванням функцій кореляції параметрів вектору $\tilde{d} = \left\{ \tilde{d}_0, \tilde{d}, \dots, \tilde{d}_{k-1} \right\}$ маємо:

$$E_{\tilde{\varphi}, \tilde{d}}^* \left\{ \left[\operatorname{Re}\left[\exp\left(-j\tilde{\varphi}\right) \sum_k \tilde{d}_k^* y\left(\tilde{v}, kT + \tilde{\tau}\right)\right] \right]^2 \right\} = \frac{D_S}{2} \sum_k \left| y\left(\tilde{v}, kT + \tilde{\tau}\right) \right|^2.$$

4 ЕКСПЕРИМЕНТИ

Усереднення виразу (4) по значенню τ , рівномірно розподіленому на інтервалі $\left[-\frac{T}{2}, \frac{T}{2}\right]$, дає кінцевий результат:

$$\tilde{\Lambda}(\tilde{\nu}) = 1 + \frac{1}{2\sigma^4} \frac{D_S}{2T} \sum_k \int_{-T/2}^{T/2} \left| y(\tilde{\nu}, kT + \tau) \right|^2 d\tau.$$

де $\tilde{\Lambda}(\tilde{\nu})$ функція правдоподібності

Співмножник $\frac{1}{2\sigma^4} \frac{D_S}{2T}$ можна виключити із роз-

гляд, так як він не залежить від $\tilde{\nu}$.

Тобто, для знаходження МП-оцінки значення $\tilde{\nu}$ достатньо знайти максимум функції

$$\tilde{\Lambda}(\tilde{\nu}) = \sum_k \int_{-T/2}^{T/2} \left| y(\tilde{\nu}, kT + \tau) \right|^2 d\tau. \quad (6)$$

Як підсумок, відмітимо, що сума в (6) еквівалентна інтегралу від функції $\left| y(\tilde{\nu}, t) \right|^2$ на інтервалі спостереження $T_0 = kT$.

Таким чином, для знаходження МП-оцінки несучої частоти сигналу необхідно максимізувати функцію

$$L(\tilde{\nu}) = \sum_k \int_{-T/2}^{T/2} \left| y(\tilde{\nu}, t) \right|^2 d\tau \quad (7)$$

По всім розглянутим частотам $\tilde{\nu}$, таким, що

$$\left| \tilde{\nu} \right| \leq \nu_{\max}.$$

Схема, що може реалізувати обрахунок функції $L(\tilde{\nu})$, подана на рис. 1 [20].

Реалізація алгоритму МП-оцінки полягає в наступному. Діапазон можливих значень $-\nu_{\max} \leq \tilde{\nu} \leq \nu_{\max}$ розбі'ємо на M інтервалів з середніми точками $\nu_k = \nu_0 + k\Delta\nu$, $\nu_0 = -\nu_{\max}$, $k = 0, 1, 2, \dots, M-1$.

Для кожної частоти ν_k обраховується значення інтеграла (7) і знаходиться максимум по всім ν_k .

На рис. 2 подана паралельна схема оцінки [20].

Застосування паралельної структури приводить до скорочення часу обрахування, але викликає ріст складності пристрою.

Схеми обрахування МП-оцінки несучої частоти, подані на рис. 1 та 2, як правило, називають схемами розімкнутого типу (open-loop frequency estimation algorithm) [4, 5, 6].

Альтернативний підхід полягає в застосуванні методів оцінювання замкнутого типу (closed-loop algorithm).

Як відомо [4, 19], знаходження максимуму функції $L(\tilde{\nu})$ досягається рішенням рівняння:

$$dL(\tilde{\nu}) / d\tilde{\nu} = 0. \quad (8)$$

В роботах [4, 5, 6, 7] показано, що на основі аналізу виразу (8) може бути отримана петля частотної автопідстройки, яка реалізує МП-оцінку несучої частоти сигналу, що приймається (closed-loop frequency estimation algorithm) при відсутності інформації про передану інформаційну послідовність, початкову фазу несучої частоти сигналу, що приймається і тактову синхронізацію. На основі аналізу лінійної моделі петлі оцінюється мінімальна гранична дисперсія МП-оцінки частоти несучого коливання, отримана на основі правила (8) [4, 21].

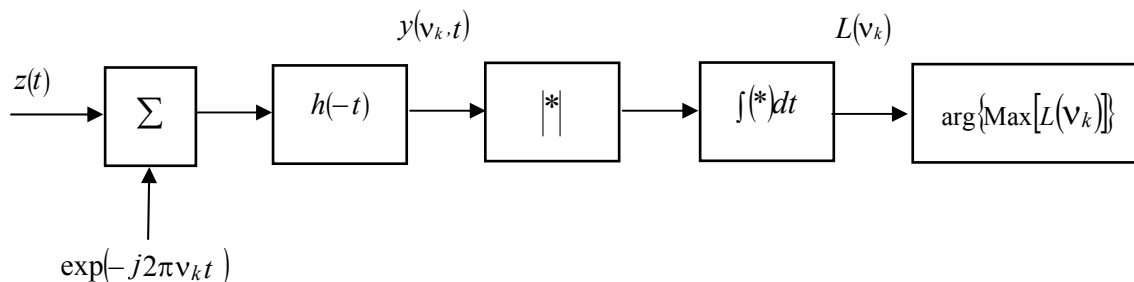


Рисунок 1 – Структурна схема послідовного обрахування МП-оцінки

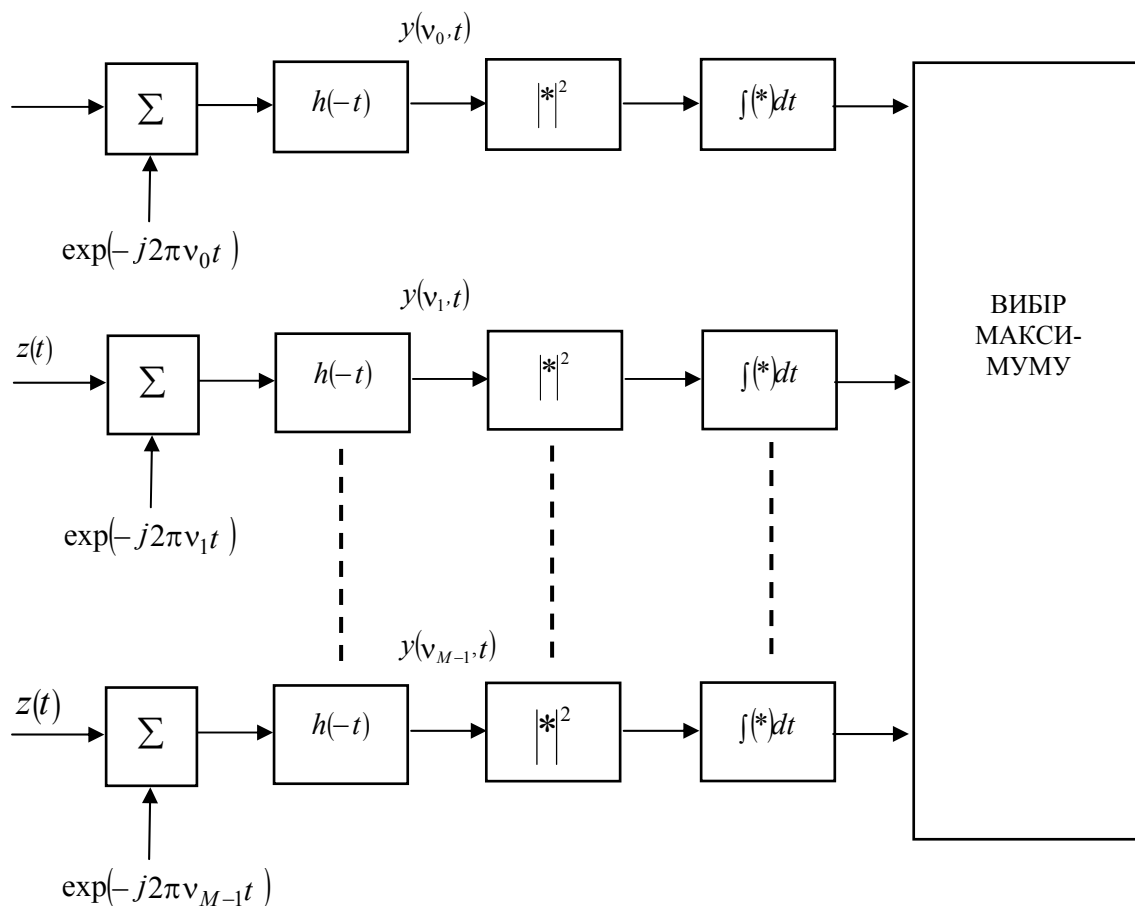


Рисунок 2 – Структурна схема паралельного обчислення МП-оцінки

На основі результатів, поданих в роботах [4, 5, 6], вираз для граничної дисперсії, нормованої до квадрату тактової частоти сигналу, що приймається може бути подано в вигляді:

$$\sigma_v^2 T^2 = \frac{2a}{\pi^2 K} \frac{1}{E_s/N_0} \left(1 + \frac{1}{E_s/N_0} \right) \quad (9)$$

5 РЕЗУЛЬТАТИ

Для ілюстрації співвідношення кордонів, отриманих на основі виразу (9), і залежності $CRLB(v) * T^2$ [4, 5], отриманої без врахування інших параметрів сигналу на рис. 3 подані залежності $\sigma_v^2 T^2$ та $CRLB(v) * T^2$ від співвідношення сигнал/шум на біт інформації E_s/N_0 для $K = 256$ (вид модуляції ФМ-4).

6 ОБГОВОРЕННЯ

Аналіз залежності, поданої на рис. 3, показує, що природно кордон $\sigma_v^2 T^2$ значно гірше, ніж нижній кордон Крамера-Рао. Це пояснюється тим, що при отриманні МП-оцінки несучої частоти при прийнятих вище початкових умовах не використана інформація про інші параметри сигналу, що приймається, а саме (φ, τ, d) .

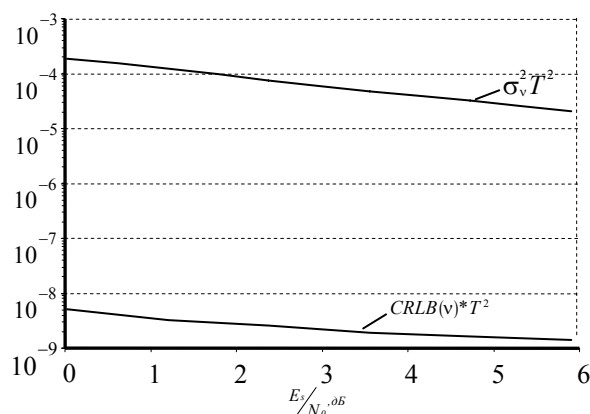


Рисунок 3 – Залежність нормованого кордону $CRLB(v) * T^2$ та $\sigma_v^2 T^2$ від співвідношення сигнал/шум, отримані на біт інформації для $K = 256$. Вид модуляції ФМ-4, $K = 256$

Величину $CRLB(v) * T^2$ в ряді робіт називають модифікованим кордоном Крамер-Рао – $MCRB(v)$ (Modified Cramer-Rao bound) [6, 22]. Величину σ_v^2 в деяких роботах найменують кордоном Крамера-Рао (Cramer-Rao bound) – $CRB(v)$ [22]. Кордон $CRB(v)$ – це кордон Крамера-Рао для дисперсії оцінки несучої

частоти сигналу, що приймається, коли ми маємо данні про всі інші параметри, крім $\tilde{\nu}$. А кордон $\sigma_{\tilde{\nu}}^2$ – кордон Крамера-Рао для дисперсії оцінки несучої частоти, коли ми не маємо даних про ці параметри. Таким чином, дисперсія будь якої «доброї» незміщеної оцінки частоти коливання фазомодульованого сигналу повинна лежати між цими кордонами і якомога ближче до $CRLB(\nu) * T^2$.

Слід зазначити, що такий підхід до визначення нижньої межі дисперсії оцінки і, загалом, до оцінки несучої частоти, певною мірою досить ефективно використовується для оцінки несучої частоти сигналу, що передається в пакетному режимі. Це обґрунтовано та представлено в роботі [23].

У вказаній роботі визначено функціональні залежності, на їх основі формується правило та подано алгоритм оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в пакетному режимі, за правилом максимальної правдоподібності за допомогою ковзного швидкого перетворення Фур'є. Цей алгоритм дає можливість оцінити частоту за правилом максимальної правдоподібності, враховуючи умови невизначеності всіх параметрів сигналу, що приймається супутниковою системою зв'язку в пакетному режимі, через короткі інтервали спостереження.

ВИСНОВКИ

В роботі визначені функціональні залежності та, на їх основі, розроблено алгоритм оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі по правилу максимальної правдоподібності.

Наукова новизна отриманих результатів полягає у тому, що в роботі вперше запропонований метод, та розроблений на його основі алгоритм, який дозволяє здійснити оцінку несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі по правилу максимальної правдоподібності з врахуванням умови невизначеності всіх параметрів сигналу.

Практична цінністю одержаних результатів роботи полягає у розробленому і поданому алгоритмі та запропонованій схемі його апаратної реалізації, яка дозволяє здійснити впровадження вказаного алгоритму в діючі системи супутникового зв'язку.

Для ілюстрації співвідношення кордонів мінімальної граничної дисперсії оцінки несучої частоти в роботі подані залежності, які дозволяють здійснити порівняння мінімальної граничної дисперсії, визначеної нижнім кордоном Крамера-Рао та мінімальної граничної дисперсії, визначеної з урахуванням всіх параметрів сигналу.

Аналіз вказаних залежностей показав, що в реальних умовах мінімальна дисперсія проведення МП-оцінки несучої частоти сигналу, що приймається супутниковою системою зв'язку в безперервному режимі

при невизначеності всіх параметрів сигналу може істотно відрізнитися від мінімальної дисперсії, отриманої на основі застосування нижнього кордону Крамера-Рао.

Перспективи подальших досліджень слід спрямувати на розробку та створення алгоритмів та методик, спрямованих на МП-оцінку при умові максимального наближення мінімальної граничної дисперсії оцінки несучої частоти реального сигналу до нижнього кордону Крамера-Рао, визначеного для оцінки несучої частоти в умовах визначеності інших параметрів сигналу.

ПОДЯКИ

Робота виконана в межах науково-дослідного проекту «Розробка методу підвищення якості функціонування космічних навігаційних радіолокаційних систем на основі застосування нелінійних адаптивних компенсаторів» (номер державної реєстрації №0117U001334), що забезпечується бюджетом Державного університету телекомунікацій.

ЛІТЕРАТУРА / LITERATURA

1. Горбатий І. В. Системи дистанційного зондування Землі з космосу : монографія / І. В. Горбатий. – Львов : СПОЛОМ, 2011. – 612 с.
2. Гришин Ю. П. Радиотехнические системы / Ю. П. Гришин, В. П. Ипатов, Ю. М. Казаринов. – М. : Высш. шк., 1990. – 496 с.
3. Левин Б. Р. Теоретические основы статистической радиотехники / Б. Р. Левин. – М. : Радио и связь, 1989. – 656 с.
4. D'Amico A. Cramer-Rao Bound for Clock Drift in UWB Ranging Systems / A. D'Amico, U. Mengali, L. Taponecco // IEEE Wireless Communication Letters. – 2013. – № 2(6). – P. 591–594. DOI: 10.1109/WCL.2013.080813.130424.
5. Khalesehossseini S. Generalized CRLB for DA and NDA Synchronization of UWB Signals with Clock Offset / S. Khalesehossseini, J. Nielsen // IEEE International Conference on Communications, 24–28 June 2007: proceedings. – Glasgow, Scotland, 2007. – Vol. 10. – P. 4305–4311. DOI: 10.1109/ICC.2007.710.
6. Wei H. Q. Influence of Phase on Cramer-Rao Lower Bounds for Joint Time Delay and Doppler Stretch Estimation / H. Wei, S. Ye, Q. Wan // 2007 9th International Symposium on Signal Processing and Its Applications (ISSPA-2007), 12–15 Feb. 2007: proceedings. – Sharjah, Piscataway, 2007. – P. 1–4. DOI: 10.1109/ISSPA.2007.4555336.
7. Zhao Tong Cramer-Rao Lower Bounds for the Joint Delay-Doppler Estimation of an Extended Target / Tong Zhao, Tianyao Huang // IEEE Global Conference on Signal and Information Processing (GlobalSIP), 14–16 Dec. 2015: proceedings. – Orlando, FL, USA, 2015. – P. 17–24. DOI: 10.1109/TSP.2015.2505681.
8. Nasir A. A. Particle filters for joint timing and carrier estimation: Improved resampling guidelines and weighted bayesian cramer-rao bounds / A. A. Nasir, S. Durrani, R. A. Kennedy // IEEE Transactions on Com-

- munications. – 2012. – № 60(5). – P. 1407–1419. DOI: 10.1109/TCOMM.2012.022912.100559.
9. Гоголев И. В. Граница Крамера-Рао оценки доплеровской деформации и задержки сигнала с произвольной шириной спектра / И. В. Гоголев // Известия высших учебных заведений России. Радиоэлектроника. – 2016. – №6 – С. 3–6.
 10. Nan Z. Cramer-Rao lower bounds of joint time delay and Doppler-stretch estimation with random stepped-frequency signals / Z. Nan, T. Zhao, T. Huang // 2016 IEEE International Conference on Digital Signal Processing (DSP), 2016: proceedings. – Beijing. – 2016. – P. 647–651. DOI: 10.1109/ICDSP.2016.7868638.
 11. Liguori C. Estimation of signal parameters in the frequency domain in the presence of harmonic interference: a comparative analysis / C. Liguori, A. Paolillo, A. Pignotti // IEEE Transactions on Instrumentation and Measurement. – 2006. – Vol. 55, Issue 2. – P. 562–569. DOI: 10.1109/TIM.2006.870116
 12. Pourhomayoun M. Cramer-Rao lower bound for frequency estimation for coherent pulse train with unknown pulse / M. Pourhomayoun, M. Fowler // IEEE Transactions on Aerospace and Electronic Systems. – 2014. – Vol. 50, № 2. – P. 1304–1312. DOI: 10.1109/TAES.2014.130024
 13. Analysis of Cramer-Rao lower bound for time delay estimation using UWB pulses / [S. Yun, S. Kim, J. Koh, J. Kang] // Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS). – 2012. – № 1. – P. 1–5. DOI: 10.1109/UPINLBS.2012.6409764.
 14. Tavares G. On the true cramer-rao lower bound for data-aided carrier-phase-independent frequency offset and symbol timing estimation / G. Tavares, L. Tavares, A. Petrolino // IEEE Transactions on Communications. – 2010. – № 58 (2). – P. 442–447. DOI: 10.1109/TCOMM.2010.02.090023
 15. Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid cramer-rao lower bound / [O. H. Salim, A. Nasir, H. Mehrpouyan et al.] // IEEE Transactions on Communications. – 2014. – No. 62 (9). – P. 3311–3325. DOI: 10.1109/TCOMM.2014.2345056
 16. Huang W. C. An investigation into the noise variance and the SNR estimators in imperfectly-synchronized OFDM systems / W. C. Huang, C. P. Li, H. J. Li // IEEE Transactions Wireless on Communications. – 2010. – № 9 (3). – P. 1159–1167. DOI: 10.1109/TWC.2010.03.090614
 17. Minimum variance of frequency estimations for a sinusoidal signal with low noise / [Y. Chibane, S. K. Lamoreaux, J. M. Pendlebury, K. F. Smith] // Measurement Science and Technology. – 1999. – Vol. 6, № 12. – P. 135–144. DOI: 10.1088/0957-0233/6/12/004
 18. Леман Э. Теория точечного оценивания / Э. Леман. – М. : Наука, 1991. – 448 с.
 19. Боровков А. А. Математическая статистика / А. А. Боровков. – Санкт-Петербург : Лань, 2010. – 704 с.
 20. Фомин А. И. Синхронизация цифровых радиосистем передачи информации / А. И. Фомин. – М. : Сайнс-Пресс, 2008. – 280 с.
 21. Стифлер Дж. Дж. Теория синхронной связи. / Дж. Дж. Стифлер. – М. : «Связь», 1975. – 484 с.
 22. Determination of the lower border of Cramer-Rao for evaluation of the carrier frequency of the radio technical communication channel signal / [O. Turovsky, L. Berkman, O. Drobyk, et al.] // International Journal of Advanced Trends in Computer Science and Engineering. – 2020. – Vol. 9, № 4. – P. 5838–5845. <https://doi.org/10.30534/ijatcse/2020/243942020>
 23. Estimates of the carrier frequency of the signal received by the satellite communication system in package mode / [O. Turovsky, O. Drobyk, A. Makarenko et al.] // International Journal of Advanced Trends in Computer Science and Engineering. – 2020. – Vol. 9, № 3. – P. 3223–3228. <https://doi.org/10.30534/ijatcse/2020/115932020>

Стаття надійшла до редакції 26.05.2020.
Після доробки 28.04.2021.

УДК 621.396

ПОРЯДОК ОЦЕНКИ НЕСУЩЕЙ ЧАСТОТЫ СИГНАЛА СПУТНИКОВОЙ СИСТЕМЫ СВЯЗИ В НЕПРЕРЫВНОМ РЕЖИМЕ

Туровский А. Л. – канд. техн. наук, доцент, профессор кафедры средств защиты информации Национального авиационного университета Киев, Украина.

Дробик А. В. – канд. техн. наук, профессор, директор Научного центра, Государственный университет телекоммуникаций, Киев, Украина.

АННОТАЦИЯ

Актуальность. Одной из особенностей систем спутниковой связи является преимущественное использование в них во время приему сигнала в непрерывном режиме фазовой модуляции сигналов, предназначенных для передачи полезной информации. Использование указанного вида модуляции требует решения задачи по оценке несущей частоты сигнала. А сама оценка сводится к задаче оценки частоты максимума в спектре фрагмента синусоидального сигнала на фоне аддитивного гауссовского шума. В статье рассмотрен процесс оценки несущей частоты сигнала спутниковой системой связи в непрерывном режиме по правилу максимального правдоподобия.

Цель. Разработка порядка оценки несущей частоты сигнала, принимаемого спутниковой системой связи в непрерывном режиме по правилу максимального правдоподобия.

Метод. Предложенный в работе порядок и разработанный на его основе алгоритм позволяет осуществить оценку несущей частоты по правилу максимального правдоподобия с учетом условий неопределенности всех параметров сигнала спутниковой системой связи в непрерывном режиме.

Результаты. С целью практического внедрения указанного алгоритма в действующие схемы спутниковой связи, в работе предложены схемы его аппаратной реализации. Для иллюстрации соотношения границ минимально предельной дисперсии оценки несущей частоты в работе представлены зависимости, которые позволяют осуществить сравнение минимально предельной дисперсии, определенной нижней границей Крамера-Рао и минимально предельной дисперсии, определенной с учетом всех параметров сигнала.

Выводы. Анализ представленных зависимостей показал, что в реальных условиях минимальная дисперсия проведения несущей частоты сигнала по правилу максимального правдоподобия, что принимается спутниковой системой связи в непрерывном режиме при неопределенности всех параметров сигнала может существенно отличаться от минимальной дисперсии, полученной на основе применения нижней границы Крамера-Рао. Перспективные дальнейших исследования следует направить на разработку и создание алгоритмов и методик, направленных на МП-оценку несущей частоты при условиях максимального приближения минимально предельной дисперсии оценки несущей частоты реального сигнала к нижней границе Крамера-Рао, определенного для оценки несущей частоты в условиях определенности других параметров сигнала.

КЛЮЧЕВЫЕ СЛОВА: принимаемый сигнал, оценка несущей частоты сигнала, минимально предельная дисперсия оценки несущей частоты, нижняя граница Крамера-Рао.

UDC 621.396

PROCEDURE FOR EVALUATION OF THE SUPPORTING FREQUENCY SIGNAL OF THE SATELLITE COMMUNICATION SYSTEM IN CONTINUOUS MODE

Turovsky A. L. – PhD, Associate Professor, Professor of the Department of Information Protection at the National Aviation University Kyiv, Ukraine.

Drobik O. V. – PhD, Professor, Director of the research center, State University of Telecommunications, Kyiv, Ukraine.

ABSTRACT

Context. One of the features of satellite communication systems is the advantageous use in them during the reception of the signal in the continuous mode of phase modulation of signals intended for the transmission of useful information. The use of this type of modulation requires solving the problem of estimating the carrier frequency of the signal. And the estimation itself is reduced to the problem of estimating the frequency of the maximum in the spectrum of a fragment of a sinusoidal signal against the background of additive Gaussian noise. The article considers the process of estimating the carrier frequency of a signal by a satellite communication system in a continuous mode according to the rule of maximum likelihood.

Objective. Development of a procedure for estimating the carrier frequency of a signal received by a satellite communication system in a continuous mode according to the maximum likelihood rule.

Method. The procedure proposed in the work and the algorithm developed on its basis allows to estimate the carrier frequency according to the rule of maximum likelihood, taking into account the conditions of uncertainty of all signal parameters by the satellite communication system in continuous mode.

The results. For the purpose of practical introduction of the specified algorithm in operating schemes of satellite communication, schemes of its hardware realization are offered in work. To illustrate the ratio of the limits of the minimum limiting variance of the carrier frequency estimate, the paper presents dependencies that allow comparing the minimum limiting variance defined by the lower Cramer-Rao boundary and the minimum limiting variance determined taking into account all signal parameters.

Conclusions. Analysis of these dependences showed that in real conditions the minimum dispersion of the carrier frequency of the signal according to the rule of maximum likelihood received by the satellite communication system in continuous mode with uncertainty of all signal parameters may differ significantly from the minimum dispersion obtained by applying the lower Kramer-Rao boundary. Prospective research, development and creation of algorithms and techniques aimed at estimating the carrier frequency at the minimum limiting variance in the conditions of uncertainty of all parameters of the received signal should be aimed at the maximum approximation of the minimum limiting variance of the estimated carrier frequency to the lower Cramer-Rao boundary to estimate the carrier frequency under conditions of certainty of other signal parameters.

KEYWORDS: the received signal, the signal carrier frequency estimate, the minimum limiting variance of the carrier frequency estimate, the lower Kramer-Rao boundary.

REFERENCES

1. Horbatyy I. V. *Sistemy dystantsiynoho zonduvannya Zemli z kosmosu: monohrafiya*. L'vov, SPOLOM, 2011, 612 p.
2. Grishin YU. P., Ipatov V. P., Kazarinov YU. M.. *Radio-tehnicheskiye sistemy*. Moscow, Vyssh. shk., 1990, 496 p.
3. Levin B. R. *Teoreticheskiye osnovy statisticheskoy radiotekhniki*. Moscow, Radio i svyaz', 1989, 656 p.

4. D'Amico A., Mengali U., Taponocco L. Cramer-Rao Bound for Clock Drift in UWB Ranging Systems, *IEEE Wireless Communication Letters*, 2013, №2 (6), pp. 591–594. DOI: 10.1109/WCL.2013.080813.130424.
5. Khalesehosseni S., Nielsen J. Generalized CRLB for DA and NDA Synchronization of UWB Signals with Clock Offset, *IEEE International Conference on Communications, 24–28 June 2007: proceedings*. Glasgow, Scotland, 2007, Vol. 10, pp. 4305–4311. DOI: 10.1109/ICC.2007.710.
6. Wei H. Q., Ye S., Wan Q. Influence of Phase on Cramer-Rao Lower Bounds for Joint Time Delay and Doppler Stretch Estimation, *2007 9th International Symposium on Signal Processing and Its Applications (ISSPA-2007), 12–15 Feb. 2007: proceedings*. Sharjah, Piscataway, 2007, pp. 1–4. DOI: 10.1109/ISSPA.2007.4555336.
7. Zhao Tong, Huang Tianyao Cramer-Rao Lower Bounds for the Joint Delay-Doppler Estimation of an Extended Target, *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 14–16 Dec. 2015: proceedings. Orlando, FL, USA, 2015, pp. 17–24. DOI: 10.1109/TSP.2015.2505681.
8. Nasir A. A., Durrani S., Kennedy R. A. Particle filters for joint timing and carrier estimation: Improved resampling guidelines and weighted bayesian cramer-rao bounds, *IEEE Transactions on Communications*, 2012, No. 60(5), pp. 1407–1419. DOI: 10.1109/TCOMM.2012.022912.100559.
9. Gogolev. I. V. Granica Kramera-Rao ocnki doplerovskoj deformacii i zaderzhki signala s proizvol'noj shirinoj spektra, *Izvestiya vysshix uchebnyx zavedenij Rossii. Radioelektronika*, 2016, No. 6, pp. 3–6.
10. Nan Z., Zhao T., Huang T. Cramer-Rao lower bounds of joint time delay and Doppler-stretch estimation with random stepped-frequency signals, *2016 IEEE International Conference on Digital Signal Processing (DSP), 2016: proceedings*. Beijing, 2016, pp. 647–651. DOI: 10.1109/ICDSP.2016.7868638.
11. Liguori C., Pignotti A. Estimation of signal parameters in the frequency domain in the presence of harmonic interference: a comparative analysis, *IEEE Transactions on Instrumentation and Measurement*, 2006, Vol. 55, Issue 2, pp. 562–569. DOI: 10.1109/TIM.2006.870116
12. Pourhomayoun M., Fowler M. Cramer-Rao lower bound for frequency estimation for coherent pulse train with unknown pulse, *IEEE Transactions on Aerospace and Electronic Systems*, 2014, Vol. 50, No. 2, pp. 1304–1312. DOI: 10.1109/TAES.2014.130024
13. Yun S., Kim S., Koh J., Kang J. Analysis of Cramer-Rao lower bound for time delay estimation using UWB pulses, *Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS)*, 2012, No. 1, pp. 1–5. DOI: 10.1109/UPINLBS.2012.6409764.
14. Tavares G., Tavares L., Petrolino A. On the true cramer-rao lower bound for data-aided carrier-phase-independent frequency offset and symbol timing estimation, *IEEE Trans. Commun*, 2010. No. 58(2), pp. 442–447. DOI: 10.1109/TCOMM.2010.02.090023
15. Salim O. H., Nasir A., Mehrpouyan H., et al. Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid cramer-rao lower bound, *IEEE Trans. Commun*, 2014, No.62(9), pp. 3311–3325. DOI: 10.1109/TCOMM.2014.2345056
16. Huang W. C., Li C. P., Li H. J. An investigation into the noise variance and the SNR estimators in imperfectly-synchronized OFDM systems, *IEEE Trans. Wireless Commun*, 2010, No. 9 (3), pp. 1159–1167. DOI: 10.1109/TWC.2010.03.090614
17. Chibane Y., Lamoreaux S. K., Pendlebury J. M., Smith K. F. Minimum variance of frequency estimations for a sinusoidal signal with low noise, *Measurement Science and Technology*, 1999, Vol. 6, No. 12, pp. 135–144. DOI: 10.1088/0957-0233/6/12/004.
18. Leman E. Teoriya tochechnogo otsenivaniya. Moscow, Nauka, 1991, 448 p.
19. Borovkov A. A. Matematicheskaya statistika. Sankt-Peterburg, Lan', 2010, 704 p.
20. Fomin A. I. Sinkhronizatsiya tsifrovyykh radiosistem peredachi informatsii. Moscow, Sayns-Press, 2008, 280 p.
21. Dzh. Dzh. Stifler. Teoriya sinkhronnoy svyazi. Moscow, «Svyaz'», 1975, 484 p.
22. Turovsky O., Berkman L., Drobyk O., et al. Determination of the lower border of Cramer-Rao for evaluation of the carrier frequency of the radio technical communication channel signal, *International Journal of Advanced Trends in Computer Science and Engineering*, 2020, Vol. 9, No. 4, pp. 5838–5845. <https://doi.org/10.30534/ijatcse/2020/243942020>
23. Turovsky O., Drobyk O., Makarenko A. et al. Estimates of the carrier frequency of the signal received by the satellite communication system in package mode, *International Journal of Advanced Trends in Computer Science and Engineering*, 2020, Vol. 9, No. 3, pp. 3223–3228. <https://doi.org/10.30534/ijatcse/2020/115932020>

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

MATHEMATICAL AND COMPUTER MODELING

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

UDC 51–74, 517.968.21

ON THE KOLMOGOROV-WIENER FILTER FOR RANDOM PROCESSES WITH A POWER-LAW STRUCTURE FUNCTION BASED ON THE WALSH FUNCTIONS

Gorev V. N. – PhD, Associate Professor of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.

Gusev A. Yu. – PhD, Associate Professor, Professor of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.

Korniienko V. I. – Dr. Sc., Professor, Head of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.

Safarov A. A. – PhD, Associate Professor of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.

ABSTRACT

Context. We investigate the Kolmogorov-Wiener filter weight function for the prediction of a continuous stationary random process with a power-law structure function.

Objective. The aim of the work is to develop an algorithm of obtaining an approximate solution for the weight function without recourse to numerical calculation of integrals.

Method. The weight function under consideration obeys the Wiener-Hopf integral equation. A search for an exact analytical solution for the corresponding integral equation meets difficulties, so an approximate solution for the weight function is sought in the framework of the Galerkin method on the basis of a truncated Walsh function series expansion.

Results. An algorithm of the weight function obtaining is developed. All the integrals are calculated analytically rather than numerically. Moreover, it is shown that the accuracy of the Walsh function approximations is significantly better than the accuracy of polynomial approximations obtained in the authors' previous papers. The Walsh function solutions are applicable in wider range of parameters than the polynomial ones.

Conclusions. An algorithm of obtaining the Kolmogorov-Wiener filter weight function for the prediction of a stationary continuous random process with a power-law structure function is developed. A truncated Walsh function expansion is the basis of the developed algorithm. In opposite to the polynomial solutions investigated in the previous papers, the developed algorithm has the following advantages. First of all, all the integrals are calculated analytically, and any numerical calculation of the integrals is not needed. Secondly, the problem of the product of very small and very large numbers is absent in the framework of the developed algorithm. In our opinion, this is the reason why the accuracy of the Walsh function solutions is better than that of the polynomial solutions for many approximations and why the Walsh function solutions are applicable in a wider range of parameters than the polynomial ones. The results of the paper may be applied, for example, to practical traffic prediction in telecommunication systems with data packet transfer.

KEYWORDS: Kolmogorov-Wiener filter weight function, continuous random process, Walsh functions, Galerkin method, power-law structure function.

NOMENCLATURE

T is a time interval on which the input process data are observed;

z is a time interval for which the forecast should be made;

$h(t)$ is a Kolmogorov-Wiener filter weight function;

H is a Hurst exponent;

σ^2 is a process variance;

α is a proportionality constant between the process structure function and the power-law term;

$H^{(2^m)}$ are the Hadamard matrices;

$W^{(2^m)}$ are the Walsh matrices;

$W_{kl}^{(2^m)}$ are the components of the Walsh matrices;

$wal_k(t)$ are the Walsh functions;
 g_m are the coefficients multiplying the Walsh functions;
 g is a column vector of the coefficients g_m ;
Left(t) is a left-hand side of the Wiener-Hopf integral equation
Right(t) is a right-hand side of the Wiener-Hopf integral equation;
 G_{mk} are the integral brackets;
 G is a matrix of the integral brackets;
 B_k are the free terms in the linear set of algebraic equations in g_m ;
 B is a column vector of the free terms B_k ;
 $V_{\beta\delta}$ are auxiliary integrals needed for the calculation of the integral brackets;
 V is a matrix of the integrals $V_{\beta\delta}$;
 X_β are auxiliary integrals needed for the calculation of the function Left(t);
 N is a number of points in the numerical integration.

INTRODUCTION

Random processes with a power-law structure function are widely used in different fields of knowledge. For example, they are used in plasma physics, statistical physics, in the study of financial markets (see corresponding references in [1]), in astrophysics [2, 3], in the description of turbulent flows [4] and so on. In particular, random processes with a power-law structure function may be used for the telecommunication traffic description (see, for example, [5–7]). The problem of telecommunication traffic prediction is very important for telecommunications because of its applications to power saving, optimal use of network resources, and detection of security attacks (see, for example, [8, 9]). There are plenty of telecommunication traffic models [10]. In the simplest models, telecommunication traffic in systems with data packet transfer is considered as a self-similar stationary random process [10]. One of the known telecommunication traffic models is the model where the traffic is considered to be a stationary random process with a power-law structure function [5]. In the case of a large amount of data, traffic may be treated as continuous random process [5].

As is known [11], the Kolmogorov-Wiener filter maybe applied to the prediction of stationary processes. The Kolmogorov-Wiener filter is a linear stationary filter, so this filter is rather simple and in our opinion it is quite natural to apply it to the prediction of stationary random process with a power-law structure function. There are plenty of rather complicated approaches to the prediction of random processes (see, for example, approaches to traffic prediction [8, 9]). However, as far as we know, a rather simple approach based on the use of the Kolmogorov-Wiener filter has not been sufficiently developed in

the literature, and so the results devoted to this approach have scientific novelty.

Our previous papers [5–7] were devoted to obtaining the Kolmogorov-Wiener filter weight function for the prediction of a random process with the power-law structure function. Papers [5–7] were based on the truncated polynomial expansion method, which has some drawbacks. To overcome these drawbacks, in the framework of this paper we propose to derive the unknown filter weight function on the basis of a truncated Walsh function expansion.

The object of study is the Kolmogorov-Wiener filter for the prediction of a random process with a power-law structure function.

The subject of study is the weight function of the corresponding filter.

The aim of the work is to obtain the weight function on the basis of expansion into a truncated Walsh function series.

1 PROBLEM STATEMENT

As is known [5], the Kolmogorov-Wiener weight function for the prediction of a random process with a power-law structure function is the solution of the following Wiener-Hopf integral equation:

$$\int_0^T d\tau h(\tau) \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) = \sigma^2 - \frac{\alpha}{2} (t + z)^{2H}. \quad (1)$$

The problem statement is as follows: to obtain the unknown filter weight function as an approximate solution of the integral equation (1) on the basis of a truncated Walsh function expansion.

2 REVIEW OF THE LITERATURE

Our previous papers [5–7] were devoted to a search for the unknown weight function on the basis of the integral equation (1) with the help of the truncated polynomial expansion method. In paper [5] the polynomials orthogonal without weight were used, and papers [6, 7] are based on the Chebyshev polynomials of the first and second kind, respectively. The behavior of polynomial solutions in [5–7] is almost identical.

It should be stressed that nowadays the truncated polynomial expansion method is rather popular for the solution of integral equation in different fields of knowledge (see, for example, [12–14]). However, in the framework of the problem under consideration polynomial solutions have several drawbacks. Some polynomial approximations in [5–7] indeed give good results, but some approximations absolutely fail. In our opinion, the fact that polynomial solutions lead to the product of very large and very small numbers may cause such failures. Moreover, the analytical calculation of the integral brackets in the framework of the polynomial method meets difficulties. Exact analytical expressions for the integral brackets may be obtained, but they are too cumbersome and, in fact,

they are not applicable in the framework of approximations of a rather large number of polynomials.

The truncated polynomial expansion method is a special case of the Galerkin method [15], in the framework of which the solution of an integral equation is sought in the form of a truncated orthogonal function series. As mentioned above, the use of orthogonal polynomials has some drawbacks. So, in order to avoid such drawbacks, we propose to use the orthogonal Walsh functions [16] instead of polynomials. The Walsh functions are step ones [16], thus allowing one to obtain analytical expressions for the integral brackets which are also applicable in the case of a rather large number of Walsh functions. Moreover, the use of the Walsh functions does not lead to the product of very large and very small numbers, so Walsh function solutions are applicable in a wider range of parameters than polynomial ones.

3 MATERIALS AND METHODS

As is known, the Hadamard matrices are introduced by the following recursive definition [16]:

$$H^{(2^{m+1})} = \begin{pmatrix} H^{(2^m)} & H^{(2^m)} \\ H^{(2^m)} & -H^{(2^m)} \end{pmatrix}, H^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (2)$$

The set of rows of the Walsh matrices coincides with that of the Hadamard matrices, but the numeration of the sets is not the same. The rows of the Walsh matrices are numerated as follows [16]: the smaller the number of sign changes in a row, the smaller the number of the row. The Walsh functions in the Walsh numeration are defined on the time interval $t \in [0, T]$ as follows:

$$\text{wal}_k(t) = \begin{cases} W_{k1}^{(2^m)}, t \in \left[0, \frac{T}{2^m}\right] \\ W_{k2}^{(2^m)}, t \in \left[\frac{T}{2^m}, \frac{2T}{2^m}\right] \\ \vdots \\ W_{k,2^m}^{(2^m)}, t \in \left[\frac{(2^m-1)T}{2^m}, T\right] \end{cases}. \quad (3)$$

where m is the least natural number for which the inequality $2^m \geq k$ holds. As can be seen, the Walsh functions (3) are step ones, they may take only the values $+1$ or -1 . The Walsh functions form a complete orthogonal function system on the time interval $t \in [0, T]$.

The unknown weight function $h(\tau)$ in the approximation of $n = 2^m$ is sought in the form

$$h(\tau) = \sum_{m=1}^n g_m \text{wal}_m(\tau). \quad (4)$$

On substituting (4) into (1) one can obtain

$$\sum_{m=1}^n g_m \int_0^T d\tau \text{wal}_m(\tau) \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) = \sigma^2 - \frac{\alpha}{2} (t+z)^{2H}, \quad (5)$$

which after multiplying by $\text{wal}_k(t)$ and integrating over t leads to the following system of linear equations:

$$\sum_{m=1}^n G_{mk} g_m = B_k, \quad k = \overline{1, n} \quad (6)$$

where

$$G_{mk} = \int_0^T \int_0^T dt d\tau \text{wal}_m(\tau) \text{wal}_k(t) \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) \quad (7)$$

and

$$B_k = \int_0^T dt \text{wal}_k(t) \left(\sigma^2 - \frac{\alpha}{2} (t+z)^{2H} \right). \quad (8)$$

The matrix form of the system (6) is as follows:

$$Gg = B, \quad (9)$$

where

$$G = \begin{pmatrix} G_{11} & G_{12} & \cdots & G_{1n} \\ G_{21} & G_{22} & \cdots & G_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n1} & G_{n2} & \cdots & G_{nn} \end{pmatrix}, B = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{pmatrix}, g = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix}. \quad (10)$$

The Walsh functions obey the property

$$\text{wal}_k\left(\frac{T}{2} - t\right) = \begin{cases} \text{wal}_k\left(\frac{T}{2} + t\right), k \text{ : 2} \\ -\text{wal}_k\left(\frac{T}{2} + t\right), k \not\text{ : 2} \end{cases}, \quad (11)$$

which with account for the fact that the absolute value is sign-independent leads to the following properties of the integral brackets:

$$G_{ks} = G_{sk}; \quad G_{ks} = 0 \text{ if } k, s \text{ are of opposite parities}, \quad (12)$$

see a similar derivation of (12) in [5]. So, a straightforward calculation is needed only for the integral brackets G_{km} where $k \geq m$ and k, m are of the same parity. Analytical expressions for the integral brackets may be obtained on the basis of the fact that the Walsh functions are step ones:

$$G_{mk} = \sum_{\beta, \delta=1}^n W_{m\beta}^{(n)} W_{k\delta}^{(n)} V_{\beta\delta} \quad (13)$$

where

$$V_{\beta\delta} = \int_{(\beta-1)T/n}^{\beta T/n} \int_{(\delta-1)T/n}^{\delta T/n} dt d\tau \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right). \quad (14)$$

The integrals $V_{\beta\delta}$ obey the properties

$$V_{\beta\delta} = V_{\beta+1, \delta+1}; \quad V_{\beta\beta} = V_{\delta\beta}. \quad (15)$$

This can be demonstrated as follows:

$$\begin{aligned} V_{\beta\delta} &= \int_{(\beta-1)T/n}^{\beta T/n} \int_{(\delta-1)T/n}^{\delta T/n} dt d\tau \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) = \\ &= \left\{ x = t + \frac{T}{n}, y = \tau + \frac{T}{n} \right\} = \\ &= \int_{\beta T/n}^{(\beta+1)T/n} \int_{\delta T/n}^{(\delta+1)T/n} dx dy \left(\sigma^2 - \frac{\alpha}{2} |x - y|^{2H} \right) = V_{\beta+1, \delta+1}. \end{aligned} \quad (16)$$

and

$$\begin{aligned} V_{\beta\delta} &= \int_{(\beta-1)T/n}^{\beta T/n} \int_{(\delta-1)T/n}^{\delta T/n} dt d\tau \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) = \\ &= \left\{ t \leftrightarrow \tau, |t - \tau| = |\tau - t| \right\} = \\ &= \int_{(\delta-1)T/n}^{\delta T/n} \int_{(\beta-1)T/n}^{\beta T/n} dt d\tau \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) = V_{\delta\beta}. \end{aligned} \quad (17)$$

So the matrix V takes the form

$$\begin{aligned} V &= \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1n} \\ V_{21} & V_{22} & \dots & V_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ V_{n1} & V_{n2} & \dots & V_{nn} \end{pmatrix} = \\ &= \begin{pmatrix} V_{11} & V_{12} & V_{13} & \dots & V_{1n} \\ V_{12} & V_{11} & V_{12} & \dots & V_{1,n-1} \\ V_{13} & V_{12} & V_{11} & \dots & V_{1,n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ V_{1n} & V_{1,n-1} & V_{1,n-2} & \dots & V_{11} \end{pmatrix}. \end{aligned} \quad (18)$$

and a straightforward calculation is needed only for the first row in (18). Let us obtain analytical results for the quantities $V_{1\beta}$. First of all, from (14) one can see that

$$\begin{aligned} V_{1\beta} &= \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2} \int_0^{\frac{T}{n}} dt \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} d\tau |t - \tau|^{2H} = \\ &= \left\{ y = t - \tau \right\} = \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2} \int_0^{\frac{T}{n}} dt \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} dy |y|^{2H}. \end{aligned} \quad (19)$$

For the quantity V_{11} the integrand is of alternating sign and

© Gorev V. N., Gusev A. Yu., Korniienko V. I., Safarov A. A., 2021
 DOI 10.15588/1607-3274-2021-2-4

$$\begin{aligned} V_{11} &= \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2} \int_0^{\frac{T}{n}} dt \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} dy |y|^{2H} = \\ &= \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2} \int_0^{\frac{T}{n}} dt \left(\int_0^0 dy (-y)^{2H} + \int_0^{\frac{T}{n}} dy y^{2H} \right) = \\ &= \left\{ x = -y \right\} = \\ &= \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2} \int_0^{\frac{T}{n}} dt \left(\int_0^{\frac{T-t}{n}} dx x^{2H} + \int_0^{\frac{T-t}{n}} dy y^{2H} \right) \end{aligned} \quad (20)$$

which after a straightforward calculation leads to

$$V_{11} = \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{(2H+1)(2H+2)} \left(\frac{T}{n} \right)^{2H+2}. \quad (21)$$

For $\beta > 1$ the integrand in (19) does not change the sign:

$$V_{1\beta} = \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2} \int_0^{\frac{T}{n}} dt \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} dy (-y)^{2H}. \quad (22)$$

which after a straightforward calculation leads to

$$\begin{aligned} V_{1\beta} &= \left(\sigma \frac{T}{n} \right)^2 - \frac{\alpha}{2(2H+1)(2H+2)} \left[\left(\frac{\beta T}{n} \right)^{2H+2} + \right. \\ &\quad \left. + \left(\frac{(\beta-2)T}{n} \right)^{2H+2} - 2 \left(\frac{(\beta-1)T}{n} \right)^{2H+2} \right]. \end{aligned} \quad (23)$$

The free terms B_k may be calculated by a similar idea:

$$B_k = \sum_{\beta=1}^n W_{k\beta}^{(n)} \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} dt \left(\sigma^2 - \frac{\alpha}{2} (t+z)^{2H} \right) \quad (24)$$

which after a straightforward calculation leads to

$$\begin{aligned} B_k &= \sum_{\beta=1}^n W_{k\beta}^{(n)} \left\{ \sigma^2 \frac{T}{n} - \frac{\alpha}{2(2H+1)} \left[\left(\frac{\beta T}{n} + z \right)^{2H+1} - \right. \right. \\ &\quad \left. \left. - \left(\frac{(\beta-1)T}{n} + z \right)^{2H+1} \right] \right\}. \end{aligned} \quad (25)$$

So the algorithm of obtaining the coefficients multiplying the Walsh functions may be formulated as follows:

1. Calculate V_{11} by formula (21).
2. Calculate $V_{1\beta}$, $\beta \geq 2$ by formula (23).
3. Form the matrix V by formula (18).
4. Calculate the integral brackets G_{km} by formula (13) where $k \geq m$ and k, m are of the same parity

5. Calculate all the other integral brackets on the basis of the properties (12) and form the matrix G of the integral brackets (see (10)).

6. Calculate the free terms B_k by formula (25) and form the column vector B of the free terms

7. Calculate the column vector g of the unknown coefficients by the following formula (26):

$$g = G^{-1}B. \quad (26)$$

In order to verify the obtained solutions, the left-hand side and the right-hand side of (1) should be numerically compared for the obtained solutions:

$$\begin{aligned} \text{Left}(t) &= \int_0^T d\tau h(\tau) \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right), \\ \text{Right}(t) &= \sigma^2 - \frac{\alpha}{2} (t + z)^{2H}. \end{aligned} \quad (27)$$

The obtained weight function $h(\tau)$ is a step one, so the function $\text{Left}(t)$ may be calculated as

$$\text{Left}(t) = \sum_{\beta=1}^n h_{\beta} X_{\beta}(t) \quad (28)$$

where

$$h_{\beta} = h\left(\frac{1}{2}\left(\frac{(\beta-1)T}{n} + \frac{\beta T}{n}\right)\right) = h\left(\frac{(2\beta-1)T}{2n}\right), \quad (29)$$

and

$$\begin{aligned} X_{\beta}(t) &= \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} d\tau \left(\sigma^2 - \frac{\alpha}{2} |t - \tau|^{2H} \right) = \\ &= \sigma^2 \frac{T}{n} - \frac{\alpha}{2} \int_{\frac{(\beta-1)T}{n}}^{\frac{\beta T}{n}} d\tau |t - \tau|^{2H} \end{aligned} \quad (30)$$

which yields

$$X_{\beta}(t) = \begin{cases} \sigma^2 \frac{T}{n} - \frac{\alpha}{2(2H+1)} \left[\left(t - \frac{(\beta-1)T}{n} \right)^{2H+1} - \left(t - \frac{\beta T}{n} \right)^{2H+1} \right], & t > \frac{\beta T}{n} \\ \sigma^2 \frac{T}{n} - \frac{\alpha}{2(2H+1)} \left[\left(\frac{\beta T}{n} - t \right)^{2H+1} - \left(\frac{(\beta-1)T}{n} - t \right)^{2H+1} \right], & t < \frac{(\beta-1)T}{n} \\ \sigma^2 \frac{T}{n} - \frac{\alpha}{2(2H+1)} \left[\left(\frac{\beta T}{n} - t \right)^{2H+1} + \left(t - \frac{(\beta-1)T}{n} \right)^{2H+1} \right], & \frac{(\beta-1)T}{n} \leq t \leq \frac{\beta T}{n} \end{cases} \quad (31)$$

The mean average error (MAE) may be calculated in order to estimate the accuracy of the obtained solutions:

$$\text{MAE} = \frac{1}{T} \int_0^T |\text{Left}(t) - \text{Right}(t)|. \quad (32)$$

The integral (32) may be numerically calculated, for example, as follows:

$$\begin{aligned} \text{MAE} &\approx \frac{1}{T} \sum_{j=1}^N \left| \text{Left}\left(\frac{(2j-1)T}{2N}\right) - \text{Right}\left(\frac{(2j-1)T}{2N}\right) \right| \cdot \frac{T}{N} = \\ &= \frac{1}{N} \sum_{j=1}^N \left| \text{Left}\left(\frac{(2j-1)T}{2N}\right) - \text{Right}\left(\frac{(2j-1)T}{2N}\right) \right|. \end{aligned} \quad (33)$$

In this paper, the value $N = 10^4$ is used.

So the algorithm of the MAE obtaining is as follows:

1. Calculate the quantities h_{β} by formulas (29).
2. Introduce the functions $X_{\beta}(t)$ by formulas (30).
3. Introduce the function $\text{Left}(t)$ by formula (28).
4. Introduce the function $\text{Right}(t)$ by formula (27).
5. Calculate the MAE by formula (33).

It should be stressed that the only integral that should be numerically calculated is the integral (33) for the MAE calculation. For all the other integrals, applicable analytical expressions are obtained.

4 EXPERIMENTS

The numerical calculations are made in the Wolfram Mathematica package. The results for the following set of parameters are investigated:

$$T = 100, \quad z = 3, \quad H = 0.8, \quad \sigma = 1.2, \quad \alpha = 3 \cdot 10^{-3}. \quad (34)$$

The results for the polynomial solutions for the parameters (34) were investigated in [5–7]. It was shown that the approximations of a number of polynomials from 9 to 15 absolutely fail (the corresponding MAE is greater than 10^2 , the graphs of $\text{Left}(t)$ and $\text{Right}(t)$ are totally different). On the basis of the Walsh functions the following MAE are obtained, see Table 1.

As can be seen, the approximations of $n = 2^m$ Walsh functions are rather accurate, and the accuracy increases with m . For graphical visualization, the graphs of $\text{Left}(t)$ and $\text{Right}(t)$ are given for the approximation of 256 Walsh functions, see Fig. 1.

As can be seen from Fig. 1, an almost ideal coincidence of the curves takes place.

Table 1 – MAE for the parameters (34) for the approximations of n Walsh functions rounded off to two significant digits

n	MAE
2	$3.2 \cdot 10^{-2}$
4	$2.4 \cdot 10^{-2}$
8	$9.2 \cdot 10^{-3}$
16	$7.4 \cdot 10^{-4}$
32	$1.1 \cdot 10^{-4}$
64	$2.0 \cdot 10^{-5}$
128	$3.8 \cdot 10^{-6}$
256	$7.6 \cdot 10^{-7}$

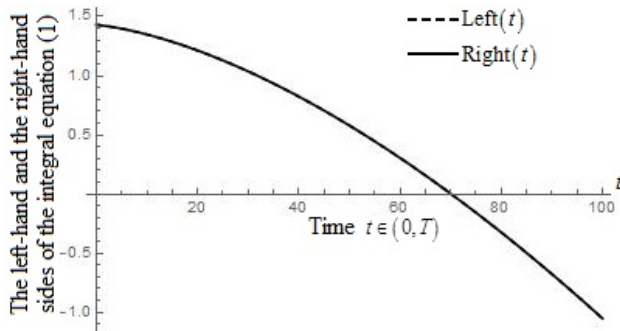


Figure 1 – Comparison of the left-hand and right-hand sides of eq. (1) for parameters (34) for the approximation of 256 Walsh functions.

The following set of parameters is also investigated:

$$T = 1000, z = 3, H = 0.8, \sigma = 1.2, \alpha = 8 \cdot 10^{-5}. \quad (35)$$

As indicated in [5–7], in fact only the approximation of two polynomials is valid for the parameters (35). As for the Walsh solutions, we have the following, see Table 2:

Table 2 – MAE for the parameters (35) for the approximations of n Walsh functions rounded off to two significant digits

n	MAE
2	$3.2 \cdot 10^{-2}$
4	$3.8 \cdot 10^{-2}$
8	$4.1 \cdot 10^{-3}$
16	$4.1 \cdot 10^{-4}$
32	$6.0 \cdot 10^{-5}$
64	$9.8 \cdot 10^{-6}$
128	$1.7 \cdot 10^{-6}$
256	$3.2 \cdot 10^{-7}$

So the solutions based on the Walsh functions give good results for the parameters (35). For graphical visualization, the graphs of $Left(t)$ and $Right(t)$ are given for the approximation of 256 Walsh functions, see Fig. 2. As can be seen from Fig. 2, an almost ideal coincidence of the curves takes place.

It should also be stressed that even for the parameters (34) the Wolfram Mathematica is not able to build a graph for $Left(t)$ for a number of polynomials greater than 18.

The method based on the Walsh functions allows one to treat several hundreds of functions.

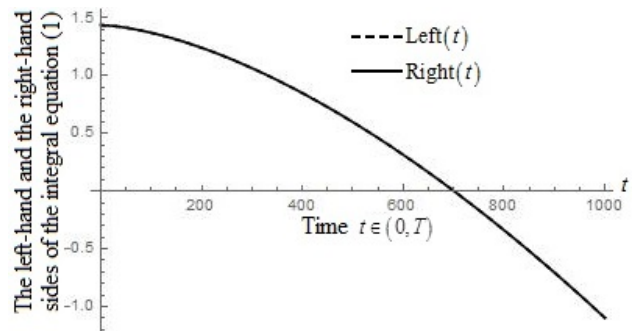


Figure 2 – Comparison of the left-hand and right-hand sides of eq. (1) for parameters (35) for the approximation of 256 Walsh functions

So one can conclude that the method based on a truncated Walsh function expansion works very well and the accuracy of the approximations of rather high numbers of Walsh functions are very accurate.

5 RESULTS

The Kolmogorov-Wiener filter weight function for the prediction of a random process with a power-law structure function is investigated. The investigation is based on the Galerkin method, in the framework of which the unknown weight function is sought in the form of a truncated Walsh function series. An algorithm of the weight function derivation is developed, this algorithm does not require a numerical calculation of the integrals.

In order to verify the results, the mean average error (MAE) of the residual (the difference of the left-hand side and the right-hand side) of the integral equation (1) is introduced. The sets of parameters (34) and (35) are investigated (the corresponding sets were also investigated for polynomial solutions [5–7]). It is shown that the Walsh function approximations give good results and the approximations of a rather large number of the Walsh functions are very accurate (they lead to a very small MAE).

It is found that the Walsh function solutions are much better than the polynomial ones in the framework of the problem under consideration.

6 DISCUSSION

In this paper we propose to realize the Galerkin method for obtaining the weight function on the basis of a truncated Walsh function expansion. The corresponding algorithm of the weight function derivation is developed. In contrast to the truncated polynomial expansion method, the proposed algorithm has the following advantages:

1. The numerical calculation of the integrals is not needed for the weight function obtaining (the only numerical calculation of the integral is needed for MAE obtaining, not for the weight function obtaining). In particular, applicable analytical expressions for the integral brackets are obtained.

2. The calculation of the left-hand side of the Wiener-Hopf integral equation does not require numerical calculation of the integral, the corresponding analytical expression for the function $\text{Left}(t)$ is obtained.

3. The use of the Walsh functions does not lead to the products of very small and very large numbers, which can hardly be calculated numerically.

It is shown that for the same parameters the solutions based on the Walsh functions are much better than the polynomial ones. They are applicable in a wider range of parameters, and they lead to far smaller values of the MAE (the MAE for the 18-polynomial approximation for the parameters (34) is of the order 10^{-3} , the MAE for the 256 Walsh function approximation is of the order 10^{-7}). The use of the Walsh functions allows one to numerically treat the approximations of several hundreds of Walsh functions (in contrast to only several tens of polynomials). Moreover, there are no Walsh function approximations that absolutely fail, in contrast to polynomial ones.

CONCLUSIONS

The Kolmogorov-Wiener filter weight function for the prediction of a random process with a power-law structure function is investigated on the basis of a truncated Walsh function expansion. It is shown that in the framework of the problem under consideration the Walsh function solution is much better than the polynomial solutions investigated in the authors' previous papers.

Random processes with a power-law structure function are widely used indifferent fields of knowledge (in particular, for the telecommunication traffic description). So the results of this paper may be useful for practical prediction of stationary random processes with a power-law structure function in various fields of knowledge (in particular, for telecommunication traffic prediction in systems with data packet transfer).

The scientific novelty of the paper is the fact that for the first time the weight function under consideration is found on the basis of a truncated Walsh function expansion. The proposed algorithm of the weight function derivation does not lead either to numerical calculation of integrals or to the product of very large and very small numbers. It is shown that for the problem under consideration the Walsh function solutions are much better than the polynomial ones.

The practical significance is that the obtained results may be applied for the practical prediction of stationary random processes with a power-law structure function in various fields of knowledge.

Prospects for further research are to obtain a practical prediction on the basis of the obtained results.

ACKNOWLEDGEMENTS

The work is a continuation of the research made in the framework of the State Budgetary Research and Development Project GP-458 "Intelligent Control Technologies of the Mining Processes in Problems of Energy Saving

and Energy Efficiency" (State Registration No. 0113U000402) of the Dnipro University of Technology.

REFERENCES

1. Gorev V. N., Gusev A. Yu., V. I. Korniienko On the analytical solution of a Volterra integral equation for investigation of fractal processes, *Radio Electronics, Computer Science, Control*, 2018, No. 4, pp. 42–50. DOI: 10.15588/1607-3274-2018-4-4
2. Akiyama K., Johnson M. D. Interstellar scintillation and the radio counterpart of the fast radio burst FRB 150418, *The Astrophysical Journal Letters*, 2016, Vol. 824, No. 1, L3 (5 pages). DOI: 10.3847/2041-8205/824/1/L3.
3. Micelotta E. R., Dwek E. and Slavin J. D. Dust destruction by the reverse shock in the Cassiopeia A supernova remnant, *Astronomy & Astrophysics*, 2016, Vol. 590, A65 (20 pages). DOI: 10.1051/0004-6361/201527350
4. Pope S. B. Turbulent Flows. Cambridge: Cambridge University Press, 2000, 777 p. DOI: 10.1017/CBO9780511840531
5. Gorev V. N., Gusev A. Yu., Korniienko V. I. Polynomial solutions for the Kolmogorov-Wiener filter weight function for fractal processes, *Radio Electronics, Computer Science, Control*, 2019, No. 2, pp. 44–52. DOI: 10.15588/1607-3274-2019-2-5
6. Gorev V., Gusev A., Korniienko V. Investigation of the Kolmogorov-Wiener filter for continuous fractal processes on the basis of the Chebyshev polynomials of the first kind, *IAPGOS*, 2020, No. 1, pp. 58–61. DOI: 10.35784/iapgos.912
7. Gorev V., Gusev A., Korniienko V. Investigation of the Kolmogorov-Wiener filter for treatment of fractal processes on the basis of the Chebyshev polynomials of the second kind, *Ceur Workshop Proceedings*, 2019, Vol. 2353, pp. 596–606.
8. Iqbal M. F., Zahid M., Habib D., John L. K. Efficient Prediction of Network Traffic for Real-Time Applications, *Journal of Computer Networks and Communications*, 2019, Vol. 2019, 4067135 (11 pages). DOI: 10.1155/2019/4067135.
9. Katris C., Daskalaki S. Comparing forecasting approaches for Internet traffic, *Expert Systems with Applications*, 2015, Vol. 42, Issue 21, pp. 8172–8183. DOI: 10.1016/j.eswa.2015.06.029.
10. Al-Azzeh J. S., Al Hadidi M., Odarchenko R., Gnatyuk S., Z. Shevchuk, Zhengbing Hu Analysis of Self-Similar Traffic Models in Computer Networks, *International Review on Modelling and Simulations*, 2017, Vol. 10, No. 5, pp. 328–336. DOI: 10.15866/iremos.v10i5.12009
11. Miller S., Childers D. Probability and Random Processes With Applications to Signal Processing and Communications. Second edition. Amsterdam, Elsevier/Academic Press, 2012, 598 p. DOI: doi.org/10.1016/B978-0-12-386981-4.50001-1
12. Sokolovsky A. I., Sokolovsky S. A., Hrinishyn O. A. On relaxation processes in a completely ionized plasma, *East European Journal of Physics*, 2020, No. 3, pp. 19–30. DOI: 10.26565/2312-4334-2020-3-03
13. Gonzalez R. G., Khalil N., Garzo V. Enskog kinetic theory for multicomponent granular suspensions, *Physical Review E*, 2020, Vol. 101, 012904 (23 pages). DOI: 10.1103/PhysRevE.101.012904
14. Wargnier Q., Laguna A. A., Scoggins J. B., Mansour N. N., Massot M., Magin T. E. Consistent transport properties in multicomponent two-temperature magnetized plasmas. Ap-

- plication to the Sun atmosphere, *Astronomy and Astrophysics*, 2020, Vol. 635, A87 (15 pages). DOI: 10.1051/0004-6361/201834686
15. Polyaniin A. D., Manzhairov A. V. Handbook of integral equations. Second edition. Boca Raton, Chapman & Hall/CRC Press, Taylor & Francis Group, 2008, 1144 p. DOI: 10.1201/9781420010558
16. Lindfield G., Penny J. Numerical Methods Using MATLAB. Fourth edition. Amsterdam, Elsevier, Academic Press, 2019, 608 p. DOI: 10.1016/C2016-0-00395-9

Received 31.03.2021.
Accepted 12.05.2021.

УДК 51–74, 517.968.21

ДО ФІЛЬТРУ КОЛМОГОВОРА-ВІНЕРА ДЛЯ ВИПАДКОВОГО ПРОЦЕСУ ЗІ СТЕПЕНЕВОЮ СТРУКТУРНОЮ ФУНКЦІЄЮ НА ОСНОВІ ФУНКЦІЙ УОЛША

Горєв В. М. – канд. фіз.-мат. наук, доцент кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

Гусєв О. Ю. – канд. фіз.-мат. наук, доцент, професор кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

Корнієнко В. І. – д-р техн. наук, професор, завідувач кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

Сафаров О. О. – канд. техн. наук, доцент кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

АНОТАЦІЯ

Актуальність. Розглянуто вагову функцію фільтра Колмогорова-Вінера для прогнозування неперервного стаціонарного випадкового процесу зі степеневу структурною функцією.

Мета роботи. Метою роботи є розробити алгоритм отримання наближеного розв'язку для вагової функції, який не містить числового обчислення інтегралів.

Метод. Вагова функція, що розглядається, підпорядковується інтегральному рівнянню Вінера-Хопфа. Пошук точного аналітичного розв'язку відповідного інтегрального рівняння стикається з труднощами, тож шукається наближений розв'язок для вагової функції в рамках методу Галеркіна, який базується на основі обірваного розвинення в ряд за функціями Уолша.

Результати. Розроблено відповідний алгоритм отримання вагової функції. Усі інтеграли обчислено аналітично, а не чисельно. Більше того, показано, що точність отриманих наближень, що базуються на функціях Уолша, є значно кращою за точність поліноміальних розв'язків, отриманих у попередніх роботах авторів. Розв'язки, що базуються на функціях Уолша, є застосовними у ширшому діапазоні параметрів, ніж поліноміальні розв'язки.

Висновки. Розроблено алгоритм отримання вагової функції фільтра Колмогорова-Вінера для прогнозування неперервного стаціонарного випадкового процесу зі степеневу структурною функцією. Основою алгоритму є розвинення за функціями Уолша. На відміну від поліноміальних розв'язків, досліджених у минулих статтях, розроблений алгоритм має наступні переваги. По-перше, усі інтеграли обчислено аналітично, і немає потреби в числовому розрахунку інтегралів. По-друге, проблема добутку дуже малих та дуже великих чисел відсутня в рамках запропонованого алгоритму. На наш погляд, це є причиною того, що точність розв'язків, що базуються на функціях Уолша, є кращою за точність поліноміальних розв'язків для багатьох наближень, і це є причиною того, що розв'язки на основі функцій Уолша є застосовними у ширшому діапазоні параметрів, ніж поліноміальні розв'язки. Результати роботи можуть бути застосовані до, наприклад, прогнозування на практиці трафіку в телекомунікаційних системах з пакетною передачею даних.

КЛЮЧОВІ СЛОВА: вагова функція фільтра Колмогорова-Вінера, неперервний випадковий процес, функції Уолша, метод Галеркіна, степенева структурна функція.

УДК 51–74, 517.968.21

К ФИЛЬТРУ КОЛМОГОВОРА-ВИНЕРА ДЛЯ СЛУЧАЙНОГО ПРОЦЕССА СО СТЕПЕННОЙ СТРУКТУРНОЙ ФУНКЦИЕЙ НА ОСНОВЕ ФУНКЦИЙ УОЛША

Горєв В. Н. – канд. физ.-мат. наук, доцент кафедры безопасности информации и телекоммуникаций, Национальный технический университет «Днепропетровская Политехника», Днепр, Украина.

Гусєв А. Ю. – канд. физ.-мат. наук, доцент, профессор кафедры безопасности информации и телекоммуникаций, Национальный технический университет «Днепропетровская Политехника», Днепр, Украина.

Корнієнко В. І. – д-р техн., профессор, заведующий кафедры безопасности информации и телекоммуникаций, Национальный технический университет «Днепропетровская Политехника», Днепр, Украина.

Сафаров А. А. – канд. техн. наук, доцент кафедры безопасности информации и телекоммуникаций, Национальный технический университет «Днепропетровская Политехника», Днепр, Украина.

АННОТАЦИЯ

Актуальность. Рассмотрена весовая функция фильтра Колмогорова-Винера для прогнозирования непрерывного стационарного случайного процесса со степенной структурной функцией.

Цель работы. Целью работы является разработать алгоритм получения приближенного решения для весовой функции, который не содержит численного вычисления интегралов.

Метод. Рассматриваемая весовая функция, подчиняется интегральному уравнению Винера-Хопфа. Поиск точного аналитического решения соответствующего интегрального уравнения затруднен, так что ищется приближенное решение для весовой функции в рамках метода Галеркина, основываемого на оборванном разложении в ряд по функциям Уолша.

Результаты. Разработан соответствующий алгоритм получения весовой функции. Все интегралы вычислены аналитически, а не численно. Более того, показано, что точность полученных приближений, базируемых на функциях Уолша, значительно лучше, чем точность полиномиальных решений, полученных в предыдущих работах авторов. Решения, которые базируются на функциях Уолша, применимы в более широком диапазоне параметров, чем полиномиальные решения.

Выводы. Разработан алгоритм получения весовой функции фильтра Колмогорова-Винера для прогнозирования стационарного непрерывного случайного процесса со степенной структурной функцией. Основным алгоритма есть разложение по функциям Уолша. В отличие от полиномиальных решений, исследованных в предыдущих статьях, разработанный алгоритм имеет следующие преимущества. Во-первых, все интегралы вычислены аналитически, и нет необходимости в численном вычислении интегралов. Во-вторых, проблема произведения очень малых и очень больших чисел отсутствует в рамках предложенного алгоритма. На наш взгляд, это является причиной того, что точность решений, основывающихся на функциях Уолша, лучше, чем точность полиномиальных решений, и это является причиной того, что решения на основе функций Уолша применимы в более широком диапазоне параметров, чем полиномиальные решения. Результаты работы могут быть применимы к, например, прогнозированию на практике трафика в телекоммуникационных системах с пакетной передачей данных.

КЛЮЧЕВЫЕ СЛОВА: весовая функция фильтра Колмогорова-Винера, непрерывный случайный процесс, функции Уолша, метод Галеркина, степенная структурная функция.

ЛИТЕРАТУРА / LITERATURA

1. Gorev V. N. On the analytical solution of a Volterra integral equation for investigation of fractal processes / V. N. Gorev, A. Yu. Gusev, V. I. Korniienko // Radio Electronics, Computer Science, Control. – 2018. – No. 4. – P. 42–50. DOI: 10.15588/1607-3274-2018-4-4
2. Akiyama K. Interstellar scintillation and the radio counterpart of the fast radio burst FRB 150418 / K. Akiyama, M. D. Johnson // The Astrophysical Journal Letters. – 2016. – Vol. 824, No. 1. – L3 (5 pages). DOI: 10.3847/2041-8205/824/1/L3.
3. Micelotta E. R. Dust destruction by the reverse shock in the Cassiopeia A supernova remnant / E. R. Micelotta, E. Dwek and J. D. Slavin // Astronomy & Astrophysics. – 2016. – Vol. 590. – A65 (20 pages). DOI: 10.1051/0004-6361/201527350
4. Pope S. B. Turbulent Flows. / S. B. Pope – Cambridge: Cambridge University Press, 2000. – 777 p. DOI: 10.1017/CBO9780511840531
5. Gorev V. N. Polynomial solutions for the Kolmogorov-Wiener filter weight function for fractal processes / V. N. Gorev, A. Yu. Gusev, V. I. Korniienko // Radio Electronics, Computer Science, Control. – 2019. – No. 2. – P. 44–52. DOI: 10.15588/1607-3274-2019-2-5
6. Gorev V. Investigation of the Kolmogorov-Wiener filter for continuous fractal processes on the basis of the Chebyshev polynomials of the first kind / V. Gorev, A. Gusev, V. Korniienko // IAPGOS. – 2020. – No. 1. – P. 58–61. DOI: 10.35784/iapgos.912
7. Gorev V. Investigation of the Kolmogorov-Wiener filter for treatment of fractal processes on the basis of the Chebyshev polynomials of the second kind / V. Gorev, A. Gusev, V. Korniienko // Ceur Workshop Proceedings. – 2019. – Vol. 2353. – P. 596–606.
8. Efficient Prediction of Network Traffic for Real-Time Applications / [M. F. Iqbal, M. Zahid, D. Habib, L. K. John] // Journal of Computer Networks and Communications. – 2019. – Vol. 2019. – 4067135 (11 pages). DOI: 10.1155/2019/4067135.
9. Katris C. Comparing forecasting approaches for Internet traffic / C. Katris, S. Daskalaki // Expert Systems with Applications. – 2015. – Vol. 42, Issue 21. – P. 8172–8183. DOI: 10.1016/j.eswa.2015.06.029.
10. Analysis of Self-Similar Traffic Models in Computer Networks / [J. S. Al-Azzeh, M. Al Hadidi, R. Odarchenko et al] // International Review on Modelling and Simulations. – 2017. – Vol. 10, No. 5. – P. 328–336. DOI: 10.15866/iremos.v10i5.12009
11. Miller S. Probability and Random Processes With Applications to Signal Processing and Communications. Second edition / S. Miller, D. Childers. – Amsterdam : Elsevier/Academic Press, 2012. – 598 p. DOI: doi.org/10.1016/B978-0-12-386981-4.50001-1
12. Sokolovsky A. I. On relaxation processes in a completely ionized plasma / A. I. Sokolovsky, S. A. Sokolovsky, O. A. Hrinishyn // East European Journal of Physics. – 2020. – No. 3. – P. 19–30. DOI: 10.26565/2312-4334-2020-3-03
13. Gonzalez R. G. Enskog kinetic theory for multicomponent granular suspensions / R. G. Gonzalez, N. Khalil, V. Garzo // Physical Review E. – 2020. – Vol. 101. – 012904 (23 pages). DOI: 10.1103/PhysRevE.101.012904
14. Wargnier Q. Consistent transport properties in multicomponent two-temperature magnetized plasmas. Application to the Sun atmosphere / [Q. Wargnier, A. A. Laguna, J. B. Scoggins et al] // Astronomy and Astrophysics. – 2020. – Vol. 635. – A87 (15 pages). DOI: 10.1051/0004-6361/201834686
15. Polyanin A. D. Handbook of integral equations. Second edition / A. D. Polyanin, A. V. Manzhirov. – Boca Raton: Chapman & Hall/CRC Press, Taylor & Francis Group, 2008. – 1144 p. DOI: 10.1201/9781420010558
16. Lindfield J. Numerical Methods Using MATLAB. Fourth edition / G. Lindfield, J. Penny. – Amsterdam : Elsevier, Academic Press, 2019. – 608 p. DOI: 10.1016/C2016-0-00395-9

COMBINED NEWTON'S THIRD-ORDER CONVERGENCE METHOD FOR MINIMIZE ONE VARIABLE FUNCTIONS

Kodnyanko V. A. – Dr. Sc., Professor, Polytechnic Institute, Siberian Federal University, Krasnoyarsk, Russia.

Grigorieva O. A. – PhD, Associate Professor, Polytechnic Institute, Siberian Federal University, Krasnoyarsk, Russia.

Strok L. V. – Postgraduate student, Polytechnic Institute, Siberian Federal University, Krasnoyarsk, Russia.

ABSTRACT

Contex. The article deals with the actual problem of numerical optimization of slowly computed unimodal functions of one variable. The analysis of existing methods of minimization of the first and second orders of convergence, which showed that these methods can be used to quickly solve these problems for functions, the values of which can be obtained without difficulty. For slowly computed functions, these methods give slow algorithms; therefore, the problem of developing fast methods for minimizing such functions is urgent.

Objective. Development of a combined third-order Newtonian method of convergence to minimize predominantly slowly computed unimodal functions, as well as the development of a database, including smooth, monotonic and partially constant functions, to test the method and compare its effectiveness with other known methods.

Method. A technique and an algorithm for solving the problem of fast minimization of a unimodal function of one variable by a combined numerical Newtonian method of the third order of convergence presented. The method is capable of recognizing strictly unimodal, monotonic and constant functions, as well as functions with partial or complete sections of a flat minimum.

Results. The results of comparison of the proposed method with other methods, including the fast Brent method, presented. 6954 problems were solved using the combined Newtonian method, while the method turned out to be faster than other methods in 95.5% of problems, Brent's method worked faster in only 4.5% of problems. In general, the analysis of the calculation results showed that the combined method worked 1.64 times faster than the Brent method.

Conclusions. A combined third-order Newtonian method of convergence proposed for minimizing predominantly slowly computed unimodal functions of one variable. A database of problems developed, including smooth, monotone and partially constant functions, to test the method and compare its effectiveness with other known methods. It is shown that the proposed method, in comparison with other methods, including the fast Brent method, has a higher performance.

KEYWORDS: unimodal function, Brent method, combined Newton minimization method, method speed.

ABBREVIATIONS

BM is a Brent method;

DS is a dichotomous search;

GSS is a golden section search;

EDS is an economical dichotomous search;

MF is a minimizing function;

N3o is a Newton's method of the third order of convergence.

NOMENCLATURE

$f(x)$ is an unimodal function;

$G = [a, b]$ is an uncertainty segment;

ε is a calculation accuracy;

$E = p^{1/\sigma}$ is an efficiency index;

p is an order of convergence of the method;

σ is a number of calculations;

δ is a number close to machine zero;

x_i is a current point of the iterative process, a_i, b_i are the current boundaries of the uncertainty segment at iteration with number i , δ is a number close to machine zero;

$h_i = \alpha(b_i - a_i)$ is a minimum gap restricting the excessive convergence between a new point and the boundaries of the uncertainty segment to hamper the slowdown of the computational process;

$\alpha < 0.5$ is a coefficient;

k_g is a number of calls of the GSS to the MF necessary to solve the problem.

INTRODUCTION

Numerical optimization of one-variable functions is one of the most common computational problems [1–4]. In practice, minimization of unimodal functions [5–8] that may have an extremum in a point or segment of the uncertainty segment where the function is constant is one of the most frequently used procedures. Most often, it is for unimodal functions that numerical minimization methods are developed.

First- and second-order convergence methods applied to solve these problems quickly and without difficulties [9–11]. The difference in their use becomes only visible when the computation time required to calculate the minimizing function becomes many times greater than the execution time of computational operations that implement the optimization method. Although the method itself operates fast, it is the slow MF computations that eventually slow down the solution algorithm.

1 PROBLEM STATEMENT

The class of slow-computed MF's is very broad. Typical examples are functions obtained by solving computational problems for theoretical models of various technical objects, where mathematical modeling involves

one or more multidimensional differential equations, often nonlinear. The numerical methods used to solve these problems usually require a large amount and high accuracy of calculations, thus consuming a significant amount of computer time to calculate a single MF value.

The proposed method is designed to find the abscissa x of the minimum of the unimodal one-variable function $f(x)$ on the segment $G = [a, b]$ up to ε .

During the search for a solution, at each iteration, a part of the current uncertainty segment G is cut off until the condition

$$|b_i - a_i| \leq \varepsilon |x_i| + \delta. \quad (1)$$

The value of the parameter δ was chosen from the following considerations. If the arithmetic type of machine variables equals to $2m$ significant decimal digits, then we can put $\delta = 10^{-m}$, which allows you to confidently operate with small numbers, without risking to affect the result of rounding errors when performing arithmetic calculations [16].

2 REVIEW OF THE LITERATURE

Even for fast-computed functions, when a large number of calls to the algorithm is required to achieve the result, it also becomes slow. Such problems include, for example, multi-parameter optimization of dynamic systems by performance criteria, where one-dimensional optimization is used [12, 13].

For algorithms with slow-computed functions and intrinsically slow algorithms, a significant efficiency factor is the number of k calls to the MF, which determines the speed of the algorithm and the corresponding method. Obviously, the less calls are required to the MF, the higher the speed of the algorithm and method.

For local minimization of one-variable functions, numerical first-order convergence methods are usually used. The preferred ones are the dichotomy search [2, 5], the golden section search [14] and the economical dichotomy search, the latter being superior in speed [15]. More often, second-order convergence methods are used, among which the leadership belongs to the Brent combined parabolic method [16]. The first-order methods are offer guaranteed reliability, however, in comparison with BM, with the same accuracy, they require a significantly larger number of k calls to the MF.

Along with BM, the Newton's analytic method [2, 9, 10], which has a second order of convergence, is used, which is highly efficient, especially when minimizing smooth functions. However, in terms of speed, when applied to slow-computed functions, a similar method based on numerical determination of the first derivative is noticeably inferior not only to BM, but even to first-order convergence methods. On this basis, it believed that for MF's requiring numerical differentiation, the Newton

method does not have any competitive advantage over the above-mentioned methods.

For optimization, we can try to apply new interpolation methods of third-order convergence, which give a quick solution to nonlinear equations [17]. When applied to smooth functions for the first derivative of the MF, these methods can yield a quick solution to the problem under consideration. The study of this opportunity is one of the objectives of this article.

In mass computing processes, monotonous, constant, or partially constant functions are often minimized. The Brent method applied to such functions does not give a gain in speed as compared to first-order methods, while monotonicity or constancy of the MF can be determined several times faster than the BM solution or one of the first-order methods. To study the opportunity to accelerate the minimization procedure for this class of functions is one of the goals of this study.

The following sections of the article discuss the method to attain the two above-mentioned objectives.

3 MATERIALS AND METHODS

The method includes two procedures to find the minimum MF

- Newton's method of third order convergence,
- method to identify its monotony and constancy segment.

The method is based on the provisions of [17], which outlines the idea of solving a nonlinear equation for smooth functions using the two-point iterative Weerakoon-Fernando method

$$\begin{aligned} x_{n+1} &= x_n - \frac{2f(x_n)}{f'(x_n) + f'(x_n^*)}, \\ x_n^* &= x_n - \frac{f(x_n)}{f'(x_n)}. \end{aligned} \quad (2)$$

In relation to the considered problem, one should search for the zero of the first derivative of the MF in which its minimum is achieved, therefore, to use the method (2), it is necessary to replace the MF in it with its first derivative, and the latter with the second derivative of the MF.

Method (2) has a third order of convergence with the efficiency index $\sqrt[3]{3} \approx 1.44$ [17]. It works faster than the Newton method since its efficiency index $\sqrt{2} \approx 1.41$ is less (by the efficiency index we mean the quantity $E = p^{1/\sigma}$).

In accordance with the proposed method, at each iteration, the MF is calculated only once; therefore, for $\sigma = 1$, the efficiency index (3) for polynomial approximation will be $E = 3$ against the Newton method for which $E = 2$.

In the process of calculations, a sequence of points $P(i) = (x_0, y_0), (x_1, y_1), \dots, (x_i, y_i)$ is formed. The necessary derivatives of the MF will be found by polynomial

approximation of the MF with respect to the last $n + 1$ points of the sequence P . Discarding the previous points, we obtain the sequence $P(n)$.

Imagine the MF polynomial

$$y(x) = \sum_{i=0}^n e_i L_i(x) = T_n(x), L_0(x) = 1; \quad (3)$$

$$L_{i+1}(x) = L_i(x)(x - x_i); \quad i = 0, 1, \dots, n-1$$

and find the unknown coefficients

$$e_0 = y_0, e_1 = \frac{y_1 - y_0}{x_1 - x_0}, \dots, e_n = \frac{y_n - T_{n-1}(x_n)}{L_n(x_n)}. \quad (4)$$

The formulas to calculate the derivatives of the function $y(x)$ are obtained using recurrence formulas similar to the Aitken scheme [18–20]

$$y'(x) = \sum_{i=0}^n e_i L_i'(x), L_0'(x) = 0, \quad (5)$$

$$L_{i+1}'(x) = L_i'(x)(x - x_i) + L_i(x),$$

$$i = 0, 1, \dots, n-1,$$

$$y''(x) = \sum_{i=0}^n e_i L_i''(x), L_0''(x) = 0, \quad (6)$$

$$L_{i+1}''(x) = L_i''(x)(x - x_i) + 2L_i'(x),$$

$$i = 0, 1, \dots, n-1.$$

Now method (6) takes the form

$$x_{n+1} = x_n - \frac{2y'(x_n)}{y''(x_n) + y''(x_n^*)}, \quad (7)$$

$$x_n^* = x_n - \frac{y'(x_n)}{y''(x_n)}.$$

The above considerations regarding the effectiveness of the methods correspond to the exact MF and its derivatives, however, in our case, the application of formulas (2)–(7) gives approximate dependences, therefore, the effectiveness of such a process requires a further study.

The solution finding process begins with determination of the initial segment of uncertainty with the boundaries $a_0 = a$, $b_0 = b$ and two initial points

$$x_0 = \frac{a_0(1+\varepsilon) + b_0(1-\varepsilon)}{2}, y_0 = f(x_0),$$

$$x_1 = \frac{a_0(1-\varepsilon) + b_0(1+\varepsilon)}{2}, y_1 = f(x_1).$$

If $y_0 = y_1$, then the solution is found in $k = 2$ MF calculations.

Otherwise, in accordance with the dichotomy method, a new segment of uncertainty $[a_1, b_1]$ is built by removing from segment $[a_0, b_0]$ segment $[a_0, x_0]$ or segment $[x_1, b_0]$ and the third point is determined

$$x_2 = \frac{a_1 + b_1}{2}, y_2 = f(x_2). \quad (8)$$

Next, a parabola built at three points

$$y(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1), \quad (9)$$

where

$$c_0 = y_0, c_1 = \frac{y_1 - y_0}{x_1 - x_0}, c_2 = \frac{y_2 - y_0 - c_1(x_2 - x_0)}{(x_2 - x_0)(x_1 - x_0)}.$$

For $c_2 > 0$, the parabola has a minimum at the point

$$x_3 = \frac{1}{2} \left(x_0 + x_1 - \frac{c_1}{c_2} \right).$$

If $a_1 < x_3 < b_1$, then the point is considered conditionally suitable, since verification of x_3 is required to meet the condition

$$a_i + h_i < x < b_i - h_i. \quad (10)$$

If and (5) is satisfied, then the abscissa of the new point is obtained and now we can calculate its ordinate $y_3 = f(x_3)$.

In cases when parabolic approximation does not yield a new point, it is determined by the half division method similarly to (8).

Then, the left or right segment of the uncertainty segment is again discarded, thereby giving its reduced analogue $[a_2, b_2]$.

Subsequent points are determined similarly, that is, using a parabolic approximation. The only difference is: if the point is suitable, then it is only the initial approximation for the Newtonian process, which is represented by formulas (4)–(7). In cases where parabolic approximation does not yield a result, the last point of the sequence P is taken for the initial approximation for method (4)–(7).

If in this case the found abscissa x falls into the current segment of uncertainty $G_i = [a_i, b_i]$, but does not satisfy condition (10), then it is corrected by the formula

$$x = \begin{cases} a_i + h, & x < a_i + h, \\ b_i - h, & x > b_i - h. \end{cases} \quad (11)$$

If for $k > 3$ at the current iteration neither the parabolic method nor the method (4)–(7) gives a result, then the new point is determined by the half division method. The process is repeated until the convergence condition (1) is satisfied.

While searching for the abscissa of the minimum, we also monitor the monotonicity and presence of constancy segment of the MF.

To monitor monotonicity, upon completion of the current iteration, sequence P is sorted by argument. If, after such sorting, the corresponding sequence of MF

values turns out to be monotonic, then starting from a certain iteration, the monotonicity of the MF is monitored.

It is known that the monotonicity of a unimodal function is detected in 4 MF calculations [8]. However, with a small number of points of the sequence P , monotonicity monitoring can slow down the optimization process, in cases where the MF is not monotonic.

Therefore, this monitoring is only advisable when a true monotonicity of the function under minimization is highly probable.

Monotonicity monitoring is based on the number of k calculations of the MF necessary to achieve the result.

If the sorted sequence P gives a monotonic sequence of ordinates, then two additional calculations of the MF at one of the ends of the segment can quickly give a solution to the problem.

So if the sequence turns out to be increasing, then it suffices to calculate at the points $x = a$ and $x = a + \varepsilon$. If, by adding the ordinates of these points to the above mentioned sequence, the monotonic sequence is obtained again, then the solution is $x = a$.

A descending sequence is identified similarly, point $x = b$ being the solution to the problem.

The procedure to identify constant functions and presence of individual sections of constant MF allows to accelerate the MF optimization. The procedure is very simple and obvious: if the sorted sequence P contains three points with the same ordinates, then the abscissa of any of them is a solution to the problem. This procedure does not require calculation of new MF values, however, for a number of functions it can significantly reduce the number and time of calculations.

4 EXPERIMENTS

To verify the method and other research goals, a problem base was built, its scope is determined by the product of the number of MF's used and the corresponding variants of uncertainty segments.

In numerical experiments, were used 26 extreme (the minimum of which is inside the uncertainty segments), two monotonic (increasing and decreasing), one constant and one partially constant functions.

Variants of the uncertainty segments for these MF's were found based on the following considerations. Each function was associated with a sufficiently wide segment $G = [a, b]$ and using the GSS, which was used as the control method, the abscissa x_m was determined by solving the problem accurate to machine zero. Then we randomly found two points c, d on the segment $[a, b]$, so that $d - c > \tau(b - a)$. This allowed us to obtain the uncertainty segments $[c, d]$ to solve the set of problems for a single function.

In the calculations, parameter τ was set from the range $0.25 \leq \tau \leq 1$, which made it possible to vary the length of the experimental segments for a certain function from a quarter to the full length of the segment G . The functions on the segments $[c, d]$ could be either monotonic, partially constant, or extreme.

When conducting a computational experiment, the solution of x problems was found accurate to $\varepsilon = 10^{-6}$ for arithmetic quantities supporting 19–20 significant digits. In the calculations, $\alpha = 0.12499$ was assumed.

The calculation results were compared with the indicators DS, EDS, GSS and BM.

The Brent method is programmed according to the original procedure [16], EDS is encoded based on the algorithm described in [15], DS and GSS are not difficult to encode.

5 RESULTS

To study the efficiency of the Newtonian third-order convergence procedure, only those segments $[c, d]$ and those functions that gave extremal problems for which the abscissa of the minimum of the MF fell inside these segments were selected.

A complete set of mass calculations was performed for all 30 functions for $\tau = 0.33$. In total, this amounted to 7.530 problems. Of these, 2.566 were identified as problems with monotonous functions. Using the combined Newtonian method, 7.215 were solved faster (95.8%), the Brent method was superior for only 4.2% of the problems. By the Brent method, the problems were solved within 127.461 calls to the MF; the combined method required 76.665 calculation of the MF. That is, the latter worked 1.66 faster than the first.

6 DISCUSSION

A typical optimization example is shown in Figure 1. Given: the MF and the boundaries of the segment of uncertainty $[c, d]$. To solve the problem, 5 methods were used: DS (Bisec), GSS (Gold), EDS (Eds), BM (Brent), N3o (N3o).

The lowest performance was shown by the DS, when used to solve the problem, it was required to calculate the minimized function 44 times. The GSS gave 31 calculations, the EDS required 26 calculations, the Brent method – 10, and the proposed Newtonian method gave the best result – 8 calculations.

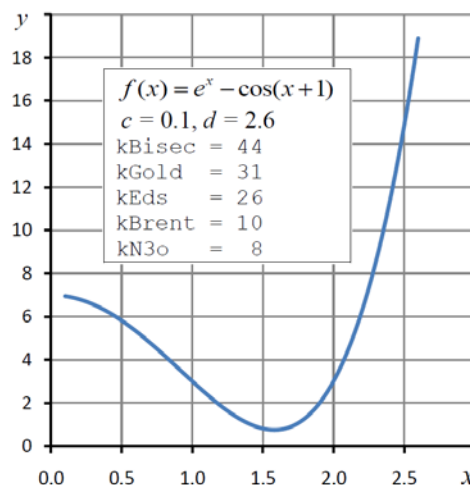


Figure 1 – Extreme function minimization results

When performing mass calculations on extreme problems for 26 functions, the minimum of which is inside the uncertainty segments, it was found that the Newtonian method is 13% faster than the Brent method. When solving problems, the Brent method required 63.092 calls to the MF, the Newtonian method – 55.873 times. In 4.830 problems out of 5.131, the Newton method worked faster (94.1%) and only in 301 cases the Brent method worked faster (5.6%). Thus, the Newtonian method showed the best result in mass calculations.

It was established that the Brent method had advantages over the Newton method only for slowly varying functions on the uncertainty segment. That was the function $f(x) = 1000(x-2.8)^4 - 1.6$. For example, on segment $G = [2.2, 3.2]$ was obtained $k_{Brent} = 25$, $k_{N3o} = 31$.

For any function in the minimization process, the monotonicity identification procedure described in Section 2.2 was applied.

As mentioned, monotonicity identification requires 2 additional function calculations. Therefore, in order not to lose speed, on the one hand, the procedure should be applied only at a high probability that the function is monotonic, and, on the other hand, the speed of the procedure should not be inferior to the Brent method for extreme functions, an example of which is shown in Figure 1.

Conducted above analysis showed that for the used extremal problems BM works approximately three times faster than the GSS. Therefore, the identification of the monotonicity of the function should be included if two conditions are satisfied: the ordinates $k_m - 2$ of the calculated points must form a monotonic sequence and the condition $k_m \geq k_g/3$ must be satisfied.

So for the one shown in Figure 1 example of $k_g = 31$. Therefore, monotonicity identification should be carried out when $k_m = 31/3 - 2 \approx 8$ ordinates of the calculated points give a monotonic sequence. Having computed the MF at two more points, we obtain the speed $k_m = 8 + 2 = 10$, which corresponds to the speed of Brent method for extreme MF.

Figure 2 shows an example of a function monotonic on a segment of the uncertainty of a function. Obviously, the Brent method's speed approximately corresponds to that of the GSS. At the same time, the monotonicity identification procedure made it possible to increase the speed by approximately a factor of three, which corresponds to the speed of the Brent method for extreme problems, the analysis of which presented above.

In order to verify the efficiency of the combined method combining Newtonian search with the monotonicity identification procedure, mass calculations were carried out for 26 extremal and two monotonic functions for the segments $[c, d]$ with the parameter $\tau = 0.33$. In total, this yielded 7.028 problems, of which 2.075 turned out to be problems with monotone functions, the rest of 4.953 were extremal.

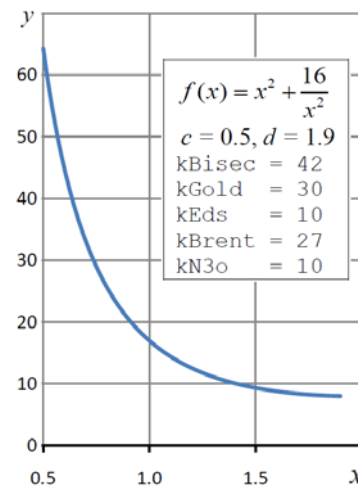


Figure 2 – Monotonic function minimization results

The combined method under consideration turned out to be the best, which solved all the problems having made 74.727 calls to the MF. The Brent method required 113.603 MF calculations. Thus, the combined method completed the problem 1.52 times faster.

Finally, the third kind of optimization problems are problems with fully and partially-constant functions. An example of such a function shown in Figure 3. The function has a flat bottom, any point of which is a solution to the problem.

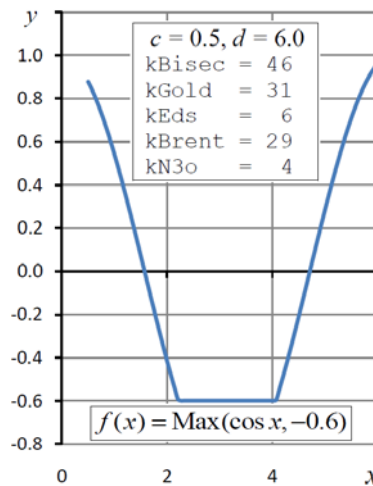


Figure 3 – Partial constant function minimization results

Identifying the minimum point of such a function is quite simple. In this case, no additional calculations are required. As mentioned above, if, during the calculation, sequence P gets three points with the same abscissas, then any of them will be a solution to the problem. For a problem with a function shown in Figure 3, the algorithm finds such a solution in just 4 calls to the function.

It is noteworthy that the EDS for monotonic and partially constant shows the same or close to the Newton method speed of solving problems. This is due to the fact that the described improvements are implemented in this method.

CONCLUSIONS

The paper considers and solves the problem of fast minimization of unimodal functions of one variable. A combined third-order Newtonian method of convergence is proposed for minimizing predominantly slowly computed unimodal functions of one variable. A database of problems has been developed, including smooth, monotonic and partially constant functions to test the method and compare its effectiveness with other known methods. The performance of the proposed method is compared with other methods, including the fast Brent method. Analysis of numerous data obtained using computational experiments showed that the proposed method has the highest performance. For smooth extremal functions, the method works 13% faster than Brent's method, for monotone functions its speed is 3 times higher, for partially constant functions the speed of the method is even higher.

The scientific novelty of the results obtained lies in the fact that for the first time a method for minimizing the minimization of unimodal functions of one variable was proposed, which in general works faster than the well-known fast methods. The most effective method for minimizing monotonic, constant and partially constant functions. The efficiency of the method is achieved through the use of a computational scheme that guarantees the third order of convergence to the minimum point of the optimized function. The practical value of the results obtained lies in the fact that the proposed method generally works faster than the known methods, which makes it possible to reduce the optimization time for unimodal functions. Prospects for further research are that the developed method can be extended to other functions, for example, to periodic, bimodal and other types of functions.

ACKNOWLEDGEMENTS

The work was carried out within the framework of the scientific research budget themes "Methods of modeling and optimizing of quality control information systems on the basis of intellectual technologies" at the Department of Standardization, Metrology and Quality Management of the Polytechnic Institute of the Siberian Federal University.

REFERENCES

1. Sahu D. R., Ansari Q. H., Yao J. C. A unified hybrid iterative method for hierarchical minimization problems, *Journal of Computational and Applied Mathematics*, 2013, Vol. 253 (1), pp. 208–221. <https://doi.org/10.1016/j.cam.2013.04.018>.
2. Rao S. S. Engineering optimization: theory and practice, John Wiley & Sons, 2009. <https://doi.org/10.1002/9781119454816>.
3. Gill P. E., Murray W., Wright M. H. Practical Optimization, Emerald Group Publishing Limited, 2019. <https://doi.org/10.1137/1.9781611975604>.
4. Aoki M. Introduction to optimization techniques: Fundamentals and Applications of Nonlinear Programming. London, Macmillan, 1971.
5. Bai Z., Tao M. Rigorous convergence analysis of alternating variable minimization with multiplier methods for quadratic programming problems with equality constraints, *BIT Numerical Mathematics*, 2016, 56, pp. 399–422. <https://doi.org/10.1007/s10543-015-0563-z>.
6. Shor N. Z. Minimization Methods for Non-Differentiable Functions, Springer Berlin Heidelberg, Softcover reprint of the original 1st ed. 1985, 2011. <https://doi.org/10.1007/978-3-642-82118-9>.
7. Himmelblau D. M. Applied Nonlinear Programming. New York, McGraw-Hill, 1972.
8. Hassin R., Reuven H. Asymptotic analysis of dichotomous search with search and travel costs, *European Journal of Operational Research*, 1992, Vol. 58(1), pp. 78–89. [https://doi.org/10.1016/0377-2217\(92\)90237-4](https://doi.org/10.1016/0377-2217(92)90237-4).
9. Glad T., Goldstein A. Optimization of functions whose values are subject to small errors, *BIT Numerical Mathematics*, 17, 1977, pp. 160–169. <https://doi.org/10.1007/bf01932287>.
10. Bazaraa M. S., Shetty C. M. Nonlinear Programming, Theory and Algorithms, New York, Wiley, 1979.
11. Chen D., Zhou Y., Song L. Fixed point algorithm based on adapted metric method for convex minimization problem with application to image deblurring, *Advances in Computational Mathematics*, 2016, 42, pp. 1287–1310. DOI: 10.1007/s10444-016-9462-3.
12. Ruszczyński A. Nonlinear Optimization, Princeton University Press, 2006. <https://doi.org/10.1515/9781400841059>.
13. Bian W., Chen X., Ye Y. Complexity analysis of interior point algorithms for non-Lipschitz and nonconvex minimization, *Mathematical Programming*, 2015, 149, pp. 301–327. DOI: 10.1007/s10107-014-0753-5.
14. Cai J., Hanb D., Chen Ch., Chen S. Application of the golden section search algorithm in the nonlinear isoconversional calculations to the determination of the activation energy from nonisothermal kinetic conversion data, *Solid State Sciences*, 2010, Vol.12 (5), pp. 829–833. DOI: 10.1021/ef7006672.
15. Kodnyanko V. A. Econometrical dichotomous search for minimizing one-variable functions, *Radio Electronics, Computer Science, Control*, 2019, No. 3, pp. 34–39. DOI: 10.15588/1607-3274-2019-3-4
16. Weerakoon S., Fernando T. G. I. A variant of Newton's method with accelerated third-order convergence, *Applied Mathematics Letters*, 2000, Vol. 13, pp. 87–93. DOI: 10.1016/S0893-9659(00)00100-2
17. Brent R. P. Algorithms for Minimization Without Derivatives, Dover, 2002. DOI: 10.2307/2005713
18. Deng Y., Glimm J., Yu Q., Eisenberg M. Global minimization for problems with multiple local minima, *Applied Mathematics Letters*, 1993, Vol. 6 (2), pp. 89–90. [https://doi.org/10.1016/0893-9659\(93\)90019-j](https://doi.org/10.1016/0893-9659(93)90019-j).
19. Gerald C. F., Wheatley P. O. Applied Numerical Analysis, fifth ed., Addison-Wesley Pub. Co., MA, 1994. <https://doi.org/10.2307/2007813>.
20. Atkinson K. Intro to Numerical Analysis, 2nd Edition, John Wiley & Sons, 1989.
21. Guimier A. Modelisation d'algorithmes d'optimisation a strategie aleatoire, *Calcolo*, 1998, 23, pp. 21–43. <https://doi.org/10.1007/bf02576906>.
22. Salgueiroda Silva M. A. A novel method for robust minimisation of univariate functions with quadratic convergence, *Journal of Computational and Applied Mathematics*, 2007, Vol. 200 (1), pp. 168–177. <https://doi.org/10.1016/j.cam.2005.12.020>.
23. Archetti F. Evaluation of random gradient techniques for unconstrained optimization, *Calcolo*, 1975, 12, pp. 83–94. <https://doi.org/10.1007/bf02576717>.

Received 12.04.2021.

Accepted 23.05.2021.

УДК 519.67

КОМБІНОВАНИЙ НЬЮТОНІВСЬКОЇ МЕТОД ТРЕТЬОГО ПОРЯДКУ ЗБІЖНОСТІ ДЛЯ МІНІМІЗАЦІЇ ФУНКЦІЙ ОДНІЄЇ ЗМІННОЮ

Коднянко В. А. – д-р техн. наук, професор кафедри стандартизації, метрології та управління якістю, Політехнічний інститут Сибірського федерального університету, Красноярськ, Росія.

Григор'єва О. А. – канд. техн. наук, доцент кафедри стандартизації, метрології та управління якістю, Політехнічний інститут Сибірського федерального університету, Красноярськ, Росія.

Рядків Л. В. – аспірант Сибірського федерального університету, Красноярськ, Росія.

АНОТАЦІЯ

Актуальність. У статті розглядається актуальна проблема чисельної оптимізації повільно обчислюваних унімодальних функцій однієї змінної. Проведено аналіз існуючих методів мінімізації першого і другого порядку збіжності, який показав, що ці методи можуть бути використані для швидкого вирішення цих завдань для функцій, значення яких можуть бути отримані без труднощів. Для повільно обчислюваних функцій ці методи дають повільні алгоритми, тому актуальною є задача розробки швидких методів для мінімізації таких функцій.

Мета. Розробка комбінованого ньютонівського методу третього порядку збіжності для мінімізації переважно повільно обчислюваних унімодальних функцій, а також розробка бази даних, що включає гладкі, монотонні і частково постійні функції, для перевірки методу і порівняння його ефективності з іншими відомими методами.

Метод. Викладається методика і алгоритм вирішення завдання швидкої мінімізації унімодальної функції однієї змінної комбінованим чисельним ньютонівським методом третього порядку збіжності. Метод здатний розпізнавати строго унімодальне, монотонні і постійні функції, а також функції з частковими або повними ділянками плоского мінімуму.

Результати. Наводяться результати порівняння запропонованого методу з іншими методами, в тому числі швидким методом Брента. За допомогою комбінованого ньютонівського методу вирішено 6954 завдання, при цьому метод виявився швидше за інших методів в 95,5% завдань, метод Брента працював швидше лише в 4,5% завдань. В цілому аналіз результатів розрахунку показав, що комбінований метод працював в 1,64 рази швидше, ніж метод Брента.

Висновки. Запропоновано комбінований ньютонівський метод третього порядку збіжності для мінімізації переважно повільно обчислюваних унімодальних функцій однієї змінної. Розроблено базу даних завдань, включаючи гладкі, монотонні і частково постійні функції, для перевірки методу і порівняння його ефективності з іншими відомими методами. Показано, що запропонований метод в порівнянні з іншими методами, включаючи швидкий метод Брента, має більш високу швидкість.

КЛЮЧОВІ СЛОВА: унімодальна функція, метод Брента, метод Ньютона, швидкість методу.

УДК 519.67

КОМБИНИРОВАННЫЙ НЬЮТОНОВСКИЙ МЕТОД ТРЕТЬЕГО ПОРЯДКА СХОДИМОСТИ ДЛЯ МИНИМИЗАЦИИ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

Коднянко В. А. – д-р техн. наук, профессор кафедры стандартизации, метрологии и управления качеством, Политехнический институт Сибирского федерального университета, Красноярск, Россия.

Григорьева О. А. – канд. техн. наук, доцент кафедры стандартизации, метрологии и управления качеством, Политехнический институт Сибирского федерального университета, Красноярск, Россия.

Строк Л. В. – аспирант Сибирского федерального университета, Красноярск, Россия.

АННОТАЦИЯ

Актуальность. В статье рассматривается актуальная проблема численной оптимизации медленно вычисляемых унимодальных функций одной переменной. Проведен анализ существующих методов минимизации первого и второго порядка сходимости, который показал, что эти методы могут быть использованы для быстрого решения этих задач для функций, значения которых могут быть получены без затруднений. Для медленно вычисляемых функций эти методы дают медленные алгоритмы, поэтому актуальной является задача разработки быстрых методов для минимизации таких функций.

Цель. Разработка комбинированного ньютонівського методу третього порядку сходимости для минимизации преимущественно медленно вычисляемых унимодальных функций, а также разработка базы данных, включающая гладкие, монотонные и частично постоянные функции, для проверки метода и сравнения его эффективности с другими известными методами.

Метод. Излагается методика и алгоритм решения задачи быстрой минимизации унимодальной функции одной переменной комбинированным численным ньютонівським методом третього порядка сходимости. Метод способен распознавать строго унимодальные, монотонные и постоянные функции, а также функции с частичными или полными участками плоского минимума.

Результаты. Приводятся результаты сравнение предлагаемого метода с другими методами, в том числе быстрым методом Брента. С помощью комбинированного ньютонівського методу решено 6954 задачи, при этом метод оказался быстрее других методов в 95,5% задач, метод Брента работал быстрее лишь в 4,5% задач. В целом анализ результатов расчета показал, что комбинированный метод работал в 1,64 раза быстрее, чем метод Брента.

Выводы. Предложен комбинированный ньютонівський метод третього порядка сходимости для минимизации преимущественно медленно вычисляемых унимодальных функций одной переменной. Разработана база данных задач,

включая гладкие, монотонные и частично постоянные функции, для проверки метода и сравнения его эффективности с другими известными методами. Показано, что предлагаемый метод по сравнению с другими методами, включая быстрый метод Брента, имеет более высокое быстродействие.

КЛЮЧЕВЫЕ СЛОВА: унимодальная функция, метод Брента, метод Ньютона, скорость метода.

ЛИТЕРАТУРА / LITERATURA

1. Sahu D. R. A unified hybrid iterative method for hierarchical minimization problems / D. R. Sahu, Q. H. Ansari, J. C. Yao // *Journal of Computational and Applied Mathematics*. – 2013. – Vol. 253 (1). – P. 208–221. <https://doi.org/10.1016/j.cam.2013.04.018>.
2. Rao S. S. *Engineering optimization: theory and practice* / S. S. Rao. – Sidney : John Wiley & Sons, 2009. – 832 p. <https://doi.org/10.1002/9781119454816>.
3. Gill P. E. *Practical Optimization* / P. E. Gill, W. Murray, M. H. Wright. – London : Academic Press Limited, 2019. – 418 p. <https://doi.org/10.1137/1.9781611975604>.
4. Aoki M. *Introduction to optimization techniques: Fundamentals and Applications of Nonlinear Programming* / M. Aoki. – London : Macmillan, 1971. – 335 p.
5. Bai Z. Rigorous convergence analysis of alternating variable minimization with multiplier methods for quadratic programming problems with equality constraints / Z. Bai, M. Tao // *BIT Numerical Mathematics* – 2016. – № 56. – P. 399–422. <https://doi.org/10.1007/s10543-015-0563-z>.
6. Shor N. Z. *Minimization Methods for Non-Differentiable Functions* / N. Z. Shor. – Berlin : Springer, 1985. – 164 p. <https://doi.org/10.1007/978-3-642-82118-9>.
7. Himmelblau D. M. *Applied Nonlinear Programming* / D. M. Himmelblau. – New York: McGraw-Hill, 1972. – 498 p.
8. Hassin R. Asymptotic analysis of dichotomous search with search and travel costs / R. Hassin, H. Reuven // *European Journal of Operational Research*. – 1992. – Vol. 58. – P. 78–89. [https://doi.org/10.1016/0377-2217\(92\)90237-4](https://doi.org/10.1016/0377-2217(92)90237-4).
9. Glad T. Optimization of functions whose values are subject to small errors / T. Glad, A. Goldstein // *BIT Numerical Mathematics* – 1977. – № 17. – P. 160–169. <https://doi.org/10.1007/bf01932287>.
10. Bazaraa M. S. *Nonlinear Programming, Theory and Algorithms* / M. S. Bazaraa, C. M. Shetty. – New York : Wiley, 1979. – 872 p.
11. Chen D. Fixed point algorithm based on adapted metric method for convex minimization problem with application to image deblurring / D. Chen, Y. Zhou, L. Song // *Advances in Computational Mathematics* – 2016. – № 42. – P. 1287–1310.
12. Ruszczyński A. *Nonlinear Optimization* / A. Ruszczyński – New Jersey : Princeton University Press, 2006. – 464 p. <https://doi.org/10.1515/9781400841059>.
13. Bian W. Complexity analysis of interior point algorithms for non-Lipschitz and nonconvex minimization / W. Bian, X. Chen, Y. Ye // *Mathematical Programming* – 2015. – № 149. – P. 301–327.
14. Application of the golden section search algorithm in the nonlinear isoconversional calculations to the determination of the activation energy from nonisothermal kinetic conversion data / [J. Cai, D. Hanb, Ch. Chen, S. Chen] // *Solid State Sciences*. – 2010. – Vol. 12 (5). – P. 829–833.
15. Kodnyanko V. A. Economical dichotomous search for minimizing one-variable functions, / V. A. Kodnyanko // *Radio Electronics, Computer Science, Control*. – 2019. – № 3. – P. 34–39.
16. Weerakoon S. A variant of Newton's method with accelerated third-order convergence / S. Weerakoon, T. G. I. Fernando // *Applied Mathematics Letters*. – 2000. – Vol. 13. – P. 87–93.
17. Brent R. P. *Algorithms for Minimization Without Derivatives* / R. P. Brent. – New York : Dover Publications, 2002. – 206 p.
18. Global minimization for problems with multiple local minima / [Y. Deng, J. Glimm, Q. Yu, M. Eisenberg] // *Applied Mathematics Letters*. – 1993. – Vol. 6 (2). – P. 89–90. [https://doi.org/10.1016/0893-9659\(93\)90019-j](https://doi.org/10.1016/0893-9659(93)90019-j).
19. Gerald C. F. *Applied Numerical Analysis, fifth ed.* / C. F. Gerald, P. O. Wheatley – Boston : Addison Wesley Publishing Company, 1994. – 748 p. <https://doi.org/10.2307/2007813>.
20. Atkinson K. *Intro to Numerical Analysis, 2nd Edition* / K. Atkinson. – Sidney : John Wiley & Sons, 1989. – 712 p.
21. Guimier A. Modelisation d'algorithmes d'optimisation a strategie aleatoire / A. Guimier // *Calcolo*. – 1986. – № 23. – P. 21–43. <https://doi.org/10.1007/bf02576906>.
22. Salgueiroda Silva M. A. A novel method for robust minimisation of univariate functions with quadratic convergence / M. A. Salgueiroda Silva // *Journal of Computational and Applied Mathematics*. – 2007. – Vol. 200 (1). – P. 168–177. <https://doi.org/10.1016/j.cam.2005.12.020>.
23. Archetti F. Evaluation of random gradient techniques for unconstrained optimization / F. Archetti // *Calcolo*. – 1975. – № 12. – P. 83–94. <https://doi.org/10.1007/bf02576717>.

DELAY MODELS BASED ON SYSTEMS WITH USUAL AND SHIFTED HYPEREXPONENTIAL AND HYPERERLANGIAN INPUT DISTRIBUTIONS

Tarasov V. N. – Dr. Sc., Professor, Head of Department of Software and Management in Technical Systems of Volga State University of Telecommunications and Informatics, Samara, Russian Federation.

Bakhareva N. F. – Dr. Sc., Professor, Head of Department of Informatics and Computer Engineering of Volga State University of Telecommunications and Informatics, Samara, Russian Federation.

ABSTRACT

Context. In the queueing theory, the study of systems with arbitrary laws of the input flow distribution and service time is relevant because it is impossible to obtain solutions for the waiting time in the final form for the general case. Therefore, the study of such systems for particular cases of input distributions is important.

Objective. Getting a solution for the average delay in the queue in a closed form for queueing systems with ordinary and with shifted to the right from the zero point hyperexponential and hypererlangian distributions in stationary mode.

Method. To solve this problem, we used the classical method of spectral decomposition of the solution of the Lindley integral equation. This method allows to obtaining a solution for the average delay for two systems under consideration in a closed form. The method of spectral decomposition of the solution of the Lindley integral equation plays an important role in the theory of systems G/G/1. For the practical application of the results obtained, the well-known method of moments of probability theory is used.

Results. For the first time, a spectral decomposition of the solution of the Lindley integral equation for systems with ordinary and with shifted hyperexponential and hypererlangian distributions is obtained, which is used to derive a formula for the average delay in a queue in closed form.

Conclusions. It is proved that the spectral expansions of the solution of the Lindley integral equation for the systems under consideration coincide; therefore, the formulas for the mean delay will also coincide. It is shown that in systems with a delay, the average delay is less than in conventional systems. The obtained expression for the waiting time expands and complements the well-known incomplete formula of queueing theory for the average delay for systems with arbitrary laws of the input flow distribution and service time. This approach allows us to calculate the average delay for these systems in mathematical packages for a wide range of traffic parameters. In addition to the average waiting time, such an approach makes it possible to determine also moments of higher orders of waiting time. Given the fact that the packet delay variation (jitter) in telecommunications is defined as the spread of the waiting time from its average value, the jitter can be determined through the variance of the waiting time.

KEYWORDS: delayed system, shifted distributions, Laplace transform, Lindley integral equation, spectral decomposition method.

ABBREVIATIONS

LIE is a Lindley integral equation;
QS is a queueing system;
PDF is a probability distribution function.

NOMENCLATURE

$a(t)$ is a density function of the distribution of time between arrivals;

$A^*(s)$ is a Laplace transform of the function $a(t)$;

$b(t)$ is a density function of the distribution of service time;

$B^*(s)$ is a Laplace transform of the function $b(t)$;

c_λ is a coefficient of variation of time between arrivals;

c_μ is a coefficient of variation of service time;

G is a arbitrary distribution law;

H_2 is a hyperexponential distribution of the second order;

H_2^- is a shifted hyperexponential distribution of the second order;

HE_2 is a hypererlangian distribution of the second order;

HE_2^- is a shifted hypererlangian distribution of the second order;

\bar{W} is an average delay in the queue;

$W^*(s)$ is a Laplace transform of delay density function;

λ_1, λ_2 is a parameters of the hyperexponential distribution law of the input flow;

μ_1, μ_2 is a parameters of the hypererlangian distribution law of service time;

ρ is a system load factor;

$\bar{\tau}_\lambda$ is a average time between arrivals;

$\bar{\tau}_\lambda^2$ is a second initial moment of time between arrivals;

$\bar{\tau}_\mu$ is an average service time;

$\bar{\tau}_\mu^2$ is a second initial moment of service time;

$\Phi_+(s)$ is a Laplace transform of the PDF of waiting time;

$\Psi_+(s)$ is a first component of spectral decomposition;

$\psi_-(s)$ is a second component of spectral decomposition.

INTRODUCTION

In the study of G/G/1 systems, an important role is played by the method of spectral decomposition of the solution of the Lindley integral equation (LIE). The most accessible this method with specific examples is described in the classic queueing theory [1].

This article is devoted to the analysis of the QS $H_2/HE_2/1$ formed by two flows described by the usual and shifted to the right from the zero point by the density functions of the hyperexponential and hypererlangian distribution laws of the second order. In the previous works of the authors, it is clearly shown that in systems formed by shifted distribution laws, with the same load factor as compared with conventional systems, the average delay becomes less.

This is achieved because the coefficients of variation of the arrival c_λ and service times c_μ for shifted distribution laws become smaller when entering the shift parameter $t_0 > 0$. Thus, the operation of shifting the distribution law transforms ordinary Markov queueing systems into a non-Markov system.

The results of works [2–7] in the domain of QS with shifted distributions together with [1] made it possible to develop the method of spectral expansion of the solution (LIE) into the considered systems $H_2/HE_2/1$.

Here the superscript “–” will mean the operation of the shift of the distribution law.

In the queueing theory, the studies of G/G/1 systems are relevant because they are actively used in modern teletraffic theory; moreover, it is impossible to obtain solutions for such systems in the final form for the general case.

The object of study is the main characteristic – the average waiting time of requirements in the queue of the queueing systems type G/G/1.

The subject of study is the average waiting time of requirements in the queue of the QS $H_2/HE_2/1$ and $H_2^-/HE_2^-/1$. Here the superscript “–” will mean the operation of the shift of the distribution law.

The purpose of the work is obtaining a solution for the average delay of requirements in the queue in closed form for these systems.

1 PROBLEM STATEMENT

The paper poses the problem of finding a solution for the average delay of requirements in a queue for QS $H_2/HE_2/1$ and $H_2^-/HE_2^-/1$.

When using the method of spectral decomposition of an LIE solution to determine the average waiting time, we will follow the approach and symbolism of the author of the classical queueing theory [1]. To solve the problem, it is necessary to find the law of waiting time distribution in the system through the spectral decomposition of the

form: $A^*(-s) \cdot B^*(s) - 1 = \psi_+(s) / \psi_-(s)$, where $\psi_+(s)$ and $\psi_-(s)$ are some fractional rational functions of s that can be factorized. Functions $\psi_+(s)$ and $\psi_-(s)$ must satisfy the following conditions according to [1]:

- 1) for $\text{Re}(s) > 0$ function $\psi_+(s)$ is analytic without zeros in this half-plane;
- 2) for $\text{Re}(s) < D$ function $\psi_-(s)$ is analytic without zeros in this half-plane, where D is some positive constant defined by the condition:

$$\lim_{t \rightarrow \infty} a(t) / e^{-Dt} < \infty. \quad (1)$$

In addition, functions $\psi_+(s)$ and $\psi_-(s)$ must have the following properties:

$$\lim_{|s| \rightarrow \infty, \text{Re}(s) > 0} \frac{\psi_+(s)}{s} = 1, \quad \lim_{|s| \rightarrow \infty, \text{Re}(s) < D} \frac{\psi_-(s)}{s} = -1. \quad (2)$$

To solve the problem, it is necessary first to construct for these systems spectral decompositions of the form $A^*(-s) \cdot B^*(s) - 1 = \psi_+(s) / \psi_-(s)$, taking into account conditions (1), (2).

2 REVIEW OF THE LITERATURE

The method of spectral decomposition of the solution of the Lindley integral equation was first presented in detail in the classic queueing theory [1], and was subsequently used in many papers, including [8, 9]. A different approach to solving Lindley’s equation has been used in Russian language literature. That work used factorization instead of the term “spectral decomposition” and instead of the functions $\psi_+(s)$ and $\psi_-(s)$ it used factorization components $\omega_+(z, t)$ and $\omega_-(z, t)$ of the function $1 - z \cdot \chi(t)$, where $\chi(t)$ is the characteristic function of a random variable ξ with an arbitrary distribution function $C(t)$, and z is any number from the interval $(-1, 1)$. This approach for obtaining results for systems under consideration is less convenient than the approach described and illustrated with numerous examples in [1].

Practical application of the method of spectral decomposition of the LIE solution for studying systems with shifted distribution laws, i.e. systems with time delay, shown in the works of the authors [2–7], as well as in other works.

In the scientific literature, the term “system with delay” does not occur, there is only one term “queues with a delay in time” [10, 11]. In [10, 11] presents the results of the approach of queues to the Internet and mobile services as queues with a delay in time. It is shown that if information is delayed long enough, a Hopf bifurcation can occur, which can cause unwanted fluctuations in the queues. However, it is not known how large the fluctuations are when the Hopf bifurcation

occurs. This is the first publication in the English-language journals about queues with a delay.

Approximation methods with respect to distribution laws are described in detail in [9, 13, 14, 23, 24], and new research in queuing theory has recently been carried out in [15–24].

3 MATERIALS AND METHODS

Let us first consider the system $H_2/HE_2/1$, formed by distribution laws with density functions

$$a(t) = p\lambda_1 e^{-\lambda_1 t} + (1-p)\lambda_2 e^{-\lambda_2 t}, \quad (3)$$

$$b(t) = 4q\mu_1^2 t e^{-2\mu_1 t} + 4(1-q)\mu_2^2 t e^{-2\mu_2 t}. \quad (4)$$

The distribution laws (3) and (4) are the most common distributions of non-negative continuous random variables, since they have a wide range of variation of the coefficient of variation.

Then we get

$$A^*(s) = p \frac{\lambda_1}{s + \lambda_1} + (1-p) \frac{\lambda_2}{s + \lambda_2},$$

$$B^*(s) = q \left(\frac{2\mu_1}{s + 2\mu_1} \right)^2 + (1-q) \left(\frac{2\mu_2}{s + 2\mu_2} \right)^2.$$

Then the spectral decomposition of the solution of the LIE for the system $H_2/HE_2/1$ takes the form

$$\frac{\psi_+(s)}{\psi_-(s)} = \left[p \frac{\lambda_1}{\lambda_1 - s} + (1-p) \frac{\lambda_2}{\lambda_2 - s} \right] \times$$

$$\times \left[q \left(\frac{2\mu_1}{2\mu_1 + s} \right)^2 + (1-q) \left(\frac{2\mu_2}{2\mu_2 + s} \right)^2 \right] - 1.$$

The first factor in the right part in square brackets is equal to

$$\left[p \frac{\lambda_1}{\lambda_1 - s} + (1-p) \frac{\lambda_2}{\lambda_2 - s} \right] = \frac{\lambda_1 \lambda_2 - [p\lambda_1 + (1-p)\lambda_2]s}{(\lambda_1 - s)(\lambda_2 - s)}$$

$$= \frac{a_0 - a_1 s}{(\lambda_1 - s)(\lambda_2 - s)},$$

where are the intermediate parameters $a_0 = \lambda_1 \lambda_2$, $a_1 = p\lambda_1 + (1-p)\lambda_2$. Similarly, we represent the second factor

$$\left[q \left(\frac{2\mu_1}{2\mu_1 + s} \right)^2 + (1-q) \left(\frac{2\mu_2}{2\mu_2 + s} \right)^2 \right] =$$

$$= \frac{q(16\mu_1^2 \mu_2^2 + 16\mu_1^2 \mu_2 s + 4\mu_1^2 s^2)}{(2\mu_1 + s)^2 (2\mu_2 + s)^2} +$$

$$+ \frac{(1-q)(16\mu_1^2 \mu_2^2 + 16\mu_1 \mu_2^2 s + 4\mu_2^2 s^2)}{(2\mu_1 + s)^2 (2\mu_2 + s)^2} =$$

$$= \frac{b_0 + b_1 s + b_2 s^2}{(2\mu_1 + s)^2 (2\mu_2 + s)^2},$$

where are the intermediate parameters $b_0 = 16\mu_1^2 \mu_2^2$, $b_1 = 16\mu_1 \mu_2 [q\mu_1 + (1-q)\mu_2]$, $b_2 = 4[q\mu_1^2 + (1-q)\mu_2^2]$.

Then the desired expression for the spectral decomposition will be

$$\frac{\psi_+(s)}{\psi_-(s)} = \frac{(a_0 - a_1 s)(b_0 + b_1 s + b_2 s^2)}{(\lambda_1 - s)(\lambda_2 - s)(2\mu_1 + s)^2 (2\mu_2 + s)^2} -$$

$$\frac{(\lambda_1 - s)(\lambda_2 - s)(2\mu_1 + s)^2 (2\mu_2 + s)^2}{(\lambda_1 - s)(\lambda_2 - s)(2\mu_1 + s)^2 (2\mu_2 + s)^2}. \quad (5)$$

The polynomial in the numerator in the right part of the decomposition (5) as a rule always has one zero [1]. In this case, the free term of the decomposition is also 0: $a_0 b_0 - 16\lambda_1 \lambda_2 \mu_1^2 \mu_2^2 = 0$. In the numerator of the fraction in the right part of the decomposition a polynomial of the sixth degree was obtained, the coefficients of which are equal to:

$$c_0 = a_0 b_1 - a_1 b_0 + b_0(\lambda_1 + \lambda_2) - 16a_0 \mu_1 \mu_2 (\mu_1 + \mu_2),$$

$$c_1 = a_0 b_2 - a_1 b_1 - b_0 - 4a_0(\mu_1^2 + \mu_2^2) +$$

$$+ 16(\lambda_1 + \lambda_2)(\mu_1 + \mu_2)\mu_1 \mu_2 - 16a_0 \mu_1 \mu_2,$$

$$c_2 = 4(\lambda_1 + \lambda_2)[(\mu_1 + \mu_2)^2 + 2\mu_1 \mu_2] -$$

$$- 4(\mu_1 + \mu_2)(a_0 + 4\mu_1 \mu_2) - a_1 b_2, \quad (6)$$

$$c_3 = 4(\lambda_1 + \lambda_2)(\mu_1 + \mu_2) - 4[(\mu_1 + \mu_2)^2 + 2\mu_1 \mu_2] - a_0,$$

$$c_4 = \lambda_1 + \lambda_2 - 4(\mu_1 + \mu_2).$$

The coefficients (6) are obtained using the Mathcad symbolic operations, since the numerator of the decomposition (5) even after the introduction of intermediate parameters contains 42 terms. Apparently, the lack of results for the system under consideration is explained by the large laboriousness of the calculations.

We select the polynomial in the numerator of the decomposition (5)

$$s^5 - c_4 s^4 - c_3 s^3 - c_2 s^2 - c_1 s - c_0, \quad (7)$$

because determining its roots is the main part of the spectral decomposition method.

The study of polynomial (7) with coefficients (6) using the Vietet formulas confirms the presence of four negative real roots and one positive root, or instead of the first – two negative real roots and two complex conjugate roots with negative real parts.

The coefficient of the polynomial (7) $c_0 > 0$ is always in the case of a stable system, when $0 < \rho < 1$. Taking into account the minus sign in the polynomial before the coefficient, Vietet formulas do not contradict the fact that there are four negative roots of the polynomial (7).

Denoting the negative roots of the polynomial (7) or their negative real parts for convenience through $-s_1, -s_2, -s_3, -s_4$, and the positive root through s_5 , the relation $\psi_+(s)/\psi_-(s)$ can be finally decomposed into the following factors:

$$\frac{\psi_+(s)}{\psi_-(s)} = \frac{-s(s+s_1)(s+s_2)(s+s_3)(s+s_4)(s-s_5)}{(\lambda_1-s)(\lambda_2-s)(2\mu_1+s)^2(2\mu_2+s)^2}. \quad (8)$$

Now we build the functions $\psi_+(s)$ and $\psi_-(s)$ taking into account the conditions (1) and (2):

$$\psi_+(s) = \frac{s(s+s_1)(s+s_2)(s+s_3)(s+s_4)}{(2\mu_1+s)^2(2\mu_2+s)^2},$$

because the zeros of the polynomial (7): $s = 0, -s_1, -s_2, -s_3, -s_4$ and double poles $s = -2\mu_1, s = -2\mu_2$ lie in the half-plane $\text{Re}(s) \leq 0$,

$$\psi_-(s) = -\frac{(\lambda_1-s)(\lambda_2-s)}{(s-s_5)},$$

because its zeros and the pole lie in the region $\text{Re}(s) > D$ defined by condition (1).

The fulfillment of conditions (1) and (2) for functions $\psi_+(s)$ and $\psi_-(s)$ is obvious, which is also confirmed by figure 1.

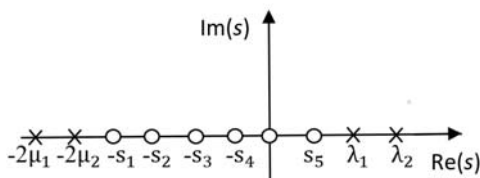


Figure 1 – Zeros and poles of the function $\psi_+(s)/\psi_-(s)$ for the system $H_2/HE_2/1$

When constructing these functions, it is more convenient to mark the zeros and poles of the relation $\psi_+(s)/\psi_-(s)$ on the complex s – plane to eliminate errors in the construction of the functions $\psi_+(s)$ and $\psi_-(s)$. In Figure 1, the poles are marked with crosses, and zeros are indicated by circles.

According to the method of spectral decomposition, the constant K is determined from the condition:

$$K = \lim_{s \rightarrow 0} \frac{\psi_+(s)}{s} = \frac{s_1 s_2 s_3 s_4}{16\mu_1^2 \mu_2^2}.$$

The constant K determines the probability that the demand entering the system finds it free. Through $\psi_+(s)$ and the constant K we define

$$\Phi_+(s) = \frac{K}{\psi_+(s)} = \frac{s_1 s_2 s_3 s_4 (s+2\mu_1)^2 (s+2\mu_2)^2}{16s\mu_1^2 \mu_2^2 (s+s_1)(s+s_2)(s+s_3)(s+s_4)}.$$

From here we get

$$W^*(s) = s \cdot \Phi_+(s) = \frac{s_1 s_2 s_3 s_4 (s+2\mu_1)^2 (s+2\mu_2)^2}{16\mu_1^2 \mu_2^2 (s+s_1)(s+s_2)(s+s_3)(s+s_4)}. \quad (9)$$

The desired average waiting time in the queue is equal to the value of the derivative of the Laplace transform (9) of the density function with a minus sign at the point $s = 0$:

$$-\left. \frac{dW^*(s)}{ds} \right|_{s=0} = \frac{1}{s_1} + \frac{1}{s_2} + \frac{1}{s_3} + \frac{1}{s_4} - \frac{1}{\mu_1} - \frac{1}{\mu_2}.$$

Finally for QS $H_2/HE_2/1$

$$\bar{W} = \frac{1}{s_1} + \frac{1}{s_2} + \frac{1}{s_3} + \frac{1}{s_4} - \frac{1}{\mu_1} - \frac{1}{\mu_2}. \quad (10)$$

From expression (9) if necessary, we can also determine higher order moments for the waiting time. The second derivative of the transformation (9) at the point $s = 0$ gives the second initial moment of the waiting time, which allows us to determine the variance of the waiting time. Taking into account the definition of jitter in telecommunications as the spread of waiting time around its average value [10], we thereby obtain the possibility of determining jitter through variance. This is an important result for analyzing latency sensitive traffic.

We now turn to the study of the $H_2/HE_2/1$ system with shifted input distributions, i.e. to the system with a delay in time. Such a system, unlike the usual system, we denote $H_2^-/HE_2^-/1$. This system will be described by distribution laws:

$$a(t) = p\lambda_1 e^{-\lambda_1(t-t_0)} + (1-p)\lambda_2 e^{-\lambda_2(t-t_0)}, \quad (11)$$

$$b(t) = 4q\mu_1^2(t-t_0)e^{-2\mu_1(t-t_0)} + 4(1-q)\mu_2^2(t-t_0)e^{-2\mu_2(t-t_0)}. \quad (12)$$

Density functions (11) and (12) are shifted to the right from the zero point by the value $t_0 > 0$ of second-order hyperexponential and hyper-Erlang distributions. To find the average waiting time in the queue for this system, we prove the following statement.

Statement. The spectral decompositions $A^*(-s) * B^*(s) - 1 = \psi_+(s)/\psi_-(s)$ of the LIE solution

for systems $H_2^- / HE_2^- / 1$ and $H_2 / HE_2 / 1$ completely coincide and have the form (8), i.e. the spectral decomposition of $A^*(-s) * B^*(s) - 1$ is invariant to the operation of the time shift of the density function.

Proof. For a $H_2^- / HE_2^- / 1$ system, the spectral decomposition will be:

$$\begin{aligned} \frac{\Psi_+(s)}{\Psi_-(s)} &= [p \frac{\lambda_1}{\lambda_1 - s} + (1-p) \frac{\lambda_2}{\lambda_2 - s}] e^{t_0 s} \times \\ &\times [q \left(\frac{2\mu_1}{2\mu_1 + s} \right)^2 + (1-q) \left(\frac{2\mu_2}{2\mu_2 + s} \right)^2] e^{-t_0 s} - 1 = \\ &= [p \frac{\lambda_1}{\lambda_1 - s} + (1-p) \frac{\lambda_2}{\lambda_2 - s}] \times \\ &\times [q \left(\frac{2\mu_1}{2\mu_1 + s} \right)^2 + (1-q) \left(\frac{2\mu_2}{2\mu_2 + s} \right)^2] - 1. \end{aligned}$$

Here, exponents with opposite signs of exponential functions are reset to zero, and thus the shift operation in the spectral decomposition is leveled. Thus, the spectral decompositions of the solution of the LIE for the two systems under consideration coincide.

Assertion is proved.

Corollary. The formula for the average waiting time for a system with shifted distributions will have exactly the same form as for system with ordinary distributions, but with changed parameters due to a time shift operation [2–7]. Consequently, the average waiting time for systems with lag actually depends on the magnitude of the shift parameter $t_0 > 0$.

Now we define the numerical characteristics, and through them the unknown parameters of the distributions (11) and (12) by the method of moments. To do this, we write their Laplace transforms:

$$A^*(s) = [p \frac{\lambda_1}{\lambda_1 + s} + (1-p) \frac{\lambda_2}{\lambda_2 + s}] e^{-t_0 s},$$

$$B^*(s) = [q \left(\frac{2\mu_1}{2\mu_1 + s} \right)^2 + (1-q) \left(\frac{2\mu_2}{2\mu_2 + s} \right)^2] e^{-t_0 s}.$$

The first derivative of the function with a minus sign at the point $s = 0$ gives the values of the $\bar{\tau}_\lambda$

$$\bar{\tau}_\lambda = p\lambda_1^{-1} + (1-p)\lambda_2^{-1} + t_0, \quad (13)$$

and the second derivative at the same point gives the $\bar{\tau}_\lambda^2$:

$$\bar{\tau}_\lambda^2 = t_0^2 + 2t_0 \left[\frac{p}{\lambda_1} + \frac{(1-p)}{\lambda_2} \right] + 2 \left[\frac{p}{\lambda_1^2} + \frac{(1-p)}{\lambda_2^2} \right]. \quad (14)$$

Then the value c_λ^2

$$c_\lambda^2 = \frac{[(1-p^2)\lambda_1^2 - 2\lambda_1\lambda_2 p(1-q) + p(2-p)\lambda_2^2]}{[t_0\lambda_1\lambda_2 + (1-p)\lambda_1 + p\lambda_2]^2}. \quad (15)$$

By doing the same with distribution (12), we determine the corresponding characteristics for the service time

$$\bar{\tau}_\mu = q\mu_1^{-1} + (1-q)\mu_2^{-1} + t_0, \quad (16)$$

$$\bar{\tau}_\mu^2 = t_0^2 + 2t_0 \left[\frac{q}{\mu_1} + \frac{(1-q)}{\mu_2} \right] + \frac{3}{2} \left[\frac{q}{\mu_1^2} + \frac{(1-q)}{\mu_2^2} \right], \quad (17)$$

$$c_\mu^2 = \frac{\mu_1^2 - 2q\mu_2(\mu_1 - \mu_2) + q(1-2q)(\mu_1 - \mu_2)^2}{2[t_0\mu_1\mu_2 + (1-q)\mu_1 + q\mu_2]^2}. \quad (18)$$

The mechanism for determining the parameters of distributions (3), (4), (11) and (12) using both the first two initial moments and the three initial moments is described in detail in [2] and [3], respectively. Here we give ready-made expressions for these parameters. For distribution (11) unknown parameters are found by expressions

$$\begin{aligned} p &= \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{(\bar{\tau}_\lambda - t_0)^2}{2[(\bar{\tau}_\lambda - t_0)^2 + c_\lambda^2 \bar{\tau}_\lambda^2]}}, \quad \lambda_1 = 2p / (\bar{\tau}_\lambda - t_0), \\ \lambda_2 &= 2(1-p) / (\bar{\tau}_\lambda - t_0), \quad \text{and for distribution (12) -} \\ q &= \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{3(\bar{\tau}_\mu - t_0)^2}{8[(\bar{\tau}_\mu - t_0)^2 + c_\mu^2 \bar{\tau}_\mu^2]}}, \quad \mu_1 = 2q / (\bar{\tau}_\mu - t_0), \\ \mu_2 &= 2(1-q) / (\bar{\tau}_\mu - t_0). \end{aligned}$$

From these expressions it follows that the shift parameter is limited by the condition $t_0 < \bar{\tau}_\mu < \bar{\tau}_\lambda$. In addition, the range of applicability of the system is determined by the non-negativity of two expressions under the square root for p and q .

The algorithm for calculating the average waiting time for given input parameters is reduced to the sequential determination of the unknown parameters of distributions (11) and (12). Next, we determine the coefficients of polynomial (7) using the above expressions (6) and find the necessary roots with negative real parts

$-s_1, -s_2, -s_3, -s_4$. Substituting the absolute values of these roots in expression (10), we determine the average waiting time.

The presence of such roots is due to the existence and uniqueness of the spectral decomposition. Conducted numerous experiments only confirm this fact.

4 EXPERIMENTS

Tables 1 and 2 below show the calculation data in the Mathcad package for the usual $H_2/HE_2/1$ system and for the system $H_2^-/HE_2^-/1$ with a delay for cases of low, medium and high load $\rho = 0,1; 0,5; 0,9$ for a wide range of variation coefficients of variation c_λ, c_μ and for the shift parameter t_0 . Results for a usual system are compared with data for a close $H_2/H_2/1$ system. Dashes in table. 1 means that at such values of the parameters the $H_2/HE_2/1$ system is not applicable.

The results for the $H_2^-/HE_2^-/1$ system with a delay are compared with the results for the conventional system $H_2/HE_2/1$. The load factor ρ in both tables is determined by the ratio of average intervals $\rho = \bar{c}_\mu / \bar{c}_\lambda$. The calculations used the normalized service time $\bar{c}_\mu = 1$.

The results for the $H_2/HE_2/1$ and $H_2/H_2/1$ systems coincide to whole parts, but the range of service parameters for the first system is wider than that of the second.

The $H_2^-/HE_2^-/1$ s system is applicable for small values of the coefficients of variation, in particular, when $\rho=0,9$ $c_\lambda=0,2$ and $t_0=0,99$, the average delay is only a $\bar{W} = 0,187$ few units of time.

Table 1 – Results of experiments for QS $H_2/HE_2/1$ and $H_2/H_2/1$

Input parameters		Average delay	
ρ	(c_λ, c_μ)	For QS $H_2/HE_2/1$	For QS $H_2/H_2/1$
0,1	(1; 0.71)	0.086	–
	(1;1)	0.111	0.111
	(2;2)	0.446	0.445
	(4;4)	1.791	1.779
	(8;8)	7.173	7.112
0,5	(1; 0.71)	0.755	–
	(1;1)	1.000	1.000
	(2;2)	4.043	4.044
	(4;4)	16.235	16.129
	(8;8)	64.844	64.178
0,9	(1; 0.71)	6.771	–
	(1;1)	9.075	9.000
	(2;2)	36.169	36.200
	(4;4)	144.773	144.833
	(8;8)	577.875	577.861

Table 2– Results of experiments for QS $H_2^-/HE_2^-/1$ and $H_2/HE_2/1$

Input parameters		Average delay			
ρ	(c_λ, c_μ)	For QS $H_2^-/HE_2^-/1$			For QS $H_2/HE_2/1$
		$t_0=0.99$	$t_0=0.5$	$t_0=0.01$	
0.1	(1;0.71)	0.03	0.04	0.09	0.09
	(1;1)	0.06	0.07	0.11	0.11
	(2;2)	0.23	0.36	0.44	0.45
	(4;4)	0.93	1.56	1.79	1.79
	(8;8)	3.74	6.38	7.16	7.17
0.5	(1;0.71)	0.26	0.48	0.75	0.76
	(1;1)	0.51	0.75	0.99	1.00
	(2;2)	2.04	3.15	4.03	4.04
	(4;4)	8.15	12.73	16.17	16.24
	(8;8)	32.62	51.07	64.58	64.84
0.9	(1;0.71)	2.49	6.00	6.77	6.77
	(1;1)	4.73	8.29	9.06	9.08
	(2;2)	18.92	33.20	36.14	36.17
	(4;4)	75.69	123.39	144.63	144.77
	(8;8)	302.78	528.43	577.29	577.88

5 RESULTS

The paper presents the spectral expansions of the solution of the Lindley integral equation for the ordinary system $H_2/HE_2/1$ and the system with a delay, which are used to derive the formula for the average waiting time in the queue for these systems in closed form.

The operation of the shift in time on the one hand, leads to an increase in system load with a delay. The time shift operation, on the other hand, reduces the variation coefficients of the interval between receipts and the service time of requirements.

Because the average waiting time in the G/G/1 system is related to the coefficients of variation of the arrival and service time intervals with the quadratic dependence, the average waiting time in the delay system will be less than in the conventional system with the same load factor.

6 DISCUSSION

Thus, the range of change of parameters for the $H_2^-/HE_2^-/1$ system is much wider than that of the conventional system $H_2/HE_2/1$.

It is obvious that the average delay in a system with a delay depends on the shift parameter t_0 .

As one would expect, a decrease in the coefficients of variation and due to the introduction of the shift parameter into the laws of the distributions of the input flow and service time, entails a decrease in the average waiting time in systems with a delay several times. Thus, we expand the range of applicability of the $H_2/HE_2/1$ system in the teletraffic theory.

The full adequacy of the presented results is confirmed by the fact that when the shift parameter t_0 tends to zero, the average waiting time in a system with a delay tends to its value in a conventional system. In general, the adequacy of the presented mathematical models of systems is achieved by correctly using the classical method of spectral decomposition of LIE for the considered systems.

The results obtained extend the theory of systems G/G/1 and complement the well-known incomplete formula of queuing theory for the average waiting time.

CONCLUSIONS

In this work, the problem of deriving a formula for the average delay of requests in the queue for the considered systems is solved.

The scientific novelty of the results is that for the first time the spectral decomposition of the solution of the Lindley integral equation for the considered systems was obtained which are used to derive expression for the average waiting time in the queue for this system in closed form.

These expressions complements and expands the well-known incomplete formula for the average waiting time in the G/G/1 systems with arbitrary laws of input flow distribution and service time.

The practical significance of the work lies in the fact that the obtained results can be successfully applied in the modern theory of teletraffic, where the delays of incoming traffic packets play a primary role. For this, it is necessary to know the numerical characteristics of the incoming traffic intervals and the service time at the level of the first two moments, which does not cause difficulties when using modern traffic analyzers [10].

Prospects for further research are seen in the continuation of the study of systems of type G/G/1 with other common input distributions and in expanding and supplementing the formulas for average waiting time.

ACKNOWLEDGEMENTS

This work was carried out as part of the author's scientific school "Methods and Models for the Research of Computing Systems and Networks", registered at the Russian Academy of Natural Sciences on 31.03.2015 and was supported by the University of PSUTI.

REFERENCES

1. Kleinrock L. *Queueing Systems, Vol. I: Theory*. New York: Wiley, 1975, 417 p.
2. Tarasov V. N., Akhmetshina, E. G. The average waiting time in a H-2/H-2/1 queueing system with delay, *Journal of Samara State Technical University, Ser. Physical and Mathematical Sciences*, 2018, No. 2, pp. 702–713. DOI: <https://doi.org/10.14498/vsgtu1607>
3. Tarasov V. N. The analysis of two queueing systems HE2/M/1 with ordinary and shifted input distributions, *Radio Electronics, Computer Science, Control*, 2019, Vol. 49, No. 2, pp. 71–79. DOI: 10.15588/1607-3274-2019-2-8
4. Tarasov V. N. Queueing systems with delay, *Radio Electronics, Computer Science, Control*, 2019, Vol. 50, No. 3, pp. 55–63. DOI: 10.15588/1607-3274-2019-3-7
5. Tarasov V. N., Bakhareva N. F. Comparative analysis of two queueing systems M/HE2/1 with ordinary and with the shifted input distributions, *Radio Electronics, Computer Science, Control*, 2019, Vol. 51, No. 4, pp. 50–58. DOI: 10.15588/1607-3274-2019-4-5
6. Tarasov V. N. Analysis of H-2/E-2/1 system and her of the analog with shifted input distributions, *Radio Electronics, Computer Science, Control*, 2020, Vol. 52, No. 1, pp. 90–97. DOI: 10.15588/1607-3274-2020-1-10
7. Tarasov V. N., Bakhareva N. F. Model of teletraffic based on queueing systems E-2/HE2/1 with ordinary and shifted input distributions, *Radio Electronics, Computer Science, Control*, 2020, Vol. 53, No. 2, pp. 65–74. DOI: 10.15588/1607-3274-2020-2-7
8. Brannstrom N. A. Queueing Theory analysis of wireless radio systems. Applied to HS-DSCH. Lulea university of technology, 2004, 79 p.
9. Whitt W. Approximating a point process by a renewal process: two basic methods, *Operation Research*, 1982, Vol. 30, No. 1, pp. 125–147.
10. Novitzky S., Pender J., Rand R. H., Wesson E. Limiting the oscillations in queues with delayed information through a novel type of delay announcement. *Queueing Systems*, 2020, Vol. 95, P. 281–330.
11. Novitzky S., Pender J., Rand R. H., Wesson E. Nonlinear Dynamics in Queueing Theory: Determining the Size of Oscillations in Queues with Delay. *SIAM J. Appl. Dyn. Syst.*, 18–1, 2019, Vol. 18, No. 1, pp. 279–311. DOI: <https://doi.org/10.1137/18M1170637>
12. RFC 3393 [IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)] Available at: <https://tools.ietf.org/html/rfc3393>. (accessed: 26.02.2016).
13. Myskja A. An improved heuristic approximation for the GI/GI/1 queue with bursty arrivals. *Teletraffic and datatraffic in a Period of Change*. ITC-13. Elsevier Science Publishers, 1991, pp. 683–688.
14. Aliev T. I. Approximation of Probability Distributions in Queueing Models, *Scientific and technical bulletin of information technologies, mechanics and optics*, 2013, No. 2, pp. 88–93.
15. Aras A. K., Chen X. & Liu Y. Many-server Gaussian limits for overloaded non-Markovian queues with customer abandonment, *Queueing Systems*, 2018, Vol. 89, No. 1, pp. 81–125. DOI: <https://doi.org/10.1007/s11134-018-9575-0>
16. Jennings O. B. & Pender J. Comparisons of ticket and standard queues, *Queueing Systems*, 2016, Vol. 84, No. 1, pp. 145–202. DOI: <https://doi.org/10.1007/s11134-016-9493-y>
17. Gromoll H. C., Terwilliger B. & Zwart B. Heavy traffic limit for a tandem queue with identical service times, *Queueing Systems*, 2018, Vol. 89, No. 3, pp. 213–241. DOI: <https://doi.org/10.1007/s11134-017-9560-z>
18. Legros B. M/G/1 queue with event-dependent arrival rates, *Queueing Systems*, 2018, Vol. 89, No. 3, pp. 269–301. DOI: <https://doi.org/10.1007/s11134-017-9557-7>
19. Bazhba M., Blanchet J., Rhee CH., et al. Queue with heavy-tailed Weibull service times, *Queueing Systems*, 2019, Vol. 93, No. 11, pp. 1–32. DOI: <https://doi.org/10.1007/s11134-019-09640-z/>
20. Adan I., D'Auria B., Kella O. Special volume on 'Recent Developments in Queueing Theory' of the third ECQT conference. *Queueing Systems*, 2019, Vol. 93, No. 1, pp. 1–190. DOI: <https://doi.org/10.1007/s11134-019-09630-1>
21. Adan I., D'Auria B., Kella O. Special volume on 'Recent Developments in Queueing Theory' of the third ECQT conference: part 2, *Queueing Systems*, 2019, pp. 1–2. DOI: <https://doi.org/10.1007/s11134-019-09637-8>
22. Tibi D. Martingales and buffer overflow for the symmetric shortest queue model. *Queueing Systems*, Vol. 93, 2019, pp. 153–190. DOI: 10.1007/s11134-019-09628-9

23. Jacobovic R., Kella O. Asymptotic independence of regenerative processes with a special dependence structure. *Queueing Systems*, Vol. 93, 2019, pp. 139–152. DOI: 10.1007/s11134-019-09606-1
24. Wang L., Kulkarni V. Fluid and diffusion models for a system of taxis and customers with delayed matching. *Queueing Systems*, 2020, Vol. 96, pp. 101–131. DOI: 10.1007/s11134-020-09659-7

Received 19.01.2021.
Accepted 18.04.2021.

УДК 621.391.1:621.395

МОДЕЛІ ЗАТРИМКИ НА БАЗІ СИСТЕМ З ЗВИЧАЙНИМИ ТА З ЗСУНУТИМИ ГІПЕРЕКСПОНЕНТНИМ ТА ГІПЕРЕРЛАНГІВСЬКИМ ВХІДНИМИ РОЗПОДІЛАМИ

Тарасов В. Н. – д-р техн. наук, професор, завідувач кафедри програмного забезпечення та управління в технічних системах Поволзького державного університету телекомунікацій та інформатики, РФ.

Бахарєва Н. Ф. – д-р техн. наук, професор, завідувач кафедри інформатики та обчислювальної техніки Поволзького державного університету телекомунікацій та інформатики, РФ.

АНОТАЦІЯ

Актуальність. У теорії масового обслуговування дослідження систем довільними законами розподілів вхідного потоку і часу обслуговування актуальні в зв'язку з тим, що не можна отримати рішення для затримки в кінцевому вигляді в загальному випадку при довільних законах розподілів вхідного потоку і часу обслуговування. Тому в сучасній теорії телетрафіка важливі дослідження таких систем для окремих випадків вхідних розподілів.

Мета роботи. Отримання рішення для середньої затримки в черзі в замкнутій формі для систем масового обслуговування зі звичайними і з зсунутими вправо від нульової точки розподілами в сталому режимі.

Метод. Для вирішення поставленого завдання був використаний класичний метод спектрального розкладання рішення інтегрального рівняння Ліндли. Цей метод дозволяє отримати рішення для середнього часу очікування для розглянутих систем в замкнутій формі. Метод спектрального розкладання рішення інтегрального рівняння Ліндли грає важливу роль в теорії систем G/G/1. Для практичного застосування отриманих результатів було використано відомий метод моментів теорії ймовірностей.

Результати. Вперше отримано спектральні розкладання рішення інтегрального рівняння Ліндли для систем зі звичайними та з зсунутими гіперекспоненціальними і гіперерлангівськими розподілами, за допомогою якого виведено розрахункове вираз для середньої затримки в черзі в замкнутій формі.

Висновки. Доведено, що спектральні розкладання рішення інтегрального рівняння Ліндли для розглянутих систем збігаються, тому формули для середньої затримки також будуть збігатися. Показано, що в системах з запізненням у часі середній час очікування менше, ніж в звичайних системах. Отримане розрахункове вираз для часу очікування розширює і доповнює відому незавершену формулу теорії масового обслуговування для середньої затримки для систем з довільними законами розподілів вхідного потоку і часу обслуговування. Такий підхід дозволяє розрахувати середньої затримки для зазначених систем в математичних пакетах для широкого діапазону зміни параметрів трафіку. Крім середньої затримки, такий підхід дає можливість також визначити моменти вищих порядків часу очікування. З огляду на той факт, що варіація затримки пакетів (джиттер) в телекомунікації визначається як дисперсія часу очікування від його середнього значення, то джиттер можна буде визначити через дисперсію затримки.

КЛЮЧОВІ СЛОВА: система з запізненням, зсунуті розподілу, перетворення Лапласа, інтегральне рівняння Ліндли, метод спектрального розкладання.

УДК 621.391.1:621.395

МОДЕЛИ ЗАДЕРЖКИ НА БАЗЕ СИСТЕМ С ОБЫЧНЫМИ И СДВИНУТЫМИ ГИПЕРЭКСПОНЕНЦИАЛЬНЫМ И ГИПЕРЭРЛАНГОВСКИМ ВХОДНЫМИ РАСПРЕДЕЛЕНИЯМИ

Тарасов В. Н. – д-р техн. наук, профессор, заведующий кафедрой программного обеспечения и управления в технических системах Поволжского государственного университета телекоммуникаций и информатики, РФ.

Бахарєва Н. Ф. – д-р техн. наук, профессор, заведующая кафедрой информатики и вычислительной техники Поволжского государственного университета телекоммуникаций и информатики, РФ.

АННОТАЦИЯ

Актуальность. В теории массового обслуживания исследования систем произвольными законами распределений входного потока и времени обслуживания актуальны в связи с тем, что нельзя получить решения для времени ожидания в конечном виде в общем случае при произвольных законах распределений входного потока и времени обслуживания. Поэтому в современной теории телетрафика важны исследования таких систем для частных случаев входных распределений.

Цель работы. Получение решения для средней задержки в очереди в замкнутой форме для систем массового обслуживания с обычными и с сдвинутыми вправо от нулевой точки гиперэкспоненциальными и гиперэрланговскими распределениями в установившемся режиме.

Метод. Для решения поставленной задачи использован классический метод спектрального разложения решения интегрального уравнения Линдли. Данный метод позволяет получить решение для средней задержки для рассматриваемых систем в замкнутой форме. Метод спектрального разложения решения интегрального уравнения Линдли играет важную

роль в теории систем G/G/1. Для практического применения полученных результатов использован известный метод моментов теории вероятностей.

Результаты. Впервые получено спектральное разложение решения интегрального уравнения Линдли для систем с обычными и с сдвинутыми гиперэкспоненциальными и гиперэрланговскими распределениями, с помощью которого выведено расчетная формула для средней задержки в очереди в замкнутой форме.

Выводы. Доказано, что спектральные разложения решения интегрального уравнения Линдли для рассматриваемых систем совпадают, поэтому формулы для средней задержки также будут совпадать. Показано, что в системах с запаздыванием во времени средняя задержка меньше, чем в обычных системах. Полученное расчетное выражение для времени ожидания расширяет и дополняет известную незавершенную формулу теории массового обслуживания для среднего времени ожидания для систем с произвольными законами распределений входного потока и времени обслуживания. Такой подход позволяет рассчитать среднее время ожидания для указанных систем в математических пакетах для широкого диапазона изменения параметров трафика. Кроме среднего времени ожидания, имеется возможность определения и моментов высших порядков времени ожидания. Учитывая тот факт, что вариация задержки пакетов (джиттер) в телекоммуникациях определяется как разброс времени ожидания вокруг его среднего значения, то джиттер можно будет определить через дисперсию времени ожидания.

КЛЮЧЕВЫЕ СЛОВА: система с запаздыванием, сдвинутые распределения, преобразование Лапласа, интегральное уравнение Линдли, метод спектрального разложения.

ЛИТЕРАТУРА / LITERATURA

1. Kleinrock L. *Queueing Systems, Vol. I: Theory* / L. Kleinrock. – New York: Wiley, 1975. – 417 p.
2. Tarasov V. N. The average waiting time in a H-2/H-2/1 queueing system with delay / V. N. Tarasov, E. G. Akhmetshina // *Journal of Samara State Technical University, Ser. Physical and Mathematical Sciences*. – 2018. – No. 2. – P. 702–713.
3. Tarasov V. N. The analysis of two queueing systems HE2/M/1 with ordinary and shifted input distributions / V. N. Tarasov // *Radio Electronics, Computer Science, Control*. – 2019. – Vol. 49, No. 2. – P. 71–79. DOI: 10.15588/1607-3274-2019-2-8
4. Tarasov V. N. Queueing systems with delay / V. N. Tarasov // *Radio Electronics, Computer Science, Control*. – 2019. – Vol. 50, No. 3. – P. 71–79. DOI: 10.15588/1607-3274-2019-3-7
5. Tarasov V. N. Comparative analysis of two queueing systems M/HE2/1 with ordinary and with the shifted input distributions / V. N. Tarasov, N. F. Bakhareva // *Radio Electronics, Computer Science, Control*. – 2019. – Vol. 51. – No. 4. – P. 50–58. DOI: 10.15588/1607-3274-2019-4-5
6. Tarasov V. N. Analysis of H-2/E-2/1 system and her of the analog with shifted input distributions / B. H. Тарасов // *Radio Electronics, Computer Science, Control*. – 2020. – Vol. 52, No. 1. – P. 90–97. DOI: 10.15588/1607-3274-2020-1-10
7. Tarasov V. N. Model of teletraffic based on queueing systems E-2/HE2/1 with ordinary and shifted input distributions / V. N. Tarasov, N. F. Bakhareva // *Radio Electronics, Computer Science, Control*. – 2020. – Vol. 53, No. 2. – P. 65–74. DOI: 10.15588/1607-3274-2020-2-7
8. Brannstrom N. *A Queueing Theory analysis of wireless radio systems / N. Brannstrom – Applied to HS-DSCH*. Lulea university of technology, 2004. – 79 p.
9. Whitt W. Approximating a point process by a renewal process: two basic methods / W. Whitt // *Operation Research*. – 1982. – № 1. – P. 125–147.
10. Limiting the oscillations in queues with delayed information through a novel type of delay announcement / [S. Novitzky, J. Pender, R. H. Rand, E. Wesson] // *Queueing Systems*. – 2020. – Vol. 95. – P. 281–330. DOI: <https://doi.org/10.1007/s11134-020-09657-9>
11. Nonlinear Dynamics in Queueing Theory: Determining the Size of Oscillations in Queues with Delay / S. Novitzky, J. Pender, R. H. Rand, E. Wesson // *SIAM J. Appl. Dyn. Syst.*, 18–1. – 2019. – Vol. 18, No. 1. – P. 279–311. DOI: <https://doi.org/10.1137/18M1170637>
12. [HTTPS://tools.ietf.org/html/rfc3393](https://tools.ietf.org/html/rfc3393). RFC 3393 IP Packet Delay Variation Metric for IP Performance Metrics (IPPM) (дата обращения: 26.02.2016).
13. Myskja A. An improved heuristic approximation for the GI/GI/1 queue with bursty arrivals / A. Myskja // *Teletraffic and datatraffic in a Period of Change, ITC-13*. Elsevier Science Publishers. – 1991. – P. 683–688.
14. Aliev T. I. Approximation of Probability Distributions in Queueing Models / T. I. Aliev // *Scientific and technical bulletin of information technologies, mechanics and optics*. – 2013. – No. 2. – P. 88–93.
15. Aras A. K. Many-server Gaussian limits for overloaded non-Markovian queues with customer abandonment / A.K. Aras, X. Chen, Y. Liu // *Queueing Systems*. – 2018. – Vol. 89, No. 1. – P. 81–125. DOI: <https://doi.org/10.1007/s11134-018-9575-0>
16. Jennings O. B. Comparisons of ticket and standard queues / O. B. Jennings, J. Pender // *Queueing Systems*. – 2016. – Vol. 84, No. 1. – P. 145–202. DOI: <https://doi.org/10.1007/s11134-016-9493-y>
17. Gromoll H. C., Terwilliger B. & Zwart B. Heavy traffic limit for a tandem queue with identical service times / H. C. Gromoll, B. Terwilliger, B. Zwart // *Queueing Systems*. – 2018. – Vol. 89, No. 3. – P. 213–241. DOI: <https://doi.org/10.1007/s11134-017-9560-z>
18. Legros B. M/G/1 queue with event-dependent arrival rates / B. Legros // *Queueing Systems*. – 2018. – Vol. 89, No. 3. – P. 269–301. DOI: <https://doi.org/10.1007/s11134-017-9557-7>
19. Bazhba M. Queue with heavy-tailed Weibull service times / M. Bazhba, J. Blanchet, C. H. Rhee // *Queueing Systems*. – 2019. – Vol. 93, No. 11. – P. 1–32. DOI: <https://doi.org/10.1007/s11134-019-09640-z/>
20. Adan I. Special volume on ‘Recent Developments in Queueing Theory’ of the third ECQT conference / I. Adan, B. D’Auria, O. Kella // *Queueing Systems*. – 2019. – Vol. 93, No. 1. – P. 1–190. DOI: <https://doi.org/10.1007/s11134-019-09630-1>
21. Adan I. Special volume on ‘Recent Developments in Queueing Theory’ of the third ECQT conference: part 2 / I. Adan, B. D’Auria, O. Kella // *Queueing Systems*. – 2019. – P. 1–2. DOI: <https://doi.org/10.1007/s11134-019-09637-8>
22. Tibi D. Martingales and buffer overflow for the symmetric shortest queue model / D. Tibi // *Queueing Systems*. – 2019. – Vol. 93. – P. 153–190. DOI: 10.1007/s11134-019-09628-9
23. Jacobovic R. Asymptotic independence of regenerative processes with a special dependence structure / R. Jacobovic, O. Kella // *Queueing Systems*. – 2019. – Vol. 93. – P. 139–152. DOI: 10.1007/s11134-019-09606-1
24. Wang L. Fluid and diffusion models for a system of taxis and customers with delayed matching / L. Wang, V. Kulkarni // *Queueing Systems*. – 2020. – Vol. 96. – P. 101–131. DOI: 10.1007/s11134-020-09659-7

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

УДК 004.8:004.032.26

ОНЛАЙН МЕТОД МОЖЛИВІСНОЇ КЛАСТЕРИЗАЦІЇ ДАНИХ НА ОСНОВІ ЕВОЛЮЦІЙНОЇ ОПТИМІЗАЦІЇ КОТЯЧИХ ЗГРАЙ

Бодяньський С. В. – д-р техн. наук, професор, професор кафедри штучного інтелекту, Харківський національний університет радіоелектроніки, Харків, Україна.

Шафроненко А. Ю. – канд. техн. наук, доцент, доцент кафедри інформатики, Харківський національний університет радіоелектроніки, Харків, Україна.

Клімова І. М. – асистент кафедри системної інженерії Харківського національного університету радіоелектроніки, Харків, Україна.

АНОТАЦІЯ

Актуальність. Проблема кластеризації великих даних сьогодні є дуже цікавою сферою штучного інтелекту. Це завдання часто зустрічається у багатьох додатках, пов'язаних з інтелектуальним аналізом даних, глибоким навчанням, веб-майнінгом тощо. Для вирішення цих проблем традиційні підходи та методи вимагають, щоб уся вибірка даних подавалася у пакетній формі.

Мета. Метою роботи було запропонувати метод нечіткої можливісної кластеризації даних з використанням еволюційної оптимізації котячих зграй, який був би позбавлений недоліків традиційних підходів до кластеризації даних.

Метод. Процедура нечіткої можливісної кластеризації даних із використанням еволюційних алгоритмів, для більш швидкого визначення екстремумів вибірки, центрів кластерів та адаптивних функцій, що дозволяють не витратити ресурси машини для зберігання проміжних розрахунків та не потребують додаткового часу для вирішення поставленої задачі кластеризації даних незалежно від розмірності та способу подачі на обробку.

Результати. Запропонований алгоритм кластеризації даних на основі еволюційної оптимізації є достатньо простим у чисельній реалізації, позбавлений недоліків, властивих традиційним методам нечіткої кластеризації та може працювати з великим розміром вихідної інформації, що обробляється, в режимі онлайн у реальному часі.

Висновки. Результати експерименту дозволяють рекомендувати розроблений метод для вирішення проблем автоматичної кластеризації та класифікації великих даних, максимально швидко знаходити екстремуми цільової функції, незалежно від способу подачі даних на обробку. Запропонований метод онлайн можливісної нечіткої кластеризації даних на основі еволюційної оптимізації котячих зграй призначений для використання у гібридних системах обчислювального інтелекту, нейро-фаззі системах, в навчанні штучних нейронних мереж, у завданнях кластеризації та класифікації.

КЛЮЧОВІ СЛОВА: нечітка кластеризація, правило навчання, оптимізація котячої зграї, режим трасування, режим пошуку.

АБРЕВІАТУРИ

FCM – метод нечітких с-середніх;

CSO – алгоритм оптимізації зграї котів.

НОМЕНКЛАТУРА

X – матриця набору даних;

k – номер вектору-спостереження;

i – номер атрибуту вектору-спостереження;

j – номер неперетинного класу;

$x(k)$ – вектор-спостереження;

$x_i(k)$ – значення попередньо обробленого вихідного k -го вектора-спостереження для i -го атрибута;

l, q – номер кластеру;

m – кількість неперетинних класів;

$U_j(k)$ – рівень належності j -го кластеру;

Cl – кластер;

D – Евклідова відстань;

E – цільова функція;

μ – скалярний параметр, який визначає відстань, на якій рівень належності дорівнює 0,5;

c – центроїд кластера;

c_p – режим p -ої кішки;
 τ – ітерація пошуку;
 β – фаззифікатор;
 α – параметр, який визначає властивості інерції режиму трасування;
 η – параметр крока пошуку;
 $\Xi(\tau)$ – випадкова складова, яка вносить додаткові стохастичні рухи у процес трасування;
 η_ξ – параметр, який визначає амплітуду рухів.

ВСТУП

В цей час методи обчислювального інтелекту широко використовуються для вирішення багатьох складних задач аналізу даних. Однією з основних областей обчислювального інтелекту є, так звані, еволюційні алгоритми, які за суттю представляють певні математичні моделі розвитку біологічних організмів.

Проблема аналізу даних, що пов'язана з кластеризацією векторних образів, часто зустрічається у багатьох додатках для інтелектуального аналізу, але останнім часом основна увага приділяється нечіткій кластеризації [1–3] при обробці векторних образів з різним рівнем ймовірності, можливості або належності можуть належати більше ніж одному класу.

Самоорганізовані мапи Кохонена [4] є дуже ефективним засобом узгодженої онлайн-кластеризації, використання якої дозволяє вирішувати завдання в режимі онлайн. Еволюційні алгоритми у цьому випадку дозволяють покращити результати кластеризації даних коли ці дані надходять послідовно у онлайн-режимі.

Об'єкт дослідження можливісна кластеризація даних в режимі онлайн.

Предмет дослідження процедура онлайн можливоїсної кластеризація даних на основі еволюційної оптимізації котячих зграй.

Мета роботи полягає у запровадженні можливоїсної процедури нечіткої кластеризації даних, яка може працювати онлайн, та швидко знаходити екстемуми (центри) кластерів, незалежно від кількості даних, що надходять.

1 ПОСТАНОВКА ЗАВДАННЯ

Проблема нечіткої кластеризації масивів даних розглядається в умовах, коли сформовані кластери довільно перетинаються в просторі ознак. Вихідною інформацією для вирішення проблеми є масив багатовимірних векторів даних, сформований на вибірці спостережень $X = \{x(1), \dots, x(k), \dots, x(N)\} \subset R^n$, де k – у загальному випадку номер спостереження у вихідному масиві, $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T \in R^n$. Результатом кластеризації є розділ цього масиву на m неперетинних класів Cl_j з прототипами-

центроїдами $Cl_j \in R^n$, $j = 1, 2, \dots, m$ та обчислення рівнів належності $0 \leq U_j(k) \leq 1$ кожного спостереження $x(k)$ до кожного кластера Cl_j .

2 ОГЛЯД ЛІТЕРАТУРИ

Основний недолік класичних імовірнісних алгоритмів кластеризації [1–3] пов'язаний з обмеженнями рівнів належності, сума яких повинна дорівнювати одиниці. Така ситуація призвела до створення можливоїсних нечітких алгоритмів кластеризації [5, 6].

Цільова функція можливоїсної кластеризації має вигляд

$$E(U_j(k), c_j, \mu_j) = \sum_{k=1}^N \sum_{j=1}^m U_j^\beta(k) D^2(\tilde{x}_k, c_j) + \sum_{j=1}^m \mu_j \sum_{k=1}^N (1 - U_j(k))^\beta. \quad (1)$$

Мінімізація функції (1) за $U_j(k)$, c_j та μ_j веде до системи рівнянь

$$\begin{cases} U_j^{(\tau+1)}(k) = \frac{1}{1 + \left(\frac{D^2(\tilde{x}(k), c_j^{(\tau)})}{\mu_j^{(\tau)}}\right)^{\frac{1}{\beta-1}}}, \\ c_j^{(\tau+1)} = \frac{\sum_{k=1}^N (U_j^{(\tau+1)}(k))^\beta \tilde{x}(k)}{\sum_{k=1}^N (U_j^{(\tau+1)}(k))^\beta}, \\ \mu_j^{(\tau+1)} = \frac{\sum_{k=1}^N (U_j^{(\tau+1)}(k))^\beta D^2(\tilde{x}(k), c_j^{(\tau+1)})}{\sum_{k=1}^N (U_j^{(\tau+1)}(k))^\beta}, \end{cases} \quad (2)$$

у випадку, коли $\beta = 2$ ми приходимо до рішення, яке має вигляд (аналог класичного FCM [1]):

$$\begin{cases} U_j^{(\tau+1)}(k) = \frac{1}{1 + \frac{\|\tilde{x}(k) - c_j^{(\tau)}\|^2}{\mu_j^{(\tau)}}}, \\ c_j^{(\tau+1)} = \frac{\sum_{k=1}^N (U_j^{(\tau)}(k))^2 \tilde{x}(k)}{\sum_{k=1}^N (U_j^{(\tau)}(k))^2}, \\ \mu_j^{(\tau+1)} = \frac{\sum_{k=1}^N (U_j^{(\tau)}(k))^2 \|\tilde{x}(k) - c_j^{(\tau+1)}\|^2}{\sum_{k=1}^N (U_j^{(\tau)}(k))^2}. \end{cases} \quad (3)$$

В режимі онлайн формули (2), (3) можна записати у вигляді [6, 7]

$$\left\{ \begin{aligned} U_j(k+1) &= \frac{1}{1 + \left(\frac{D^2(\tilde{x}(k+1), c_j(k))}{\mu_j(k)} \right)^{\frac{1}{\beta-1}}}, \\ c_j(k+1) &= c_j(k) + \eta(k+1)U_j^\beta(k+1)(\tilde{x}(k+1) - c_j(k)), \\ \mu_j(k+1) &= \frac{\sum_{p=1}^{k+1} U_j^\beta(p) D^2(\tilde{x}(p), c_j(k+1))}{\sum_{p=1}^{k+1} U_j^\beta(p)} \end{aligned} \right. \quad (4)$$

або

$$\left\{ \begin{aligned} U_j(k+1) &= \frac{1}{1 + \frac{\|\tilde{x}(k) - c_j(k)\|^2}{\mu_j(k)}}, \\ c_j(k+1) &= c_j(k) + \eta(k+1)U_j^2(k+1)(\tilde{x}(k+1) - c_j(k)), \\ \mu_j(k+1) &= \frac{\sum_{p=1}^{k+1} U_j^2(p) \|\tilde{x}(p) - c_j(k+1)\|^2}{\sum_{p=1}^k U_j^2(p)}, \end{aligned} \right. \quad (5)$$

що дозволяє вирішувати нечітку задачу кластеризації в режимі онлайн.

Для пошуку глобального екстремуму (1) доцільно використовувати, так звані, алгоритми оптимізації рою частинок, а саме еволюційних частинок [8], що розробляються та досліджуються в цей час в рамках загальної теорії та практики обчислювального інтелекту. Серед алгоритмів роїння одним з найшвидших є так звані алгоритми зграї котів [9–12], які виявилися ефективними у вирішенні широкого кола завдань з обробки даних.

3 МАТЕРІАЛИ І МЕТОДИ

Оптимізація котячої зграї передбачає, що кожна кішка зграї може перебувати в одному з двох станів: режимі пошуку та режимі трасування (відстеження). У цьому випадку режим пошуку пов'язаний з повільними рухами з невеликою амплітудою навколо початкової позиції (сканування простору в районі поточної позиції) та режимом трасування, який визначається швидкими стрибками з великою амплітудою і дозволяє кожній конкретній кішці вискочити з локального екстремуму, якщо вона там є.

Поєднання локального сканування та швидких змін поточного стану дозволяє збільшити ймовірність знаходження глобального екстремуму порівняно з традиційними методами багатоекстремальної оптимізації.

У загальному випадку обидва ці режими для кожного із зграї котів можуть бути описані рекурентною процедурою оптимізації

$$c_p(\tau+1) = c_p(\tau) - \alpha(c_p(\tau) - c_p(\tau-1)) - \eta \hat{\nabla} E_M(c_p(\tau)) + \eta_\xi \Xi(\tau). \quad (6)$$

Так, коли $\alpha = 0$, приходимо до стандартного градієнтного пошуку, тобто режиму пошуку у зграї, при $0 < \alpha < 1$ процес оптимізації набуває інерційних властивостей типу «важкої кулі», але він може не зупинятися поблизу глибокого екстремуму.

Таким чином, кожна кішка може одночасно знаходитись у режимі пошуку та відстеження та, маючи достатню кількість котів у зграї, можна забезпечити пошук глобального екстремуму.

4 ЕКСПЕРИМЕНТИ

Онлайн алгоритм можливої нечіткої кластеризації на основі еволюційної оптимізації котячого рою був реалізований на двох масивах даних: табл. 1. Порівняльний аналіз кластеризації вихідних даних був проведений за допомогою нечітких С-середніх, адаптивного можливої нечіткого методу, метода Гауфсона-Кесселя та можливої нечіткої кластеризації на основі еволюційної оптимізації котячої зграї. Параметри моделювання представлені у табл. 2.

Таблиця 1 – Зразки даних

Data set	Data Number	Attributes Number	Cluster Number	Data Source
Gas	296	2	6	Box and Jenkins (1970)
Glass	214	9	6	Mackey and Glass (1977)

Таблиця 2 – Параметри алгоритму CSO

Parameters	Value
SRD	Random [0,1]
Seeking memory Pool (SMP)	5
Population size	Number of clusters
r1	Random in [0,1]
c1	Const
SPC	Random in [0,1]
Number of iteration	Manually

5 РЕЗУЛЬТАТИ

Було проведено порівняльний аналіз якості даних кластеризації за основними характеристиками рейтингів якості таких, як швидкість кластеризації даних та середня похибка.

В табл. 3 та табл. 4 наведені результати роботи алгоритмів, запропонованих для порівняння з різною кількістю спостережень.

Таблиця 3 – Порівняльні характеристики середньої похибки для вибірки Gas з різною кількістю спостережень в відсотках

Метод	50	Час	100	Час	150	Час
FCM	1,62	1,19	1,35	2,55	0,98	3,03
GK	1,66	1,62	1,32	2,72	0,99	3,12
Адаптивний можливої нечіткої кластеризації	1,22	1,15	1,02	2,02	0,75	2,10
Онлайн алгоритм можливої нечіткої кластеризації даних на основі еволюційної оптимізації котячих зграї	0,69	1,02	0,49	1,33	0,14	1,41

Таблиця 4 – Порівняльні характеристики середньої похибки для вибірки Glass з різною кількістю спостережень у відсотках

Метод	50	Час	100	Час	150	Час
FCM	1,74	1,21	1,44	2,40	0,86	3,10
GK	1,85	1,74	1,53	2,82	0,99	3,27
Адаптивний можливістьний метод нечіткої кластеризації	1,43	1,36	1,22	2,55	0,65	2,60
Онлайн алгоритм можливістьної кластеризації даних на основі еволюційної оптимізації котячих зграй	1,11	1,17	1,00	1,23	0,54	1,11

Аналізуючи отримані результати, можна зробити висновок, що незалежно від розміру вихідної інформації, що подається на обробку запропонованим методом для порівняння працездатності та ефективності, запропонований підхід до можливістьної кластеризації даних на основі еволюційного методу котячих зграй не поступається швидкодією та якістю кластерування у порівнянні з відомими алгоритмами.

Порівняльний аналіз продемонстровано на діаграмах залежності похибки та часу від кількості спостережень для різних вибірок даних.

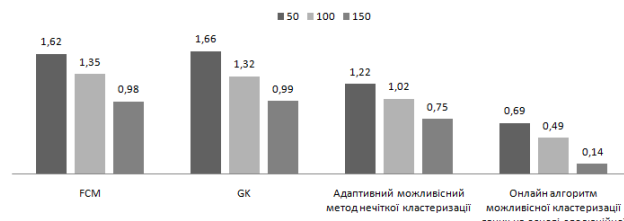


Рисунок 1 – Діаграма залежності похибки від кількості спостережень (50, 100, 150) для вибірки Gas



Рисунок 2 – Діаграма залежності похибки від кількості спостережень (50, 100, 150) для вибірки Glass

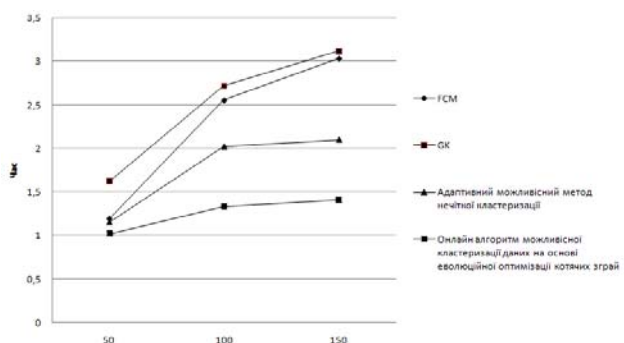


Рисунок 3 – Діаграма залежності часу кластеризації від кількості спостережень (50, 100, 150) для вибірки Gas

© Бодяньський С. В., Шафроненко А. Ю., Клімова І. М., 2021
DOI 10.15588/1607-3274-2021-2-7

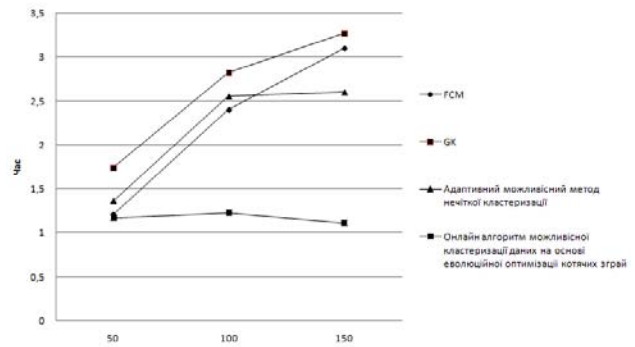


Рисунок 4 – Діаграма залежності часу кластеризації від кількості спостережень (50, 100, 150) для вибірки Glass

Як видно із діаграм, наведених на рисунках, швидкість та похибка в запропонованому методі можливістьної кластеризації даних на основі еволюційного підходу демонструє достатньо високі показники.

6 ОБГОВОРЕННЯ

Результати кластеризації наборів даних, наведених в таблицях, демонструють швидкодію та якість роботи методів кластеризації. Запропонований онлайн алгоритм можливістьної кластеризації даних на основі еволюційної оптимізації котячих зграй демонструє гарні результати роботи.

Порівняльний аналіз запропонованого алгоритма проводився з відомими на сьогодні алгоритмами кластеризації такими, як FCM, Густафсон-Кессель-алгоритм і адаптивний можливістьний метод нечіткої кластеризації.

На рисунках, що наведені вище, продемонстрована робота алгоритмів у порівнянні із запропонованим онлайн алгоритмом можливістьної кластеризації даних на основі еволюційної оптимізації котячих зграй.

Завдяки своїй адаптивності та функціям еволюційної оптимізації котячих зграй, алгоритм не потребує багато часу для обробки даних, що надходять у реальному часі, та не завантажує себе проміжними розрахунками за рахунок функцій адаптивності. Це досить яскраво демонструють діаграми залежності часу кластеризації від кількості спостережень та залежності похибки від кількості спостережень.

ВИСНОВКИ

Запропонований підхід є досить простим в чисельній реалізації, має високу продуктивність і забезпечує високу якість нечіткої кластеризації великих масивів. Показано, що задачу можна вирішити на основі самоорганізовної нейро-нечіткої мережі Кохонена «Переможець отримує більше» для нечіткої кластеризації та алгоритмів оптимізації котячої зграї.

Наукова новизна: вперше запропонована можливістьна нечітка кластеризація даних, яка може працювати онлайн, та швидко знаходити екстемуми (центри) кластерів, не залежно від кількості даних, що надходять.

Практичне значення: результати експерименту дозволяють рекомендувати запропоновані методи для використання на практиці для вирішення проблем автоматичної кластеризації великих даних.

Перспективи подальших досліджень методи нечіткої кластеризації даних для широкого класу практичних проблем.

ПОДЯКА

Робота виконана в рамках науково-дослідного проекту державного бюджету Харківського національного університету радіоелектроніки «Глибокі гібридні системи обчислювального інтелекту для аналізу потоків даних та їх швидке навчання» (номер державної реєстрації 0119U001403).

ЛІТЕРАТУРА / LITERATURA

1. Bezdek J. C. Pattern recognition with fuzzy objective function algorithms / J. C. Bezdek. – New York : Springer, 1981. – 253p. DOI <https://doi.org/10.1007/978-1-4757-0450-1>.
2. Höppner F. Fuzzy Clustering Analysis: Methods for Classification, Data Analysis and Image Recognition / F. Höppner, F. Klawonn, R. Kruse, T. Runkler. – Chichester : John Wiley & Sons, 1999. – 300 p.
3. Xu R. Clustering / R. Xu, D. C. Wunsch. – Hoboken N. J.: John Wiley & Sons, Inc., 2009. – 398 p.
4. T. Kohonen. Self-Organizing Maps / T. Kohonen. – Berlin : Springer-Verlag, 1995.
5. Krishnapuram R. A Possibilistic Approach to Clustering / R. Krishnapuram, J. M. Keller // IEEE Transactions on

- Fuzzy Systems, May 1993: Proceedings. – IEEE, 1993. – Vol. 1. – P. 98–110. DOI: 10.1109/91.227387.
6. Bodyanskiy, Ye. Computational intelligence techniques for data analysis / Ye. Bodyanskiy // Lecture Notes in Informatics. – Bonn : Gesellschaft für Informatik, 2005. – P. 15–36.
7. Grosan C. Swarm intelligence in Data Mining / C. Grosan, A. Abraham, M. Chis // Studies in Computational Intelligence. – 2006. – №34. – P. 1–20.
8. Chu S.-C. Cat swarm optimization / S.-C. Chu, P.-W. Tsai, J. S. Pan // Lecture Notes in Artificial Intelligence. – 4099. – Berlin Heidelberg : Springer-Verlag, 2006. – P. 854–858.
9. Chu S.-C. Computational Intelligence based on the behavior of cats / S.-C. Chu, P.-W. Tsai // International Journal of Innovative Computing, Information, and Control. – 2007. – Vol. 3, № 1. – P. 163–173.
10. Online neuro fuzzy clustering of data with omissions and outliers based on completion strategy [Electronic resource] / [A. Shafronenko, Ye. Bodyanskiy, D. Rudenko] // Proceedings of The Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019), 2019. – Zaporizhzhia, 2019. – P. 18–27.
11. Shafronenko A. Yu. The Fast Modification of Evolutionary Bioinspired Cat Swarm Optimization Method [Electronic resource] / A. Yu. Shafronenko, Ye. V. Bodyanskiy, I. P. Pliss // 2019 IEEE 8th International Conference on Advanced Optoelectronics and Lasers (CAOL), 2019. – Sozopol, Bulgaria, 2019. – P. 548–552. DOI: 10.1109 /CAOL46282.2019.9019583

Received 16.02.2021.

Accepted 23.04.2021.

УДК: 004.8:004.032.26

ОНЛАЙН МЕТОД ВОЗМОЖНОСТНОЙ КЛАСТЕРИЗАЦИИ ДАННЫХ НА ОСНОВЕ ЭВОЛЮЦИОННОЙ ОПТИМИЗАЦИИ КОШАЧЬИХ СТАЙ

Бодянский Е. В. – д-р техн. наук, профессор, профессор кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники, Харьков, Украина.

Шафроненко А. Ю. – канд. техн. наук, доцент, доцент кафедры информатики Харьковского национального университета радиоэлектроники, Харьков, Украина.

Климова И. Н. – ассистент кафедры системной инженерии Харьковского национального университета радиоэлектроники, Харьков, Украина.

АННОТАЦИЯ

Актуальность. Проблемы кластеризации больших данных на сегодняшний день являются актуальной сферой искусственного интеллекта. Данная задача часто встречается во многих приложениях, связанных с интеллектуальным анализом данных, глубоким обучением и тому подобное. Для решения данных проблем традиционные подходы и методы требуют, чтобы вся выборка данных подавалась в пакетной форме.

Цель. Целью работы было предложить метод нечеткой вероятностной кластеризации данных с использованием эволюционной оптимизации кошачьих стай, который был бы лишен недостатков традиционных подходов кластеризации данных.

Метод. Процедура нечеткой вероятностной кластеризации данных с использованием эволюционных алгоритмов для быстрого определения экстремумов целевой функции, центроидов кластеров и адаптивных функций принадлежности, позволяющих не тратить ресурсы машины для хранения промежуточных расчетов и не требующих дополнительного времени для решения поставленной задачи кластеризации данных, независимо от размерности и способа подачи на обработку.

Результаты. Предложенный алгоритм кластеризации данных на основе эволюционной оптимизации достаточно прост в численной реализации, лишен недостатков, присущих традиционным методам нечеткой кластеризации, и может работать с большим объемом входной информации, обрабатываемой в режиме онлайн в реальном времени.

Выводы. Результаты эксперимента позволяют рекомендовать разработанный метод для решения проблем автоматической кластеризации и классификации больших данных, максимально быстро находить экстремумы выборки, независимо от способа подачи данных на обработку. Предложенный метод онлайн возможностной нечеткой кластеризации данных на основе эволюционной оптимизации кошачьих стай предназначен для использования в гибридных системах вычислительно-интеллекта, нейро-фаззи системах, в обучении искусственных нейронных сетей, в задачах кластеризации и классификации.

КЛЮЧЕВЫЕ СЛОВА: нечеткая кластеризация, правило обучения, оптимизация кошачьей стаи, режим трассировки, режим поиска.

UDC 004.8:004.032.26

ONLINE PROBABILISTIC FUZZY CLUSTERING METHOD BASED ON EVOLUTIONARY OPTIMIZATION OF CAT SWARM

Bodyanskiy Ye. V. – Dr. Sc., Professor at the Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Shafronenko A. Yu. – PhD, Associated Professor at the Department of Informatics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

Klymova I. N. – Assistant at the Department of System Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

ABSTRACT

Context. The problems of big data clustering today is a very relevant area of artificial intelligence. This task is often found in many applications related to data mining, deep learning, etc. To solve these problems, traditional approaches and methods require that the entire data sample be submitted in batch form.

Objective. The aim of the work is to propose a method of fuzzy probabilistic data clustering using evolutionary optimization of cat swarm, that would be devoid of the drawbacks of traditional data clustering approaches.

Method. The procedure of fuzzy probabilistic data clustering using evolutionary algorithms, for faster determination of sample extrema, cluster centroids and adaptive functions, allowing not to spend machine resources for storing intermediate calculations and do not require additional time to solve the problem of data clustering, regardless of the dimension and the method of presentation for processing.

Results. The proposed data clustering algorithm based on evolutionary optimization is simple in numerical implementation, is devoid of the drawbacks inherent in traditional fuzzy clustering methods and can work with a large size of input information processed online in real time.

Conclusions. The results of the experiment allow to recommend the developed method for solving the problems of automatic clustering and classification of big data, as quickly as possible to find the extrema of the sample, regardless of the method of submitting the data for processing. The proposed method of online probabilistic fuzzy data clustering based on evolutionary optimization of cat swarm is intended for use in hybrid computational intelligence systems, neuro-fuzzy systems, in training artificial neural networks, in clustering and classification problems.

KEYWORDS: fuzzy clustering, learning rule, cat swarm optimization, tracing mode, seeking mode.

REFERENCES

1. Bezdek J. C. Pattern recognition with fuzzy objective function algorithms. New York, Springer, 1981, 253 p. DOI <https://doi.org/10.1007/978-1-4757-0450-1>.
2. Höppner F., Klawonn F., Kruse R., Runkler T. Fuzzy Clustering Analysis: Methods for Classification, Data Analysis and Image Recognition. Chichester, John Wiley & Sons, 1999, 300 p.
3. Xu R., Wunsch D. C. Clustering. Hoboken N. J., John Wiley & Sons, Inc., 2009, 398 p.
4. Kohonen T. Self-Organizing Maps. Berlin, Springer-Verlag, 1995.
5. Krishnapuram R., Keller J. M. A Possibilistic Approach to Clustering, *IEEE Transactions on Fuzzy Systems*, May 1993: *Proceedings*, IEEE, 1993, Vol. 1, pp. 98–110. DOI: 10.1109/91.227387.
6. Bodyanskiy, Ye. Computational intelligence techniques for data analysis, *Lecture Notes in Informatics*. Bonn, Gesellschaft für Informatik, 2005, pp. 15–36.
7. Grosan C., Abraham A., Chis M. Swarm intelligence in Data Mining, *Studies in Computational Intelligence*, 2006, No. 34, pp. 1–20.
8. Chu S.-C., Tsai P.-W., Pan J. S. Cat swarm optimization, *Lecture Notes in Artificial Intelligence*, 4099. Berlin Heidelberg, Springer-Verlag, 2006, pp. 854–858.
9. Chu S.-C., Tsai P.-W. Computational Intelligence based on the behavior of cats, *International Journal of Innovative Computing, Information, and Control*, 2007, Vol. 3, No. 1, pp. 163–173.
10. Shafronenko A., Bodyanskiy Ye., Rudenko D. Online neuro fuzzy clustering of data with omissions and outliers based on completion strategy [Electronic resource], *Proceedings of The Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019)*, 2019. Zaporizhzhia, 2019, pp. 18–27.
11. Shafronenko A. Yu., Bodyanskiy Ye. V., Pliss I. P. The Fast Modification of Evolutionary Bioinspired Cat Swarm Optimization Method [Electronic resource], *2019 IEEE 8th International Conference on Advanced Optoelectronics and Lasers (CAOL)*, 2019. Sozopol, Bulgaria, 2019, pp. 548–552. DOI: 10.1109/CAOL46282.2019.9019583

EXPERIMENTAL ANALYSIS OF MULTINATIONAL GENETIC ALGORITHM AND ITS MODIFICATIONS

Gulayeva N. M. –PhD, Associate Professor at the Department of Informatics, National University of “Kyiv-Mohyla Academy”, Kyiv, Ukraine.

Yaremko S. A. – Ms. Sc., Assistant Lecturer at the Department of Informatics, National University of “Kyiv-Mohyla Academy”, Kyiv, Ukraine.

ABSTRACT

Context. Niching genetic algorithms are one of the most popular approaches to solve multimodal optimization problems. When classifying niching genetic algorithms it is possible to select algorithms explicitly analyzing topography of fitness function landscape; multinational genetic algorithm is one of the earliest examples of these algorithms.

Objective. Development and analysis of the multinational genetic algorithm and its modifications to find all maxima of a multimodal function.

Method. Experimental analysis of algorithms is carried out. Numerous runs of algorithms on well-known test problems are conducted and performance criteria are computed, namely, the percentage of convergence, real (global, local) and fake peak ratios; note that peak ratios are computed only in case of algorithm convergence.

Results. Software implementation of a multinational genetic algorithm has been developed and experimental tuning of its parameters has been carried out. Two modifications of hill-valley function used for determining the relative position of individuals have been proposed. Experimental analysis of the multinational genetic algorithm with classic hill-valley function and with its modifications has been carried out.

Conclusions. The scientific novelty of the study is that hill-valley function modifications producing less number of wrong identifications of basins of attraction in comparison with classic hill-valley function are proposed. Using these modifications yields to performance improvements of the multinational genetic algorithm for a number of test functions; for other test functions improvement of the quality criteria is accompanied by the decrease of the convergence percentage. In general, the convergence percentage and the quality criterion values demonstrated by the algorithm studied are insufficient for practical use in comparison with other known algorithms. At the same time using modified hill-valley functions as a post-processing step for other niching algorithms seems to be a promising improvement of performance of these algorithms.

KEYWORDS: multimodal optimization problem, niching genetic algorithms, multinational genetic algorithm, hill-valley function, genetic algorithm convergence, real peak ratio, fake peak ratio.

ABBREVIATIONS

ASD is an adaptive species discovery;
GA is a genetic algorithm;
HillValIEA is a hill-valley evolutionary algorithm;
HTS is a history-based topological speciation;
HVCMO is a hill-valley-clustering-based variable mesh optimization;
HVF is a hill-valley function;
MMOP is a multimodal optimization problem;
MNGA is a multinational genetic algorithm;
MNGA_{c_h_v} is a MNGA with $F_{c_h_v}$;
MNGA_{m_h_v} is a MNGA with $F_{m_h_v}$;
MNGA_{m_h_v_rand} is a MNGA with $F_{m_h_v_rand}$;
MNGAⁿ is a MNGA with national selection to the mating pool;
MNGA^w is a MNGA with weighted selection to the mating pool;
MNGA^m is a MNGA with mixed selection to the mating pool;
NBC is a nearest-better clustering;
RS-CMSA is a covariance matrix self-adaptation evolution strategy with repelling subpopulations;
SCGA is a species conserving GA;
TCGM_S2 is a GA of tournament crowding based on Gaussian mutation, with S2 parameter set;
TS is a topographical selection;
TSC is a topological species conservation.

NOMENCLATURE

$d(x,y)$ is an Euclidean distance between individuals x and y ;
 f is a fitness function;
 $f(s)$ is a fitness value of an individual s ;
 $F_{c_h_v}$ is a classic HVF;
 $F_{m_h_v}$ is a modified HVF;
 $F_{m_h_v_rand}$ is a modified HVF with randomized intermediate points;
 FPR is a ratio of the number of fake peaks found by the algorithm to the total number of species formed by the final population;
 $g(k, t)$ is a number of politicians of the k -th nation at the t -th algorithm step;
 GP is a number of real global peaks found in one run of the algorithm;
 GPR is a ratio of the number of real global peaks found by the algorithm to the total number of global peaks to be localized;
 $gr=[gr_1, gr_2, \dots, gr_k]$ is a sample array, $0 < gr_1 < gr_2 < \dots < gr_k < 1$;
 i_j is a j -th politician;
 K is a size of gr and a number of intermediate points used by HVF;
 LP is a number of real local peaks found in one run of the algorithm;

LPR is a ratio of the number of real local peaks found by the algorithm to the total number of local peaks to be localized;

n is a search space dimension;

NP is a number of real peaks found in one run of the algorithm;

$NSeeds$ is a power of $Seeds$, i.e. a number of species or, equivalently, a number of peaks (real and fake) found in one run of the algorithm;

$O_\varepsilon(x^*)$ is an ε -neighborhood of the point x^* ;

$\langle p_0, p_1, p_2, \dots, p_K, p_{K+1} \rangle$ is an ordered list of points used by HVF, $p_i \in X \subseteq R^n$;

p_0 is a first test point;

P_c is a probability of crossover;

p_i is an i -th intermediate point, $0 < i < K+1$;

p_i^j is a j -th coordinate of the i -th point;

p_{K+1} is a second test point;

$pl_k(t)$ is a policy of the k -th nation at the t -th algorithm step;

P_m is a probability of mutation;

PR is a ratio of the number of real peaks found by the algorithm to the total number of peaks to be localized;

S is a set of individuals encoding $X \subseteq R^n$;

s_i is an individual encoding point p_i , $0 \leq i \leq K+1$;

$Seeds$ is a set of the best individuals of each species;

$SucRuns$ is a percentage of successful runs of the algorithm;

δ is a function accuracy parameter;

ξ is a random value from $(0, 1)$;

σ is an argument accuracy parameter.

INTRODUCTION

Niching GAs, based on the phenomenon of speciation and specialization in natural ecosystems, are one of the most popular approaches to solve MMOPs. These problems aim to find multiple extremums (global, local) of a given function and arise in many areas of science and technology, see [1–5] for examples.

It can be shown that solving an arbitrary, including multimodal, optimization problem in the form

$$F(x) \rightarrow \max, x \in X \subseteq R^n \quad (1)$$

reduces to solving by a GA an optimization problem in the form

$$f(s) \rightarrow \max, s \in S. \quad (2)$$

Here with, an arbitrary feasible solution of problem (2) is called an individual of the population, and it is said about genotype and phenotype of the individual. The objective function of (2) is used to estimate the quality of solution of (1); crossover (the exchange of segments between different solutions) and mutation (disturbance) operators ensure transition from one solution to another.

A niche in GA is a subdomain of the search space (a region around a certain optimum of the fitness function), and species is a subpopulation of individuals located in a given niche. Niching GAs tend to structure population into stable subpopulations (species) in the search space in a way that each subpopulation is formed around one of the sought optimums. To date, several dozen niching GAs of different performance reflecting various approaches to solve MMOPs have been developed. Surveys of niching GAs and their classifications are available at [1, 6, 7].

In recent years topological niching GAs are of growing interest. To capture the landscape topography and identify basins of attraction for given individuals the topological niching GAs analyze relationships between locations and fitness values of individuals. Based on the basin identification method these algorithms can be roughly classified at the algorithms based on TS [8, 9], algorithms based on NBC [10–13], and algorithms using HVF. Last class includes but not limited to MNGA [14, 15], TSC and TSC2 [16, 17], ASD [18], HTS [19], Hill-Vallea [20, 21], HVcMO [22, 23], RS-CMSA [24, 25].

The object of study is a niching GA as a method to solve MMOPs.

The subjects of study are the HVF, MNGA and how MNGA parameters affect algorithm performance.

The purpose of the work is the development and performance analysis of MNGA and its modifications. Recall that the experimental analysis of heuristic algorithm performance is to evaluate statistical data collected as a result of conducting a series of independent runs of the algorithm for each problem from the benchmark suite.

1 PROBLEM STATEMENT

In this paper the MMOP (1) is considered in the following formulation. Let $F: X \rightarrow R$ be a function defined on some set $X \subseteq R^n$. A point $x^* \in X$ is called a point of local maximum of F over X if there exists a number $\varepsilon > 0$ such that $\forall x \in X \cap O_\varepsilon(x^*): F(x^*) \geq F(x)$. A point $x^* \in X$ is called a point of global maximum of F over X if $\forall x \in X: F(x^*) \geq F(x)$. The problem is to find all points of local and global maxima and the values of function F at these points.

The MNGA analysis is to compute well-known criteria characterizing the number and accuracy of problem solutions found by the MNGA and to compare the criterion values obtained for various values of the MNGA parameters.

2 REVIEW OF THE LITERATURE

A significant drawback of many niching algorithms is so-called niche radius problem. The performance of radii-based niching algorithms heavily depends on the niche radius value while estimation of this value is a complex task requiring prior knowledge of the search space landscape [1, 6]. As opposed to radii-based niching algorithms, the MNGA presented in [14, 15] divides popula-

tion into subpopulations without using the niche radius concept. To determine whether two individuals occupy the same niche (whether the points encoded by these individuals are in the neighborhood of the same extremum) the fitness-topology function HVF is used. This function analyzes the search space landscape between two points based on the fitness values of individuals encoding points located on the straight line that connects two test points.

In MNGA the following metaphor is used. The population of individuals represents the world (the entire population of the planet), each subpopulation represents the nation, and the fittest individuals in the subpopulation represent the government of the nation; these individuals are referred to as the politicians. The government determines the policy of the nation, which is a single point representing the peak the nation is formed around; this point is the centroid of the subpopulation fittest individuals. Policy calculation is needed to determine possible migrations of individuals between nations as well as to distinguish nations from each other. The evolution of the world obeys the following rules.

1. Migration. This rule regulates movement of individuals between nations and creation of new nations in "uninhibited" areas of the search space.

2. Merge. Nations are merged when they are formed around the same extremum.

3. Selection to the mating pool. The following modifications of binary tournament selection are considered.

1) Weighted selection. In this case fitness value of an individual is divided by the total number of individuals in its nation. This approach reduces the probability for a nation to die out because of small subpopulation size.

2) National selection is conducted within each nation, therefore, the number of individuals in a nation after selection remains unchanged. This implies that migration is the only way to change the nation size.

3) Mixed selection is a combination of weighted and national selections.

4. Election. This rule describes how the government of a nation is elected and how its policy is calculated. Note

that formula $pl_k(t) = \frac{1}{g(k,t)} \sum_{j=1}^{g(k,t)} i_j$ is computed coordi-

natewise. If the number of individuals in a nation is less than predefined value g , then the number of politicians is equal to the number of individuals in the subpopulation.

5. Mating. Only individuals belonging to the same nation may produce offspring. This limitation is because the crossover of two individuals located in the neighborhoods of two different optima may produce an offspring much worse than each of the parents. The mutation operator that adds noise generated by normal distribution to an individual is called in [15] the mutation based on the distance policy; this name looks reasonable as the operator provides minor changes to individuals that are close to the nation policy and significant changes to individuals located far from it. In literature, this operator is also known as Gaussian mutation. In [14] it is proposed to perform mutations within nations, i.e. an offspring is accepted

only if it occupies the same niche as its parent. To reduce the algorithm execution time, this restriction was not applied in our research.

6. Initialization of the start nation. At the start of the algorithm, all individuals belong to the same initial nation.

MNGA works as follows. In every generation, each individual is compared to the policy of its nation. If the individual and the policy are located around different optima, the individual migrates to the nation with policy in the individual's peak neighborhood. If such a nation does not exist, the individual founds a new one; this nation corresponds to a potentially new peak the individual is approaching. If a nation with a very small number of individuals is formed at the end of the migration process, this nation is strengthened by new individuals generated from the nation policy with the use of Gaussian mutation. Instead, the worst fitness individuals of other nations are removed from the population. After completion of all necessary migrations, the pairwise comparisons of policies of all nations are carried out to ensure there are no nations following the same peak; if two nations around the same peak are found, the corresponding subpopulations are merged. Thereafter, standard actions are performed to move to the next generation, i.e. selection to the mating pool and applying genetic operators. The MNGA scheme is provided below.

0. Encoding of solutions.

1. Initialization. The initial population is generated randomly.

2. Fitness calculation for all individuals in the population.

3. Initialization of the start nation. Determining politicians and calculation of the policy.

4. Migrations of individuals.

5. Recalculation of all nation policies.

6. Strengthening small nations.

7. Merging nations.

8. Selection to the mating pool.

9. Applying genetic operators (crossover, mutation).

10. Fitness calculation for all obtained offspring.

11. If the stop condition is fulfilled, then goto step 12, otherwise goto step 4.

12. Stop the algorithm.

To implement processes of migration of individuals and merging of nations the HVF is used. This function determines the positional relationship of two selected individuals, in fact of points encoded by these individuals. HVF returns TRUE if these points are located in the neighborhoods of different maxima (there is a "valley" in the fitness function landscape between the points). HVF returns FALSE if these points are located in the neighborhood of the same maximum (there is a "hill" in the fitness function landscape between the points).

The idea of the HVF is as follows. A set of points disposed on the line connecting two test points is generated; to calculate these intermediate points a predefined array gr is used. If fitness of all intermediate individuals is not less than fitness of the worst of the test individuals, these

test individuals are considered to be located around the same maximum (the HVF returns FALSE); otherwise they are considered to be located around different maxima (the HVF returns TRUE).

The $F_{c_{h_v}}$ computation scheme proposed in [14, 15] is given below.

1. Put $i = 1$; $found = FALSE$.
2. If $i \leq K$ and $found == FALSE$, then goto step 3 (intermediate point calculation); otherwise goto step 6.
3. For each j from 1 to n do: $p_i^j = p_0^j + (p_{K+1}^j - p_0^j) * gr_i$.
4. If $f(s_i) < \min(f(s_0), f(s_{K+1}))$, then put $found = TRUE$.
5. Put $i = i + 1$.
6. Return $found$.

For example, let $n=2$ and $gr=[0.25, 0.5, 0.75]$. Then, to determine the relative position of points (0, 1) and (1, 2), the intermediate points (0.25, 1.25), (0.5, 1.5) and (0.75, 1.75) will be analyzed by the algorithm above.

In [15] the array [0.25, 0.5, 0.75] is used to ensure the migration process, and the array [0.02, 0.25, 0.5, 0.75, 0.98] is used to ensure the merging process. The extension of the sample array by two extra points is explained by the fact that merging subpopulations is more drastic operation than migration of an individual from one nation to another.

3 MATERIALS AND METHODS

In some cases the $F_{c_{h_v}}$ return values are wrong. Let us consider functions F_1 and F_2 of dimension 1, see Section 4, and determine the positional relationship of the points 0.13 and 0.97 by the $F_{c_{h_v}}$, $gr=[0.02, 0.25, 0.5, 0.75, 0.98]$. Notice that each of the functions has five peaks and test points are located on the first and the fifth peaks with respect to the left-to-right peak numeration alongside the X -axis. But $F_{c_{h_v}}(0.13, 0.97)=FALSE$. This means that points 0.13 and 0.97 are in the neighborhood of the same peak according to $F_{c_{h_v}}$. Table 1 provides F_1 and F_2 values at the test and intermediate points.

Table 1 – Values of F_1 and F_2 at the points analyzed by HVF

i	p_i	F_1	F_2
0	0.13	0.50036	0.49939
1	0.1468	0.16654	0.16575
2	0.34	0.28038	0.24749
3	0.55	0.125	0.08061
4	0.76	0.04124	0.01605
5	0.9532	0.09102	0.01881
6	0.97	0.00876	0.00170

The first proposed HVF modification $F_{m_{h_v}}$ is the following. Test points are considered to be located in the neighborhood of the same maximum if, moving along the list $\langle p_0, p_1, p_2, \dots, p_K, p_{K+1} \rangle$ from p_i to p_{i+1} , $0 \leq i \leq K+1$, fitness changes of individuals encoding these points conform to one of the rules below:

- from a smaller value to a bigger one;
- from a bigger value to a smaller one;
- from a smaller value to a bigger one and again to a smaller value.

Test points are considered to be located in the neighborhoods of different maxima in all other cases.

As an example, let test points p_0, p_{K+1} such that $f(s_0)=2$ and $f(s_{K+1})=1$ be given. The above modification considers these points as points located around the same maximum if the list of fitness values $\langle 2, 3, 4, 5, 7, 3, 1 \rangle$ is analyzed, and as points located around different maxima if the list of fitness values $\langle 2, 3, 5, 3, 7, 4, 1 \rangle$ is analyzed.

It is readily seen that $F_{m_{h_v}}$ returns correct value for the above example defined for F_1 and F_2 test functions: $F_{m_{h_v}}(0.13, 0.97)=TRUE$.

Assume that all the points (test as well as intermediate) are related to individuals of equal fitness. Notice that in this case the $F_{c_{h_v}}$ treats test points as being located around the same maximum (assuming a plateau), and the $F_{m_{h_v}}$ treats them as being located around different maxima (assuming existence of neighbor peaks). In general, the presence of plateaus in the fitness function landscape should be avoided since in this case GA can not compare individuals properly.

There exist cases when both $F_{c_{h_v}}$ and $F_{m_{h_v}}$ return wrong values. For instance, points 0.09 and 0.91 lie on different peaks of functions F_1 and F_2 of dimension 1, but they are in the neighborhood of the same peak according to both $F_{c_{h_v}}$ and $F_{m_{h_v}}$ when $gr=[0.02, 0.25, 0.5, 0.75, 0.98]$ is used: $F_{c_{h_v}}(0.09, 0.91)=FALSE$ and $F_{m_{h_v}}(0.09, 0.91)=FALSE$. Table 2 provides F_1 and F_2 values at the test and intermediate points. The F_1 graph with circles designating all target points is given in Fig. 1.

Table 2 – Values of F_1 and F_2 at the points analyzed by HVF

i	p_i	F_1	F_2
0	0.09	0.92837	0.92817
1	0.1064	0.97009	0.97000
2	0.2950	0.98165	0.90403
3	0.5	1.0	0.70711
4	0.705	0.98165	0.44425
5	0.8936	0.97009	0.24794
6	0.91	0.92837	0.22414

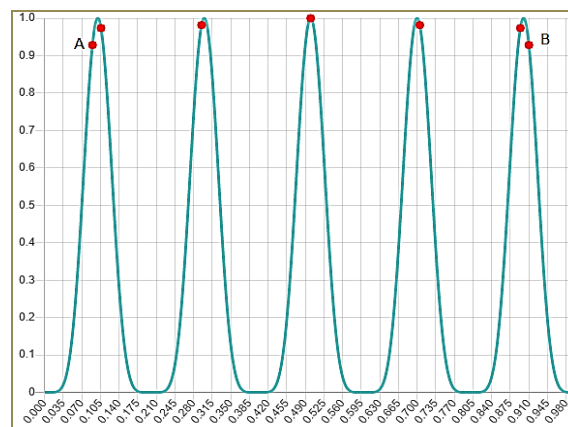


Figure 1 – Graph of F_1 with points from Table 2

The second proposed HVF modification $F_{m_{h_v_rand}}$ is to add random values ξ_1, ξ_2 to the sample array $gr=[0.25, 0.5, 0.75]$; these random values should be generated at each iteration of the algorithm. Such a dynamic array re-

duces the probability of wrong values returned by the $F_{m_h_v}$ as far as different sets of intermediate points may be used for two test points. In particular, in the example above there could be generated the sample array $gr=[0.25, 0.35, 0.5, 0.6, 0.75]$ ensuring the $F_{m_h_v}$ correct return value. Table 3 provides F_1 and F_2 values at the test and intermediate points. The F_1 graph with circles designating all target points is given in Fig. 2.

Table 3 – Values of F_1 and F_2 at the points analyzed by HVF

i	p_i	F_1	F_2
0	0.09	0.92837	0.92817
1	0.295	0.98165	0.90403
2	0.377	0.00195	0.00165
3	0.5	1	0.70711
4	0.582	0.00047	0.00029
5	0.705	0.98165	0.44425
6	0.91	0.92837	0.22414

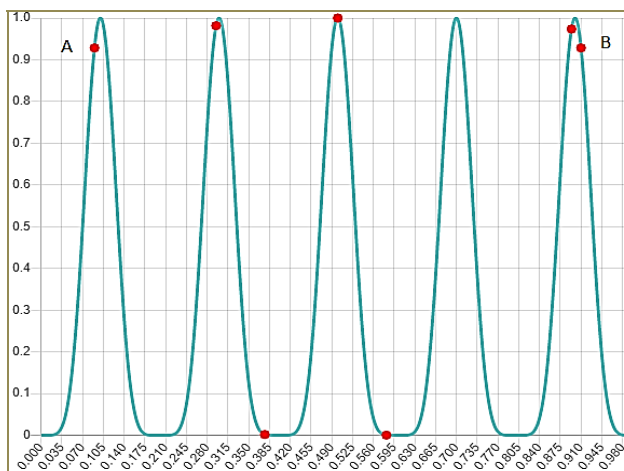


Figure 2 – Graph of F_1 with points from Table 3

Preliminary tests revealed significant differences in the $F_{c_h_v}$, $F_{m_h_v}$ and $F_{m_h_v_rand}$ responses. The statistics of these function calls for one MNGA run on the F_1 test function is shown in Table 4. Following sections provide the benchmarking of the MNGA with $F_{c_h_v}$, $F_{m_h_v}$ and $F_{m_h_v_rand}$.

Table 4 – Differences in responses of $F_{c_h_v}$, $F_{m_h_v}$, $F_{m_h_v_rand}$ for one MNGA run on F_1

Number of different responses for:	$n=1$	$n=3$
$F_{c_h_v}$ and $F_{m_h_v}$	2 616	1 020 942
$F_{c_h_v}$ and $F_{m_h_v_rand}$	1 900	1 114 164
$F_{m_h_v}$ and $F_{m_h_v_rand}$	872	253 978
Total number of HVF calls:	23 265	5 209 793

4 EXPERIMENTS

Recall that in this paper the experimental analysis of performance of the MNGA and its modifications is used. Thus, test problems and performance criteria should be defined.

Test suite of benchmark functions used in this research is given below.

1) Equal Maxima function (Deb's 1):

$$y = F_1(x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n \sin^6(5\pi x_i),$$

$$0 \leq x_i \leq 1, i = \overline{1, n}.$$

The function has 5^n equally spaced global maxima.

2) Decreasing Maxima function (Deb's 2):

$$y = F_2(x_1, x_2, \dots, x_n) = \sum_{i=1}^n e^{-2(\ln 2) \left(\frac{x_i - 0.1}{0.8} \right)^2} \sin^6(5\pi x_i),$$

$$0 \leq x_i \leq 1, i = \overline{1, n}.$$

The function has 5^n equally spaced maxima of different heights.

3) Uneven Maxima Function (Deb's 3):

$$y = F_3(x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n \sin^6(5\pi(x_i^{0.75} - 0.05)),$$

$$0 \leq x_i \leq 1, i = \overline{1, n}.$$

The function has 5^n global maxima that are at different distances from each other, and the distance between the points of maximum increases with increasing value of the argument.

4) Uneven Decreasing Maxima function (Deb's 4):

$$y = F_4(x_1, x_2, \dots, x_n) = \sum_{i=1}^n e^{-2(\ln 2) \left(\frac{x_i - 0.08}{0.854} \right)^2} \sin^6(5\pi(x_i^{0.75} - 0.05)),$$

$$0 \leq x_i \leq 1, i = \overline{1, n}.$$

The function has 5^n maxima of different heights that are at different distances from each other, and the distance between the points of maximum increases with increasing value of the argument.

5) Six-Hump Camel Back function:

$$y = F_5(x_1, x_2) = - \left(\left(4 - 2.1x_1^2 + \frac{x_1^4}{3} \right) x_1^2 + x_1 x_2 + 4(x_2^2 - 1)x_2^2 \right),$$

$$-3 \leq x_1 \leq 3, -2 \leq x_2 \leq 2.$$

The function has 2 global and 4 local maxima.

6) Griewangk's function is considered in the following form:

$$y = F_6(x_1, x_2, \dots, x_n) = n - \left(\sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \right),$$

$$-600 \leq x_i \leq 600, i = \overline{1, n}.$$

The function has 1 global and many local maxima, and the height of the local maxima decreases with increasing distance from the global maximum.

7) Rastrigin's function is considered in the following form:

$$y = F_7(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \left(10 \cos(2\pi x_i) - x_i^2 \right) - 10n,$$

$$-5.12 \leq x_i \leq 5.12, \quad i = \overline{1, n}.$$

The function has 1 global and 10^n local maxima, and the height of the local maxima decreases with increasing distance from the global maximum.

8) Modified Rastrigin's function:

$$y = F_8(x_1, x_2) = -10 - 9 \cos(6\pi x_1) - 10 - 9 \cos(8\pi x_2),$$

$$0 \leq x_1, x_2 \leq 1.$$

The function has 12 global maxima.

9) Xin-She Yang's function:

$$y = F_9(x_1, x_2, \dots, x_n) = \left(\sum_{i=1}^n |x_i| \right) \exp \left(- \sum_{i=1}^n x_i^2 \right),$$

$$-10 \leq x_i \leq 10, \quad i = \overline{1, n}.$$

The function has several global maxima located at a small distance from each other. In particular, for $n=1$ the function has 2 global maxima, and for $n=2$ the function has 4 global maxima.

Detailed description and plots of the functions above can be found in [26, 27].

To specify the performance criteria we begin with two definitions: the algorithm stop conditions and whether the extremum is found when the algorithm stops.

GA stops if either population convergence is detected, i.e. the changes in the average population fitness value do not exceed 0.0001 over the last 10 generations, or 40000 algorithm iterations are carried out. This means that if convergence was not detected the algorithm has been terminated after the specified number of iterations. The algorithm run stopped under the convergence condition is called *successful*.

The *maximum is found* if the algorithm convergence is detected and there is at least one individual of the final population such that the individual fitness value differs from the sought maximum value at most δ and the point encoded by this individual is located within the maximum's neighborhood of radius σ . Let us set $\delta=0.01$ and $\sigma=0.01$ for all test functions.

As stated in [28], when algorithm stops under any condition including convergence, individuals of the final population may be located not only in the small radius neighborhoods of function peaks but individuals may also form clusters on the peak slopes and even in the valleys. Thus, to analyze the algorithms, standard performance criteria are used as well as proposed in [28] *FPR* criterion that in fact estimates number of clusters formed by individuals located far from the sought maxima. Note that big value of *FPR* criterion is a significant drawback of a niching algorithm as far as it makes use of such algorithms impractical for problems of finding all (local, global) maxima.

To determine species (and, consequently, niches) formed by the final population the algorithm of [28] presented below is used; $\varepsilon=0.03$.

1. Put *Seeds* = \emptyset .
2. Choose the best unprocessed individual s from the population; mark it as the processed one: s^* .
3. Put *found* = FALSE.
4. If there exists an individual $s \in \text{Seeds}$ such that $d(s^*, s) \leq \varepsilon$, then put *found* = TRUE; otherwise, create a new species with the best individual s^* : put *Seeds* = *Seeds* $\cup \{s^*\}$.
5. If unprocessed individuals are remained in the population, then goto step 2.

To separate real species from the fake ones the elements from *Seeds* should be compared with real peaks of the fitness function using parameters δ and σ .

The criteria calculated for each run of the algorithm are the following ones: *NSeeds*, *NP*, *GP*, *LP*, *PR*, *GPR*, *LPR*, *FPR*. Note that $NP=GP+LP$, $0 \leq PR \leq 1$, $0 \leq GPR \leq 1$, $0 \leq LPR \leq 1$, $0 \leq FPR \leq 1$. The *FPR* criterion is calculated as

$$FPR = \frac{NSeeds - NP}{NSeeds}.$$

For algorithm analysis average values of the *PR*, *GPR*, *LPR*, *FPR* criteria are computed over all runs as well as the *SucRuns* criterion; the *SucRuns* is computed as the ratio of the number of successful runs to the total number of runs, in percent. Criterion values of unsuccessful runs are not used to calculate averages; recall that the unsuccessful runs are the runs stopped under the condition of reaching maximum iteration number.

The software implementation of the MNGA with the classic HVF and its both modifications was developed. The criteria above were computed with the following purposes. First, to tune the parameters of the MNGA with $F_{c_h_v}$. Secondly, to conduct comparative analysis of the MNGA with $F_{c_h_v}$, $F_{m_h_v}$ and $F_{m_h_v_rand}$.

To calculate averages of the criteria above there were conducted 10 runs of the algorithm for each set of parameter values and for each benchmark problem. Note that the initial population is the same for all sets of parameter values at the i -th run; this guarantees the same start point for corresponding runs of the algorithms using different sets of parameter values.

The authors of this research tried to follow the [14, 15] recommendations on setting the algorithm parameter values. If such recommendations were not provided in [14, 15], the algorithm parameter values have been set based on the authors' experience or by conducting experiments. The predefined MNGA parameter values used in this research are listed below.

- Number of individuals in the population: 500.
- Number of individuals in the government: 8.
- Array to implement migration of individuals and merging of nations:
 - [0.02, 0.25, 0.5, 0.75, 0.98] for functions $F_{c_h_v}$ and $F_{m_h_v}$;

– $[0.25, 0.5, 0.75] \cup \{\xi_1, \xi_2\}$ for function $F_{m_h_v_rand}$.

In [15], small nations are strengthened by the nation policy mutants, i.e. by addition of a small noise to the policy of the nation. To reduce the algorithm computation complexity the zero noise is used in our research, i.e. small nations are supported by copies of their policies. Since there is no information in [14, 15] about nations to be small or large, we define the small nation as a nation with population size less than 30% of the average number of individuals per nation at the current algorithm iteration, and the large nation as a nation with population size exceeding 50% of the average number of individuals per nation at the current algorithm iteration. This implies that the number of policy copies strengthening a small nation ensures population size of this nation to be not less than 30% of the average number of individuals per nation at the current algorithm iteration.

Let us remark that step 5 of the MNGA (the recalculation of all nation policies) is not specified in [14, 15]. We have added this step to ensure the use of actual nation policies while strengthening small nations and merging nations.

Further on, the following MNGA parameter values are analyzed experimentally.

– Encoding methods: real numbers and Gray codes encoding sample points to three decimal places. Notice that binary codes are used in [14] and real numbers are used in [15].

– Selection to the mating pool: the weighted selection, the national selection and the mixed selection. For mixed selection 50% of individuals are selected using weighted selection and 50% of individuals are selected using national selection [15].

– Genetic operators used for binary encoding are one-point crossover and density mutation. The analyzed values of P_c are 0.6, 0.8, 0.9 and 1. The analyzed values of P_m are 0.001, 0.01 and 0.025. In [14] values $P_c=0.9$ and $P_m=0.025$ are recommended. Gaussian density mutation is the only genetic operator used for encoding in real numbers, $P_m=1$ [15]. In this case noise determined by standard normal distribution is added to each gene encoded variable.

Generally, 39 parameter sets are analyzed in this research.

For MNGA parameter tuning the F_1-F_4 functions are used. These functions have different properties, e.g. existence of only global maxima (F_1, F_3) or global and local maxima (F_2, F_4), location of the maxima at equal (F_1, F_2) or at different (F_3, F_4) distances from each other. Therefore, Deb's benchmark functions are widely used for experimental analysis of algorithms solving MMOPs. Having regard to the function properties two test suites for parameter tuning are composed:

– T1 includes functions F_2 and F_4 having one global and many local maxima;

– T2 includes functions F_1 and F_3 having many global and no local maxima.

5 RESULTS

The convergence of the MNGA with $F_{c_h_v}$ was not observed for any parameter set using binary codes; experiments were conducted for both T1 and T2 test suites. Therefore, parameter sets with binary codes were excluded from further analysis.

Results obtained for real number encoding and Gaussian density mutation, $P_m=1$, are given in Table 5. It is easy to see from the table that the percentage of successful runs is poor in most cases. Moreover, it is worthwhile to consider only national selection method as far as $MNGA_{c_h_v}^w$ converges only when $n=1$ while for $MNGA_{c_h_v}^m$ the *SucRuns* criterion value is close to 0%.

We assume that proposed HVF modifications return wrong values more rarely in comparison with $F_{c_h_v}$; thus, using $F_{m_h_v}$ or $F_{m_h_v_rand}$ instead of $F_{c_h_v}$ should improve the overall performance of the MNGA. With regard to above, the MNGA runs for classic HVF and its modifications were conducted using national selection method; results obtained for test suites T1 and T2 are given in Table 6. It is easy to see from the table that using HVF modifications gives higher percentage of successful runs for test suite T1, but decreases this percentage for test suite T2. Anyway, the *SucRuns* criterion value is still poor in most cases.

Assuming that another selection method tuned for appropriate HVF modification could improve the MNGA performance, the MNGA runs for different HVFs and different selection methods were conducted; both T1 and T2 test suites were used. Experiments have shown the thorough improvements of algorithm convergence. In particular, best results of *SucRuns* criterion were achieved by the $MNGA_{m_h_v}^w$ (98.33%) and $MNGA_{m_h_v_rand}^w$ (91.67%) on T1 test suite (functions having one global and many local maxima), and by the $MNGA_{m_h_v}^m$ (100%) and $MNGA_{m_h_v_rand}^m$ (100%) on T2 test suite (functions having many global and no local maxima). On the other hand, the quality criteria become worse with increasing the convergence criterion value, in particular, the *PR* criterion value decreased 1.7–8.6 times. Supporting data (averages computed over all runs for functions F_1-F_4) are provided in Table 7. Note that computation of averages over all benchmark functions makes sense as far as the function surfaces are not known in advance for most practical problems.

We have also developed the software implementation of the SCGA [29] intending to analyze this algorithm and to develop its hybridization with the MNGA; this seems to be a promising area [16, 17]. Parameter tuning conducted for this algorithm led to the following parameter values: SUS selection, one-point mutation, $P_m=0.01$, no crossover, Euclidean distance, $\sigma_s=0.2$. Experimental analysis of the SCGA with mentioned parameter values was conducted for functions F_1-F_4 ; results are given in Table 7. It is obvious from the table that though the SCGA provides satisfactory percentage of the algorithm convergence the quality parameter values of this algorithm are poor.

Table 5 – Results of experiments: MNGA with classic HVF

Parameters	Test suite T1					Comment	Test suite T2			Comment
	Criteria: average values over 10 runs						Criteria: average values over 10 runs			
	<i>SucRuns</i>	<i>PR</i>	<i>GPR</i>	<i>LPR</i>	<i>FPR</i>		<i>SucRuns</i>	<i>PR=GPR</i>	<i>FPR</i>	
Weighted	31.67%	0.2	0	0.25	0	Convergence only if $n=1$	16.67%	1	0	Convergence only if $n=1$
National	45%	0.0898	0.4444	0.0749	0.7567		93.33%	0.5051	0.4287	
Mixed	0%						2%	0.9667	0.0040	Convergence only for F_1 , if $n=1$ or $n=2$

Table 6 – Results of experiments: MNGA with classic HVF and its modifications

Algorithm	Test suite T1					Test suite T2		
	Criteria: average values over 10 runs					Criteria: average values over 10 runs		
	<i>SucRuns</i>	<i>PR</i>	<i>GPR</i>	<i>LPR</i>	<i>FPR</i>	<i>SucRuns</i>	<i>PR=GPR</i>	<i>FPR</i>
MNGA ⁿ _{c_h_v}	45%	0.0898	0.4444	0.0749	0.7567	93.33%	0.5051	0.4287
MNGA ⁿ _{m_h_v}	63.33 %	0.6158	0.5790	0.6173	0.4173	68.33%	0.6092	0.4361
MNGA ⁿ _{m_h_v_rand}	63.33 %	0.5737	0.6053	0.5680	0.3542	53.33 %	0.79	0.1919

Table 7 – Results of experiments: MNGA with classic HVF and its modifications, SCGA, TCGM S2; functions F_1-F_4

Algorithm	Criteria: average values over all runs					Comment
	<i>SucRuns</i>	<i>PR</i>	<i>GPR</i>	<i>LPR</i>	<i>FPR</i>	
MNGA ⁿ _{c_h_v}	69.17%	0.3700	0.4854	0.0749	0.5354	
MNGA ⁿ _{m_h_v}	65.83%	0.6124	0.5946	0.6173	0.4270	
MNGA ⁿ _{m_h_v_rand}	58.33 %	0.6726	0.6897	0.5680	0.280	
MNGA ^m _{m_h_v}	99.17%	0.1904	0.1799	0.0863	0.4622	
MNGA ^w _{m_h_v_rand}	95.83 %	0.1908	0.2023	0.0735	0.2783	
SCGA	90.83%	0.2680	0.4481	0.025	0.6773	
TCGM S2	100%	0.9563	0.8908	0.4744	0.1538	see [28]

6 DISCUSSION

From the Table 6 it follows that using HVF modifications improves the performance of the MNGA solving optimization problems for functions having global and local maxima. Indeed, the *SucRuns* increased by nearly 41%, the *PR* increased nearly 6.86 times for $F_{m_h_v}$ and 6.39 times for $F_{m_h_v_rand}$ (mainly due to the greater number of local peaks found), the *FPR* decreased significantly, to be exact by 55% for $F_{m_h_v}$ and by 47% for $F_{m_h_v_rand}$. Let us remark that $F_{m_h_v_rand}$ gives better results in comparison with $F_{m_h_v}$ in terms of *GPR* and *FPR*, but worse results in terms of *LPR*. Fig. 3 illustrates the forgoing via a histogram.

Now let us analyze the influence of $F_{m_h_v}$ and $F_{m_h_v_rand}$ on performance of the MNGA solving optimization problems for functions having only global maxima. From the Table 6 we see that *SucRuns* decreased considerably, namely by 27% for $F_{m_h_v}$ and by 43% for $F_{m_h_v_rand}$. At the same time, the *PR* increased by 21% for $F_{m_h_v}$ and by 56% for $F_{m_h_v_rand}$. Note that $F_{m_h_v_rand}$ gives better results in comparison with $F_{m_h_v}$ in terms of *GPR* and *FPR*, but provides worse algorithm convergence. Fig. 4 illustrates the forgoing via a histogram.

Consequently, the performance improvement of the MNGA using HVF modifications can be observed when the objective function has global and local maxima. If the objective function has only global maxima, the quality criteria improvement is accompanied by the decrease of the convergence criterion value.

Let us consider other test functions. Functions F_8 and F_9 have only global maxima. Experiments conducted for these functions of dimensions 1 and 2 showed 100% of convergence and *GPR* value close to 1 for all three kinds

of the algorithm: MNGAⁿ_{c_h_v}, MNGAⁿ_{m_h_v}, MNGAⁿ_{m_h_v_rand}.

Functions F_5 , F_6 and F_7 have global and local maxima. Experiments conducted for functions F_5 of dimension 2 and F_7 of dimensions 1, 2 and 3 revealed the advantage of modified HVFs over classic HVF in terms of *PR* (mainly due to the greater number of local peaks found). However, modified HVFs provide worse algorithm convergence. Function F_6 is defined on the wide range of argument values; therefore, to find maxima of F_6 the population size parameter of the MNGA was increased to 1000 individuals. Experiments were conducted for function F_6 of dimensions 1, 2 and 3. The MNGAⁿ_{m_h_v} algorithm demonstrated significantly better results both in terms of convergence and quality of found solutions: *SucRuns*=90% for MNGAⁿ_{m_h_v} and *SucRuns*=50% for MNGAⁿ_{c_h_v}, *GPR*=0.4074 for MNGAⁿ_{m_h_v} and *GPR*=0.2667 for MNGAⁿ_{c_h_v}.

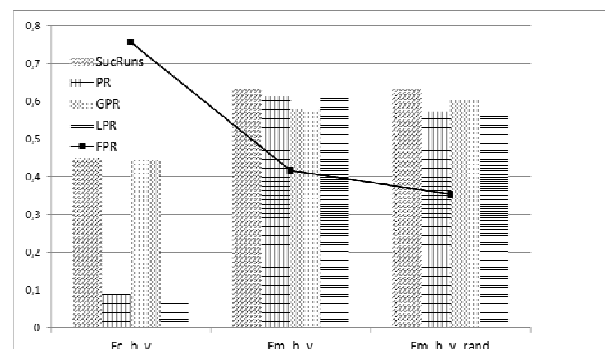


Figure 3 – Histogram with criteria values obtained for MNGA using different HVFs, test suite T1

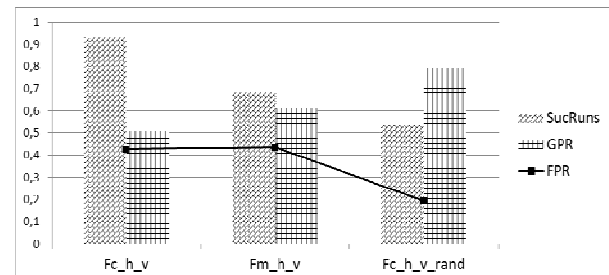


Figure 4 – Histogram with criteria values obtained for MNGA using different HVFs, test suite T2

From the stated above it can be concluded that $MNGA_{m,h,v}^n$ performance is either close to the performance of $MNGA_{c,h,v}^n$ or is better in terms of PR ; at the same time improvements in PR are accompanied by the decrease of $SucRuns$ for some functions. This is also the case for the $MNGA_{m,h,v,rand}^n$, though $MNGA_{m,h,v,rand}^n$ often gives worse results than the $MNGA_{m,h,v}^n$, especially in terms of convergence.

By analyzing data from Table 7 providing benchmarking results for functions F_1 – F_4 we can state that generally all kinds of MNGA studied in this research are unacceptable for practical use as far as they are characterized either by low value of convergence criterion (that is less than 90%) or by low ratio of real peaks and high ratio of fake peaks found by the algorithm. This is also true for other test functions. As for F_9 , there are observed $SucRuns=100\%$ and GPR value close to 1 for all kinds of MNGA. This can be explained by the specific surface of F_9 : the function curve is a plateau with four peaks close to each other. Such a surface is a challenge for most optimization algorithms but turned to be solvable for MNGA.

Note that all experiments in our study are conducted for test functions of small dimensions ($1 \leq n \leq 3$). The reason is the long algorithm execution time caused by a number of factors including the low convergence percentage. For example, the $MNGA_{m,h,v}^n$, $MNGA_{m,h,v}^w$, $MNGA_{m,h,v}^m$ running time to find optimums of F_5 – F_9 took about 48 hours. The computer configuration used to run these algorithms was the following: processor Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.71 GHz; 8.00 GB RAM (7.88 GB available); 64-bit Windows 10 operating system, x64 processor, 2 cores, 4 logic processors.

CONCLUSIONS

In this paper, the MNGA parameter tuning was carried out and experimental analysis of the MNGA performance was conducted. Two modifications of the HVF used to determine the relative position of individuals were proposed. Benchmarking of MNGA with classic HVF and with its modifications was carried out. Experiments showed that proposed modifications had increased the number of real extrema found by the algorithm for most test functions. Stated above determines the **scientific novelty** of the obtained results.

The **practical significance** of the obtained results lies in the following statements.

– First, as HVF modifications produce less number of wrong basin identifications in comparison with classic HVF, it is recommended to use $F_{m,h,v}$, $F_{m,h,v,rand}$ in other GAs using HVF [16–25].

– Secondly, as the overall performance of MNGA is poor in comparison with other known GAs, the MNGA is not recommended for practical use.

Indeed, GAs solving multimodal optimization problems, i.e. the problems of finding several optimums, contribute to grouping individuals of final population into species so that the best individual of each species represents one of the sought optimums. Obviously, it is worthwhile to analyze groups of individuals only in case of GA convergence. Generally, the number and location of function optimums are not known a priori for practical problems. It is known, that individuals of the GA final population may be located in the neighborhoods of function peaks as well as far from them, clustering on peak slopes or even in the valleys. Therefore, to choose an algorithm to solve practical MMOPs, in particular problems of finding global and local peaks, it is important to analyze criteria representing number of real peaks found by the algorithm as well as number of fake species formed by the algorithm. These are the PR , GPR , LPR and FPR criteria.

Experimental analysis of the MNGA with both classic HVF and its modifications has revealed the unacceptably low value of the convergence criterion ($SucRuns < 90\%$) and unacceptable for practical use values of the quality criteria, i.e. small values of PR , GPR , LPR accompanied by the high value of the FPR . Generally, depending on algorithm parameter values there is observed either improvement of quality criteria values accompanied by the worse algorithm convergence or higher convergence percentage accompanied by deterioration of quality criteria values. Note that quality parameter values of the SCGA algorithm developed for farther hybridization [16, 17] with MNGA are also poor. Let us remark that in [28] there was proposed TCGM_S2 algorithm; criterion values computed in [28] for this algorithm are given in Table 7. It is readily seen that this algorithm provides better criterion values in terms of convergence percentage as well as in terms of quality of the solutions found. Thus it is reasonable to say that it is impractical to use MNGA to find global and local optima of a multimodal function. On the other hand, hybridizations of MNGA with other algorithms could be an option.

Let us remark that there still exist situations when $F_{m,h,v}$, $F_{m,h,v,rand}$ return wrong values. Let us consider function F_1 and sample array $gr=[0.02, 0.25, 0.5, 0.75, 0.98]$ from [15]. Points 0.0 and 1.0 are wrongly attributed to the same basin of attraction by both $F_{c,h,v}$ and $F_{m,h,v}$; return value of $F_{m,h,v,rand}$ is wrong with a certain probability. To overcome this, modifications of $F_{m,h,v}$ can be proposed. Modifications regarding a method to construct sequences of sample points are an option. Several of these were proposed in [18–20, 24, 30], e.g. replacement of equidistant test points by a golden section search or changing quantity of test points dependent on the distance between individuals.

In [14] it is stated that MNGA forms only one sub-population in the neighborhood of every peak. Preliminary experiments have shown that using HVF to determine species (niches) formed by the final population instead of the algorithm used in this research (the algorithm from [28]) gives better results in terms of *FPR*. At the same time *SucRuns*, *PR*, *GPR* and *LPR* criteria values obviously stay the same. We believe it is worth to conduct benchmarking of different niching GAs using HVF or its modifications as the algorithm to determine species (or, equivalently, as a post-processing step for every algorithm after its convergence).

Thus, we see two main **directions for further research**: improvements of the HVF and using HVF as a post-processing step for other niching GAs. This will be the object of another paper.

ACKNOWLEDGEMENTS

Authors express great appreciations to their colleges from the National University of “Kyiv-Mohyla Academy”, Head of the Department of Informatics S. S. Gorokhovskiy and Professor of the Department of Informatics M. M. Glybovets, for their valuable remarks on the obtained results.

REFERENCES

1. Li X., Epitropakis M. G., Deb K. et al. Seeking Multiple Solutions: An Updated Survey on Niching Methods and Their Applications, *IEEE Transactions on Evolutionary Computation*, 2017, Vol. 21, No. 4, pp. 518–538. DOI: 10.1109/TEVC.2016.2638437.
2. Shen Z.-H., Zhao Y.-K., Wu W.-W. Niche Pseudo-Parallel Genetic Algorithms for Path Optimization of Autonomous Mobile Robot, *Journal of Shanghai University (English Edition)*, 2006, Vol. 10, No. 5, pp. 449–453. DOI: 10.1007/s11741-006-0089-3.
3. Preuss M., Liapis A., Togelius J. Searching for Good and Diverse Game Levels, *IEEE Conference on Computational Intelligence and Games, Dortmund, 26–29 August, 2014: proceedings*. Piscataway, IEEE, 2014, pp. 1–8. DOI: 10.1109/CIG.2014.6932908.
4. Chica M., Barranquero J., Kajdanowicz T. et al. Multimodal Optimization: an Effective Framework for Model Calibration, *Information Sciences*, 2017, Vol. 375, pp. 79–97. DOI: 10.2139/ssrn.2828069.
5. Leoshchenko S. D., Oliinyk A. O., Subbotin S. A. et al. Modification and Parallelization of Genetic Algorithm For Synthesis of Artificial Neural Networks, *Radio Electronics, Computer Science, Control*, 2019, No. 4, pp. 68–82. DOI: 10.15588/1607-3274-2019-4-7.
6. Glybovets M. M., Gulayeva N. M., Butenko S., Pardalos P. M., Shylo V. (eds.). Evolutionary Multimodal Optimization, *Optimization methods and applications. In honor of Ivan V. Sergienko's 80th birthday*. Cham, Springer, 2017, pp. 129–173. DOI: 10.1007/978-3-319-68640-0_8. (Series “Springer Optimization and Its Applications”)
7. Preuss M. Multimodal Optimization by Means of Evolutionary Algorithms. Berlin, Springer, 2015, 175 p. DOI: 10.1007/978-3-319-07407-8.
8. Törn A., Viitanen S., Floudas C. A., Pardalos P. M. (eds.), Topographical Global Optimization, *Recent Advances in Global Optimization*. Princeton University Press, 1992, pp. 384–398. DOI: 10.1515/9781400862528.384.
9. Wessing S., Preuss M. The True Destination of EGO is Multi-Local Optimization, *IEEE Latin American Conference on Computational Intelligence (LA-CCI), Arequipa, 8–10 November, 2017, proceedings*. Piscataway, IEEE, 2017, pp. 1–6. DOI: 10.1109/LA-CCI.2017.8285677.
10. Preuss M. Niching the CMA-ES via Nearest-Better Clustering, *Genetic and Evolutionary Computation Conference (GECCO), Portland, 7–11 July, 2010: proceedings*. New York, ACM, 2010, pp. 1711–1718. DOI: 10.1145/1830761.1830793.
11. Preuss M., Di Chio C., Agapitos A., Cagnoni S. et al. (eds.) Improved Topological Niching for Real-Valued Global Optimization, *Applications of Evolutionary Computation (EvoApplications), Málaga, 11–13 April, 2012: proceedings*. Berlin, Springer, 2012, Vol. 7248, pp. 386–395. DOI: 10.1007/978-3-642-29178-4_39. (Series “Lecture Notes in Computer Science”)
12. Li Y., Yu J., Takagi H. Niche Method Complementing the Nearest-Better Clustering, *IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, 6–9 December, 2019: proceedings*. Piscataway, IEEE, 2019, pp. 3065–3071. DOI: 10.1109/SSCI44817.2019.9002742.
13. Luo W., Qiao Y., Lin X. et al. Hybridizing Niching, Particle Swarm Optimization, and Evolution Strategy for Multimodal Optimization, *IEEE Transactions on Cybernetics: early access article*, 2020. DOI: 10.1109/TCYB.2020.3032995.
14. Ursem R. K. Multinational Evolutionary Algorithms / R. K. Ursem // *Congress on Evolutionary Computation (CEC99), Washington, 6–9 July, 1999, proceedings*. Piscataway, IEEE, 1999, Vol. 3, pp. 1633–1640. DOI: 10.1109/CEC.1999.785470.
15. Ursem R. K. Multinational GAs: Multimodal Optimization Techniques in Dynamic Environments, *Genetic and Evolutionary Computation Conference (GECCO), Las Vegas, 8–12 July, 2000, proceedings*. San Francisco, Morgan Kaufmann, 2000, Vol. 1, pp. 19–26.
16. Stoean C., Preuss M., Stoean R. et al. Disburdening the Species Conservation Evolutionary Algorithm of Arguing with Radii, *Genetic and Evolutionary Computation Conference (GECCO), London, 7–11 July, 2007: proceedings*. New York, ACM, 2007, pp. 1420–1427. DOI: 10.1145/1276958.1277220.
17. Stoean C., Preuss M., Stoean R. et al. Multimodal Optimization by Means of a Topological Species Conservation Algorithm, *IEEE Transactions on Evolutionary Computation*, 2010, Vol. 14, No. 6, pp. 842–864. DOI: 10.1109/TEVC.2010.2041668.
18. Cioppa A. D., Marcelli A., Napoli P. Speciation in Evolutionary Algorithms: Adaptive Species Discovery, *Genetic and Evolutionary Computation Conference (GECCO), Dublin, 12–16 July, 2011: proceedings*. New York, ACM, 2011, pp. 1053–1060. DOI: 10.1145/2001576.2001719.
19. Li L., Tang K. History-Based Topological Speciation for Multimodal Optimization, *IEEE Transactions on Evolutionary Computation*, 2015, Vol. 19, No. 1, pp. 136–150. DOI: 10.1109/TEVC.2014.2306677.
20. Maree S. C., Alderliesten T., Thierens D. et al. Real-Valued Evolutionary Multi-Modal Optimization Driven by Hill-Valley Clustering, *Genetic and Evolutionary Computation Conference (GECCO), Kyoto, 15–19 July, 2018: proceedings*. New York, ACM, 2018, pp. 857–864. DOI: 10.1145/3205455.3205477.

21. Maree S. C., Alderliesten T., Bosman P. A. N. Benchmarking HillValleA for the GECCO 2019 Competition on Multimodal Optimization [Electronic resource], 2019. Access mode: <https://arxiv.org/abs/1907.10988>.
22. Navarro R., Kim C. H. Niching Multimodal Landscapes Faster Yet Effectively: VMO and HillValleA Benefit Together, *Mathematics*, 2020, Vol. 8, No. 5 (665). DOI: 10.3390/math8050665.
23. Navarro R., Kim C. H. Improved Population Control for More Efficient Multimodal Optimizers, *Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI), Bogotá, 30 September – 2 October, 2020: proceedings*. Piscataway, IEEE, 2020, pp. 1–6. DOI: 10.1109/CONIITI51147.2020.9240365.
24. Ahrari A., Deb K., Preuss M. Multimodal Optimization by Covariance Matrix Self-Adaptation Evolution Strategy with Repelling Subpopulations, *Evolutionary Computation*, 2017, Vol. 25, No. 3, pp. 439–471. DOI: 10.1162/evco_a_00182.
25. Ahrari A., Deb K., Preuss M. Benchmarking Covariance Matrix Self Adaptation Evolution Strategy with Repelling Subpopulations for GECCO 2017 Competition on Multimodal Optimization : technical report : 2017014, *Computational Optimization and Innovation Laboratory (COIN), Michigan State University*. Michigan, 2017, 5 p.
26. Jamil M., Yang X.-S. A Literature Survey of Benchmark Functions for Global Optimization Problems / M. Jamil, // *International Journal of Mathematical Modelling and Numerical Optimisation*, 2013, Vol. 4, No. 2, pp. 150–194. DOI: 10.1504/IJMMNO.2013.055204.
27. Li X., Engelbrecht A., Epitropakis M. G. Benchmark Functions for CEC'2013 Special Session and Competition on Niching Methods for Multimodal Function Optimization : report, *Evolutionary Computation and Machine Learning Group, RMIT University*. Melbourne, 2013, 10 p.
28. Shylo V. P., Glybovets M. M., Gulayeva N. M. et al. Genetic Algorithm of Tournament Crowding Based on Gaussian Mutation, *Cybernetics and Systems Analysis*, 2020, Vol. 56, No. 2, pp. 231–242. DOI: 10.1007/s10559-020-00239-4.
29. Li J.-P., Balazs M. E., Parks G. T. et al. A Species Conserving Genetic Algorithm for Multimodal Function Optimization, *Evolutionary Computation*, 2002, Vol. 10, No. 3, pp. 207–234. DOI: 10.1162/106365602760234081.
30. Yao J., Kharna N., Grogono P. Bi-Objective Multipopulation Genetic Algorithm for Multimodal Function Optimization, *IEEE Transactions on Evolutionary Computation*, 2010, Vol. 14, No. 1, pp. 80–102. DOI: 10.1109/TEVC.2009.2017517.

Received 25.01.2021.
Accepted 23.04.2021.

УДК 004.023

ЕКСПЕРИМЕНТАЛЬНИЙ АНАЛІЗ БАГАТОНАЦІОНАЛЬНОГО ГЕНЕТИЧНОГО АЛГОРИТМУ ТА ЙОГО МОДИФІКАЦІЙ

Гулаєва Н. М. – канд. фіз.-мат. наук, доцент, доцент кафедри інформатики Національного університету «Києво-Могилянська академія», Київ, Україна.

Яремко С. А. – магістр, асистент кафедри інформатики Національного університету «Києво-Могилянська академія», Київ, Україна.

АНОТАЦІЯ

Актуальність. Генетичні алгоритми утворення ніш є одним з найпоширеніших підходів до розв'язання задач багатоекстремальної оптимізації. При проведенні класифікації цих алгоритмів можна виділити алгоритми, що ґрунтуються на явному аналізі топографії ландшафту функції пристосованості. Одним з ранніх прикладів таких алгоритмів є багатонаціональний генетичний алгоритм.

Мета. Розробка та аналіз багатонаціонального генетичного алгоритму та його модифікацій. Алгоритм застосовується для розв'язання задачі пошуку всіх максимумів багатоекстремальної функції.

Метод. Виконано експериментальний аналіз алгоритмів. Проведено численні прогони алгоритмів на відомих тестових задачах та обчислено критерії ефективності роботи алгоритмів, а саме, відсоток збіжності, частка реальних (глобальних, локальних) та хибних піків; зауважимо, що частки піків обчислюються тільки в разі збіжності алгоритму.

Результати. Виконано програму реалізацію багатонаціонального генетичного алгоритму та проведено експериментальне налаштування його параметрів. Запропоновано дві модифікації функції долин і пагорбів, яка використовується в алгоритмі для визначення взаємного розташування особин. Проведено експериментальний аналіз багатонаціонального генетичного алгоритму з класичною функцією долин і пагорбів та з її модифікаціями.

Висновки. Наукова новизна роботи полягає в тому, що були запропоновані модифікації функції долин і пагорбів, які продукують меншу кількість помилкових ідентифікацій зон притягання порівняно з класичним варіантом цієї функції. Як наслідок, використання цих модифікацій призводить до покращення продуктивності багатонаціонального генетичного алгоритму для низки тестових задач. Втім, для деяких тестових задач поліпшення критеріїв якості супроводжується зменшенням відсотка збіжності. Загалом, відсоток збіжності та значення критеріїв якості, продемонстровані дослідженням алгоритмом, є недостатніми для практичного використання багатонаціонального генетичного алгоритму у порівнянні з іншими відомими алгоритмами. У той же час, використання модифікованих функцій долин і пагорбів як етапу постобробки в інших алгоритмах утворення ніш видається перспективним підходом до покращення роботи цих алгоритмів.

КЛЮЧОВІ СЛОВА: задача багатоекстремальної оптимізації, генетичні алгоритми утворення ніш, багатонаціональний генетичний алгоритм, функція долин і пагорбів, збіжність генетичного алгоритму, частка реальних піків, частка хибних піків.

ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ МНОГОНАЦИОНАЛЬНОГО ГЕНЕТИЧЕСКОГО АЛГОРИТМА И ЕГО МОДИФИКАЦИЙ

Гулаева Н. М. – канд. физ.-мат. наук, доцент, доцент кафедры информатики Национального университета «Киево-Могилянская академия», Киев, Украина.

Яремко С. А. – магистр, ассистент кафедры информатики Национального университета «Киево-Могилянская академия», Киев, Украина.

АННОТАЦИЯ

Актуальность. Генетические алгоритмы образования ниш являются одним из самых распространенных подходов к решению задач многоэкстремальной оптимизации. При проведении классификации генетических алгоритмов образования ниш можно выделить алгоритмы, основанные на явном анализе топографии ландшафта функции приспособленности. Многонациональный генетический алгоритм является одним из ранних примеров таких алгоритмов.

Цель. Разработка и анализ многонационального генетического алгоритма и его модификаций. Алгоритм применяется для решения задачи поиска всех максимумов многоэкстремальной функции.

Метод. Выполнен экспериментальный анализ алгоритмов. Проведены многочисленные прогоны алгоритмов на известных тестовых задачах и вычислены критерии эффективности работы алгоритмов, а именно, процент сходимости, доля реальных (глобальных, локальных) и ложных пиков; отметим, что доля пиков вычисляется только в случае сходимости алгоритма.

Результаты. Выполнена программная реализация многонационального генетического алгоритма и проведена экспериментальная настройка его параметров. Предложены две модификации функции холмов и долин, используемой в алгоритме для определения взаимного расположения особей. Выполнен экспериментальный анализ многонационального генетического алгоритма с классической функцией холмов и долин и с ее модификациями.

Выводы. Научная новизна работы состоит в том, что были предложены модификации функции холмов и долин, приводящие к меньшему количеству неправильных определений зон притяжения по сравнению с классическим вариантом этой функции. Как следствие, использование этих модификаций приводит к повышению производительности многонационального генетического алгоритма для ряда тестовых задач. Однако для некоторых тестовых задач улучшение критериев качества сопровождается уменьшением процента сходимости. В целом, процент сходимости и значения критериев качества, продемонстрированные исследуемым алгоритмом, недостаточны для практического использования многонационального генетического алгоритма по сравнению с другими известными алгоритмами. В то же время, использование модифицированных функций холмов и долин в качестве шага постобработки в других алгоритмах образования ниш представляется многообещающим подходом к улучшению производительности этих алгоритмов.

КЛЮЧЕВЫЕ СЛОВА: задача многоэкстремальной оптимизации, генетические алгоритмы образования ниш, многонациональный генетический алгоритм, функция холмов и долин, сходимость генетического алгоритма, доля реальных пиков, доля ложных пиков.

ЛИТЕРАТУРА / LITERATURA

1. Li X. Seeking Multiple Solutions: An Updated Survey on Niching Methods and Their Applications / [X. Li, M. G. Epitropakis, K. Deb et al.] // IEEE Transactions on Evolutionary Computation. – 2017. – Vol. 21, № 4. – P. 518–538. DOI: 10.1109/TEVC.2016.2638437.
2. Shen Z.-H. Niche Pseudo-Parallel Genetic Algorithms for Path Optimization of Autonomous Mobile Robot / Z.-H. Shen, Y.-K. Zhao, W.-W. Wu // Journal of Shanghai University (English Edition). – 2006. – Vol. 10, № 5. – P. 449–453. DOI: 10.1007/s11741-006-0089-3.
3. Preuss M. Searching for Good and Diverse Game Levels / M. Preuss, A. Liapis, J. Togelius // IEEE Conference on Computational Intelligence and Games, Dortmund, 26–29 August, 2014: proceedings. – Piscataway: IEEE, 2014. – P. 1–8. DOI: 10.1109/CIG.2014.6932908.
4. Chica M. Multimodal Optimization: an Effective Framework for Model Calibration / [M. Chica, J. Barranquero, T. Kajdanowicz et al.] // Information Sciences. – 2017. – Vol. 375. – P. 79–97. DOI: 10.2139/ssrn.2828069.
5. Leoshchenko S. D. Modification and Parallelization of Genetic Algorithm For Synthesis of Artificial Neural Networks / [S. D. Leoshchenko, A. O. Oliinyk, S. A. Subbotin et al.] // Radio Electronics, Computer Science, Control. – 2019. – № 4. – P. 68–82. DOI: 10.15588/1607-3274-2019-4-7.
6. Glybovets M. M. Evolutionary Multimodal Optimization / M. M. Glybovets, N. M. Gulayeva // Optimization methods and applications. In honor of Ivan V. Sergienko's 80th birthday / S. Butenko, P. M. Pardalos, V. Shylo (eds.). – Cham : Springer, 2017. – P. 129–173. DOI: 10.1007/978-3-319-68640-0_8. – (Series “Springer Optimization and Its Applications”)
7. Preuss M. Multimodal Optimization by Means of Evolutionary Algorithms / M. Preuss. – Berlin : Springer, 2015. – 175 p. DOI: 10.1007/978-3-319-07407-8.
8. Törn A. Topographical Global Optimization / A. Törn, S. Viitanen ; C. A. Floudas, P. M. Pardalos (eds.) // Recent Advances in Global Optimization. – Princeton University Press, 1992. – P. 384–398. DOI: 10.1515/9781400862528.384.
9. Wessing S. The True Destination of EGO is Multi-Local Optimization / S. Wessing, M. Preuss // IEEE Latin American Conference on Computational Intelligence (LA-CCI), Arequipa, 8–10 November, 2017: proceedings. – Piscataway: IEEE, 2017. – P. 1–6. DOI: 10.1109/LA-CCI.2017.8285677.
10. Preuss M. Niching the CMA-ES via Nearest-Better Clustering / M. Preuss // Genetic and Evolutionary Computation Conference (GECCO), Portland, 7–11 July, 2010: proceedings. – New York: ACM, 2010. – P. 1711–1718. DOI: 10.1145/1830761.1830793.
11. Preuss M. Improved Topological Niching for Real-Valued Global Optimization / M. Preuss // Applications of Evolutionary Computation (EvoApplications), Málaga, 11–13

- April, 2012: proceedings / Di Chio C., Agapitos A., Cagnoni S. et al. (eds.). – Berlin: Springer, 2012. – Vol. 7248. – P. 386–395. DOI: 10.1007/978-3-642-29178-4_39. – (Series “Lecture Notes in Computer Science”)
12. Li Y. Niche Method Complementing the Nearest-Better Clustering / Y. Li, J. Yu, H. Takagi // IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, 6–9 December, 2019: proceedings. – Piscataway: IEEE, 2019. – P. 3065–3071. DOI: 10.1109/SSCI44817.2019.9002742.
 13. Luo W. Hybridizing Niching, Particle Swarm Optimization, and Evolution Strategy for Multimodal Optimization / [W. Luo, Y. Qiao, X. Lin et al.] // IEEE Transactions on Cybernetics: early access article. – 2020. DOI: 10.1109/TCYB.2020.3032995.
 14. Ursem R. K. Multinational Evolutionary Algorithms / R. K. Ursem // Congress on Evolutionary Computation (CEC99), Washington, 6–9 July, 1999: proceedings. – Piscataway: IEEE, 1999. – Vol. 3. – P. 1633–1640. DOI: 10.1109/CEC.1999.785470.
 15. Ursem R. K. Multinational GAs: Multimodal Optimization Techniques in Dynamic Environments / R. K. Ursem // Genetic and Evolutionary Computation Conference (GECCO), Las Vegas, 8–12 July, 2000: proceedings. – San Francisco: Morgan Kaufmann, 2000. – Vol. 1. – P. 19–26.
 16. Stoean C. Disburdening the Species Conservation Evolutionary Algorithm of Arguing with Radii / [C. Stoean, M. Preuss, R. Stoean et al.] // Genetic and Evolutionary Computation Conference (GECCO), London, 7–11 July, 2007: proceedings. – New York: ACM, 2007. – P. 1420–1427. DOI: 10.1145/1276958.1277220.
 17. Stoean C. Multimodal Optimization by Means of a Topological Species Conservation Algorithm / [C. Stoean, M. Preuss, R. Stoean et al.] // IEEE Transactions on Evolutionary Computation. – 2010. – Vol. 14, № 6. – P. 842–864. DOI: 10.1109/TEVC.2010.2041668.
 18. Cioppa A. D. Speciation in Evolutionary Algorithms: Adaptive Species Discovery / A. D. Cioppa, A. Marcelli, P. Napoli // Genetic and Evolutionary Computation Conference (GECCO), Dublin, 12–16 July, 2011: proceedings. – New York, ACM, 2011. – P. 1053–1060. DOI: 10.1145/2001576.2001719.
 19. Li L. History-Based Topological Speciation for Multimodal Optimization / L. Li, K. Tang // IEEE Transactions on Evolutionary Computation. – 2015. – Vol. 19, № 1. – P. 136–150. DOI: 10.1109/TEVC.2014.2306677.
 20. Maree S. C. Real-Valued Evolutionary Multi-Modal Optimization Driven by Hill-Valley Clustering / [S. C. Maree, T. Alderliesten, D. Thierens et al.] // Genetic and Evolutionary Computation Conference (GECCO), Kyoto, 15–19 July, 2018: proceedings. – New York, ACM, 2018. – P. 857–864. DOI: 10.1145/3205455.3205477.
 21. Maree S. C. Benchmarking HillValLEA for the GECCO 2019 Competition on Multimodal Optimization [Electronic resource] / S. C. Maree, T. Alderliesten, P. A. N. Bosman. – 2019. – Access mode: <https://arxiv.org/abs/1907.10988>.
 22. Navarro R. Niching Multimodal Landscapes Faster Yet Effectively: VMO and HillValLEA Benefit Together / R. Navarro, C. H. Kim // Mathematics. – 2020. – Vol. 8, №5 (665). DOI: 10.3390/math8050665.
 23. Navarro R. Improved Population Control for More Efficient Multimodal Optimizers / R. Navarro, C. H. Kim // Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI), Bogotá, 30 September – 2 October, 2020: proceedings. – Piscataway: IEEE, 2020. – P. 1–6. DOI: 10.1109/CONIITI51147.2020.9240365.
 24. Ahrari A. Multimodal Optimization by Covariance Matrix Self-Adaptation Evolution Strategy with Repelling Subpopulations / A. Ahrari, K. Deb, M. Preuss // Evolutionary Computation. – 2017. – Vol. 25, № 3. – P. 439–471. DOI: 10.1162/evco_a_00182.
 25. Ahrari A. Benchmarking Covariance Matrix Self Adaption Evolution Strategy with Repelling Subpopulations for GECCO 2017 Competition on Multimodal Optimization : technical report : 2017014 / A. Ahrari, K. Deb, M. Preuss / Computational Optimization and Innovation Laboratory (COIN), Michigan State University. – Michigan, 2017. – 5 p.
 26. Jamil M. A Literature Survey of Benchmark Functions for Global Optimization Problems / M. Jamil, X.-S. Yang // International Journal of Mathematical Modelling and Numerical Optimisation. – 2013. – Vol. 4, № 2. – P. 150–194. DOI: 10.1504/IJMMNO.2013.055204.
 27. Li X. Benchmark Functions for CEC'2013 Special Session and Competition on Niching Methods for Multimodal Function Optimization : report / X. Li, A. Engelbrecht, M. G. Epitropakis / Evolutionary Computation and Machine Learning Group, RMIT University. – Melbourne, 2013. – 10 p.
 28. Шило В. П. Генетичні алгоритми турнірного витиснення з гаусовою мутацією / [В. П. Шило, М. М. Глибовець, Н. М. Гуласва та ін.] // Кибернетика и системный анализ. – 2020. – Том 56, № 2. – С.75–88.
 29. Li J.-P. A Species Conserving Genetic Algorithm for Multimodal Function Optimization / [J.-P. Li, M. E. Balazs, G. T. Parks et al.] // Evolutionary Computation. – 2002. – Vol. 10, № 3. – P. 207–234. DOI: 10.1162/106365602760234081.
 30. Yao J. Bi-Objective Multipopulation Genetic Algorithm for Multimodal Function Optimization / J. Yao, N. Kharna, P. Grogono // IEEE Transactions on Evolutionary Computation. – 2010. – Vol. 14, № 1. – P. 80–102. DOI: 10.1109/TEVC.2009.2017517.

THE AUTOMATIC SYNTHESIS OF PETRI NETS BASED ON THE FUNCTIONING OF ARTIFICIAL NEURAL NETWORK

Gurskiy A. A. – PhD, Associate Professor of the Department of Technological Processes Automation and Robot-technical Systems, Odessa National Academy of Food Technologies, Odessa, Ukraine.

Denisenko A. V. – Lecturer of the Department of Information Systems, Odessa National Polytechnic University, Odessa, Ukraine.

Dubna S. M. – Lecturer of the Department of Technological Processes Automation and Robot-technical Systems, Odessa National Academy of Food Technologies, Odessa, Ukraine.

ABSTRACT

Context. The important task was solved during the scientific research related to the development of the methods for automatic synthesis of Petri nets while tuning up of the coordinating automatic control systems. The importance of development of these methods is due to the evolution of intelligent systems. These systems provide the automation of labor intensive processes in the particular case this is the tuning of the certain type of complex control systems.

Objective. The purpose of the scientific work is to minimize the time and automation of process in tuning of the multilevel coordinating automatic control systems.

Method. The principle of automatic synthesis of Petri nets and the implementation of certain algorithms for tuning complex control systems based on the functioning of an artificial neural network are proposed. The mathematical description of the method for changing the coefficients in neural connections of network in the synthesis of Petri nets is presented.

Results. The experiments were conducted in the Matlab/Simulink 2012a environment. These experiments were bound to the joint functioning of an artificial neural network and Petri nets. The functioning of Petri nets was presented in the Matlab \ Simulink environment using Statflow diagrams.

As a result of the experiments we have obtained the temporal characteristics of the functioning of artificial neural network providing the composition of Petri nets. The fundamental suitability of using artificial neural network to provide the automatic composition of Petri nets was determined on the basis of analysis of temporal characteristics.

Conclusion. The problem linked to the development of system for the joint functioning of neural network and Petri nets for the formation of algorithms and sequential calculations was solved in this work. Thus the method of automatic synthesis of Petri nets and the method of developing of the certain algorithms based on the functioning of a neural network were further developed.

KEYWORDS: Petri net, artificial neural networks, coordinating automatic control system, algorithms of tuning, automatic synthesis.

ABBREVIATIONS

PID is a proportional-integral-derivative control;
ANN is an artificial neural network;
ACS is an automatic control system.

NOMENCLATURE

M_k is an $m \times 1$ vector, the marking after the k -th firing;

$|A|$ is an $n \times m$ matrix, the incidence matrix of Petri net;

U_k is an $n \times 1$ vector, the control vector indicating the transition fired at the k -th firing;

$\mu(p_i)$ is a value of place p_i marking;

$w_{i,j}(t_k)$ is the synaptic weight, which determines the corresponding connection in the formed Petri net at step t_k ;

$\varphi(t)$ is a deviation from ratio of the parameters at time t ;

X_1 is an actual value of the controlled variable;

X_{3Z} is a set point of the controlled variable;

$e(t)$ is a deviation of the controlled variable within time from set point;

A^T is a coefficient matrix;

J is an integral criterion of system;

u_1 , is a control action;

u_a is a control action to change the ratio;

$|W|$ is an incidence matrix of the discrete-event part of the system;

$v_L^d(t_k)$ is a control vector which is formed depending on the fulfillment of conditions for triggering of Petri net transitions;

$u_L^d(t_k)$ is an input action.

INTRODUCTION

The algorithms or sequences of certain implementing computer instructions are usually presented in the form of block diagrams. Such graphical describing of the algorithm is most widespread.

Petri nets are also known as forms of the graphical representation of parallel algorithms and computations. Nevertheless in general, Petri nets are mathematical models of the discrete dynamical systems.

Automatic synthesis of Petri nets is the field of algorithms formation. These algorithms can represent the functioning of discrete or hybrid systems [1, 2].

In a particular case the automatic synthesis of Petri nets is expedient in the formation of algorithms for the step-by-step tuning of the certain kind of multi-level automatic control systems. In this case the formation of al-

gorithms and Petri nets is a necessary component for the automation of the process of tuning in multi-level coordinating control systems [3].

Automation of the tuning process for multi-level automatic control systems is an urgent task as the tuning process is quite difficult. The complexity of the tuning is due to the fact that as a result it is necessary to obtain the temporal subordination of the control processes in the system. If there is no temporal subordination of control processes then the automatic control system will cease to be coordinating.

The object of study is the processes of automatic synthesis of Petri nets while tuning up of the coordinating automatic control systems.

The subject of study is the methods of automatic synthesis of Petri nets while tuning up of the coordinating automatic control systems.

The purpose of the scientific work is to minimize the time and automate of process in tuning of the multi-level coordinating automatic control systems.

1 PROBLEM STATEMENT

In order to achieve this purpose we need to develop methods for automatic synthesis of Petri nets which will represent the processes of tuning in the control systems. Considering that the formation of an algorithm is an intelligent process, then it is expedient to synthesize Petri nets based on the functioning of an artificial neural network. In this case the artificial neural network represents an intelligent technology for forming the specific algorithm. Obviously, the training of the neural network should carry out on the basis of the incident matrix of the generated Petri net and on the basis of the values of the performance indicators of the tuning system.

The positive result of automatic synthesis of the Petri net and the corresponding algorithm is the minimization of the values of integral indicators J_1, J_2, J_3 of functioning at various stages of the tuning.

In this case, $J_1 = \int_0^{\infty} |\varphi(t)| dt$ is an integral indicator of functioning for the coordinating level of the control system;

$J_2 = \int_0^{\infty} (|e(t)| + \alpha \cdot |\varphi(t)|) dt$ is an integral indicator of functioning for the stabilizing level in the control system;

$J_3 = \int_0^{\infty} (|e(t)| + \beta \cdot |\varphi(t)|) dt$ is an integral indicator repre-

senting the quality of the functioning of the system for tuning to the temporal subordination of the control processes; α, β is coefficients determining the temporal subordination of the control processes ($\alpha < \beta$); $\varphi(t)$ is deviation from the ratio of the values of regulated variables; $e(t)$ is the deviation of some variable during time from the given value.

As a result we have to determine all values of the tuning parameters $K \in k_{ij}$ of various levels in the control

system according to the formed algorithm. In this case we obtain the minimum values of the J_1, J_2, J_3 integral indicators of the multilevel system.

2 REVIEW OF THE LITERATURE

The relevance of developing methods for automatic synthesis of Petri nets was presented in the scientific work of J. Peterson 1981 [4]. After J. Peterson's work it has appeared a number of scientific works linked to the automatic generation of Petri nets [5–7]. The certain synthesis methods based on the composition of individual Petri nets are presented in these works. There are not many scientific publications in this field, however, we can see the certain trend in the development of the corresponding scientific direction.

Taking into account the peculiarities of the synthesis of Petri nets, we have determined to form them on the basis of the functioning of an artificial neural network [8]. Thus, the artificial neural network is an intelligent system for the automatic synthesis of Petri nets and certain algorithms.

Currently there are lots of scientific publications linked to the integration of mathematical apparatus of Petri nets and artificial neural networks [9–11]. However, this integration is not linked to the automatic synthesis of Petri nets. Thereby, the organization of the joint functioning of the artificial neural network and Petri nets for the automatic synthesis of algorithms is presented in this article [8]. The first attempt in the development of methods for training of the artificial neural network linked to the synthesis of Petri nets is also presented in the scientific work [8].

The mathematical description of the process for training of the artificial neural network for the synthesis of Petri nets is considered in this paper. This synthesized Petri net represents the step-by-step algorithm for tuning the special class of coordinating automatic control systems shown in papers [12, 13].

3 MATERIALS AND METHODS

The features of automatic synthesis of Petri nets based on the functioning of the artificial neural network can be determined from Figure 1. The simplified block diagram of the coordinating control system with automatic tuning of its parameters is shown in Figure 1. This tuning of parameters in the regulators is occurred on the basis of the joint functioning of the artificial neural network with the set of Petri nets. Petri nets marking determines the increment of the values for the corresponding parameters in the regulators to achieve the minimum value of the integral criteria of the control system. As shown in Figure 1 the artificial neural network interacts on the principles of feedback with synchronously functioning Petri nets. We can form the composition from these acting simultaneously nets. This composition will represent the certain algorithm of actions implemented by the artificial neural network. Figure 2a shows Petri net composition. This composition presents the algorithm for tuning the coordi-

nating automatic control system. The functioning of such Petri net we can describe by the following equation:

$$M_k = M_{k-1} + |A| \cdot U_k, \quad (1)$$

Where $M_k = [\mu(p_0), \dots, \mu(p_n)]^T$; M_{k-1} is the marking of Petri net at $k-1$ step.

In this case, if $\mu(p_i) = 1$ (where $i = 1 \dots n$), then the value of the corresponding parameter changes during tuning up of the system.

If we consider that the simplified structural diagram shown in Figure 1 represents logical-dynamic model, then the discrete-event part of the model can be described by the following equation:

$$X_L^d(t_k) = X_L^d(t_{k-1}) + |W| \cdot v_L^d(t_k) + u_L^d(t_k), \quad (2)$$

where $X_L^d(t_k) = [\mu(p_1^2), \mu(p_2^2), \dots, \mu(p_8^2)]^T$; $X_L^d(t_k)$, $X_L^d(t_{k-1})$ is discrete states in times t_k, t_{k-1} .

If the place p_i is marked – $\mu(p_i) = 1$ then k_i parameter should be increased i.e., $\mu(p_1) = 1 \Rightarrow \uparrow k_1$ in this way $\mu(p_2) = 1 \Rightarrow \downarrow k_1$, $\mu(p_3) = 1 \Rightarrow \uparrow k_2$, $\mu(p_4) = 1 \Rightarrow \downarrow k_2$,

$$\mu(p_5) = 1 \Rightarrow \uparrow k_3, \quad \mu(p_6) = 1 \Rightarrow \downarrow k_3, \quad \mu(p_7) = 1 \Rightarrow \uparrow k_4, \\ \mu(p_8) = 1 \Rightarrow \downarrow k_4.$$

According to Figure 1, k_1 is the tuning parameter of the Controller 1 in the coordinating control level, k_2 is the tuning parameter of the Controller 2, k_3 is the tuning parameter of the Controller 3 in the stabilizing control level.

The equations (2) describe the joint functioning of Petri nets and the artificial neural network.

The incidence matrix of Petri net $|A|$ has the certain analogy with the matrix of the weights of the output layer in the neural network. So the artificial neural network generates the output signals $\bar{V} = |A| \cdot U_k$ corresponding to the values of the incidence matrix of the formed Petri net.

When tuning up of the system if the certain algorithm of actions is wrong then we must indicate the element of the Petri net which is erroneous. This is necessary for the reconfiguring of artificial neural network. For example, as shown in Figure 2, if the value of the integral indicators of the system increases, then it is necessary to eliminate the connection between the transition t_i and the place p_i respectively.

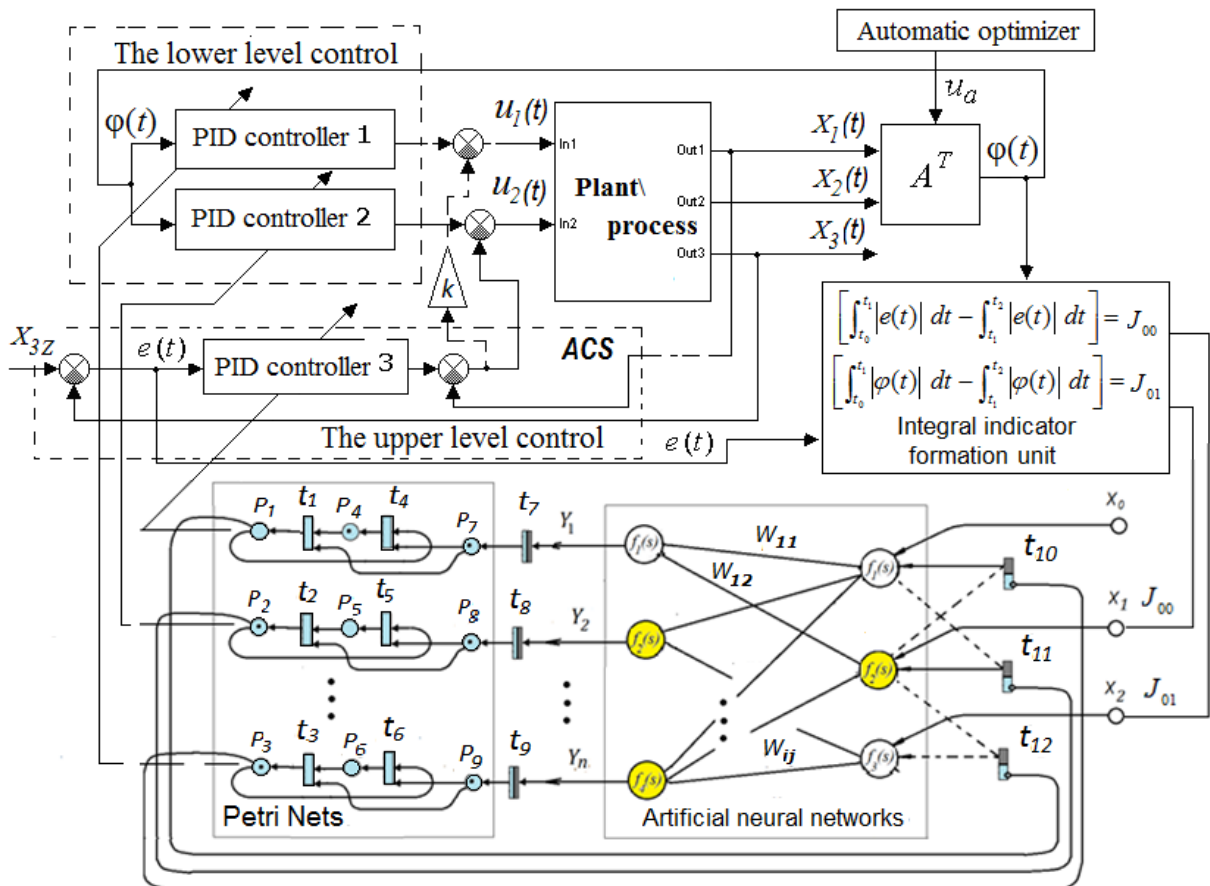


Figure 1 – The simplified block diagram of the coordinating automatic control system with automatic tuning of its parameters

But in this case it is necessary to add the new connection between the transition t_i and the neighboring place p_{i+1} . Thus, the Petri net shown in Figure 2a changes to the Petri net shown in Figure 2b. The coefficients of interneural connections in the artificial neural network must change accordingly.

It can be represented mathematically like this:

$$w_{i,j}(t_{k+1}) = w_{i,j}(t_k) - w_{i,j}(t_k) \cdot (\delta_{p,i} + \delta_{p,i+1}) + w_{i-1,j}(t_k) \cdot \delta_{p,i} + w_{i+1,j}(t_k) \cdot \delta_{p,i+1} \quad \text{if } i=1,3,5\dots \quad (3)$$

$$w_{i,j}(t_{k+1}) = w_{i,j}(t_k) - w_{i,j}(t_k) \cdot (\delta_{p,i} + \delta_{p,i-1}) + w_{i+1,j}(t_k) \cdot \delta_{p,i} + w_{i-1,j}(t_k) \cdot \delta_{p,i-1} \quad \text{if } i=2,4,6\dots \quad (4)$$

$$w_{i,j}(t_{k+1}) = w_{i,j}(t_k) - \delta_{t,j} \cdot w_{i,j}(t_k) + \delta_{i,j} \cdot w_{i-1,j}(t_k) \quad \text{if } i=1,3,5\dots \quad (5)$$

$$w_{i,j}(t_{k+1}) = w_{i,j}(t_k) - \delta_i \cdot w_{i,j}(t_k) + \delta_i \cdot w_{i+1,j}(t_k) \quad \text{if } i=2,4,6\dots \quad (6)$$

Where $w_{i,j}(t_{k+1})$ is the synaptic weight at step t_{k+1} .

In this instance the matrix of the coefficients for interneural connections of the output layer of the neural network we can show as:

$$N = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1j} & w_{1,j+1} & \dots \\ w_{21} & w_{22} & \dots & w_{2,j} & w_{2,j+1} & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots \\ w_{i,1} & w_{i,2} & \dots & w_{i,j} & w_{i,j+1} & \dots \\ w_{i+1,1} & w_{i+1,2} & \dots & w_{i+1,j} & w_{i+1,j+1} & \dots \end{bmatrix}$$

The matrix N of coefficients of interneural connections has a certain analogy with the incident matrix of Petri net. For example, the matrix N is analogous to the incident matrix W_1 of Petri net shown in Figure 2a. If the place p_1 has the error then the incident matrix W_1 changes accordingly into the matrix W_2 as follows:

$$W_1 = \begin{matrix} & t_1 & t_2 & t_3 & t_4 \\ p_1 & +1 & -1 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & +1 & -1 \\ p_4 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 \\ p_6 & 0 & +1 & -1 & 0 \\ p_7 & -1 & 0 & 0 & 0 \\ p_8 & 0 & 0 & 0 & +1 \\ p_9 & 0 & 0 & 0 & 0 \end{matrix} \Rightarrow W_2$$

$$W_2 = \begin{matrix} & t_1 & t_2 & t_3 & t_4 \\ p_1 & 0 & 0 & 0 & 0 \\ p_2 & +1 & -1 & 0 & 0 \\ p_3 & 0 & 0 & +1 & -1 \\ p_4 & 0 & 0 & 0 & 0 \\ p_5 & 0 & 0 & 0 & 0 \\ p_6 & 0 & +1 & -1 & 0 \\ p_7 & -1 & 0 & 0 & 0 \\ p_8 & 0 & 0 & 0 & +1 \\ p_9 & 0 & 0 & 0 & 0 \end{matrix}$$

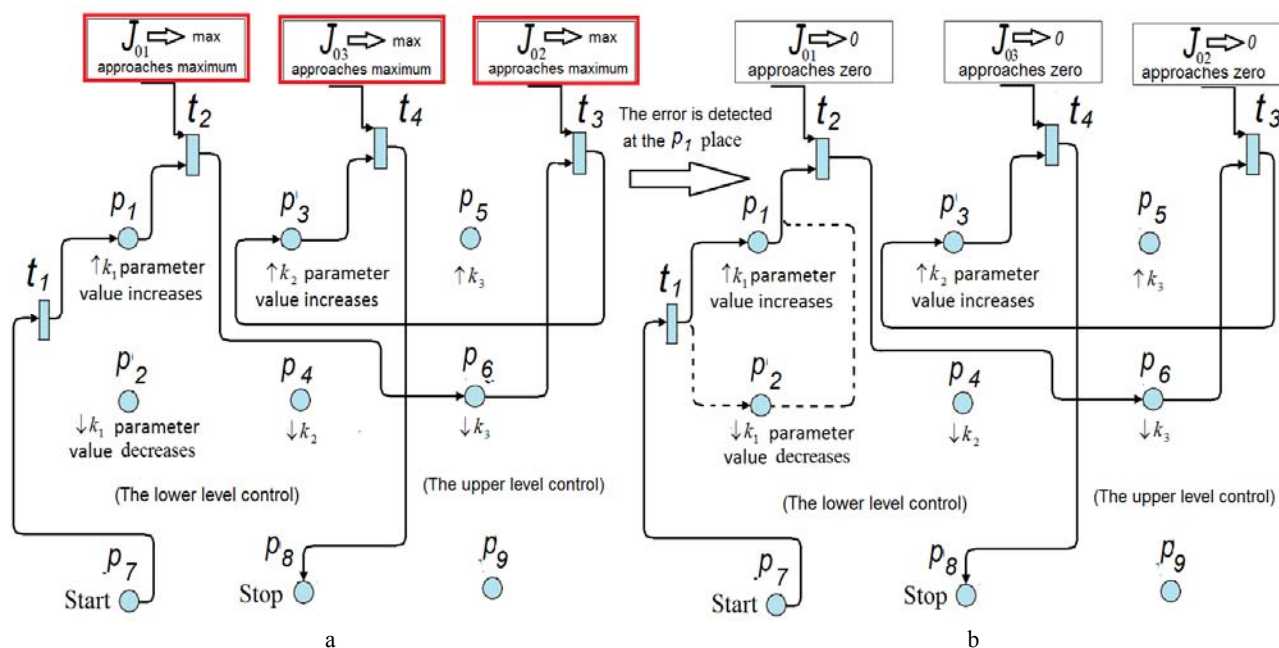


Figure 2 – The Petri nets representing the tuning algorithms of the coordinating automatic control system

The p_1 and p_2 rows of W_1 matrix were changed according to expressions (3).

if an error is detected at the p_i or p_{i+1} place then we have $\delta_{p,i} = 1$ or $\delta_{p,i+1} = 1$ accordingly. In this instance if there is a connection between the transition t_i and the place p_i then $w_{i,j}(t_k) \cdot (\delta_{p_i} + \delta_{p_{i+1}}) = |1|$. So according to the expression, the coefficients of interneuronal connections at a t_{k+1} step will become equal to zero $w_{i,j}(t_{k+1}) = 0$. Thus the corresponding connections with the place p_i in the forming Petri net will disappear. In so doing the new connections must appear with the place p_i if $w_{i-1,j}(t_k) \cdot \delta_{p,i} = 1$ or if $w_{i+1,j}(t_k) \cdot \delta_{p,i+1} = 1$.

Similarly, if an error is detected at the transition t_j then we have $\delta_{t,j} = 1$ accordingly. In this instance if there is a connection $\delta_{t,j} \cdot w_{i,j}(t_k) = 1$ and $w_{i,j}(t_{k+1}) = 0$. In the absence of connection $\delta_{t,j} \cdot w_{i+1,j}(t_k) = 1$ and $w_{i,j}(t_{k+1}) = 1$.

4 EXPERIMENTS

In the software environment MATLAB \ Simulink 2012 the experiments were carried out linked to the joint functioning of the artificial neural network with the Petri nets. The functioning of Petri nets in the MATLAB \ Simulink software environment was implemented using a Stateflow diagram. The fragment of Stateflow diagram representing the functioning of Petri nets is shown in Figure 3. State 1, State 2, State 3 and State 4 are states of the same Petri net. The artificial neural network is linked with the functioning of three such Petri nets at once. This is shown in Figure 4. Figure 4 is also shown the two-layer

artificial neural network implemented by means of the MATLAB \ Simulink environment. This artificial neural network consists of six neurons and has three output signals $Xp1, Xp2, Xp3$ connected with Petri nets.

The structural diagram of the neural network and Petri nets is similar to the simplified diagram shown in Figure 1. In Figures 3 and 4 we can see all the necessary parameters of a system representing the joint functioning of the neural network and Petri nets. This system is intended for the automatic formation of Petri nets and certain algorithms.

5 RESULTS

The system shown by the block diagram in Figure 4 can represent the functioning of different Petri nets.

For example, Figures 4 and 5 we can see timing diagrams of the functioning of Petri net consisting of three positions and three transitions. The functioning of such Petri net is a system of joint function of the artificial neural network and Stateflow diagrams.

The output signals of the artificial neural network $Xp1, Xp2, Xp3$ correspond to the incidence matrix of the generated Petri net. The output signals $\mu(p_1), \mu(p_2), \mu(p_3)$ show the change in marking of Petri net for time. In this case the work of the formed Petri net is realized as shown in Figure 5.

According to timing diagrams, when we have the signal at J_1 input ($J_{01} > 0$) then transition t_2 is triggered. If we have the signal at J_2 input ($J_{01} > 1$) then transition t_3 is triggered. The simultaneous triggered of t_2 and t_3 transition corresponds to the conflict situation in the functioning of the Petri net.

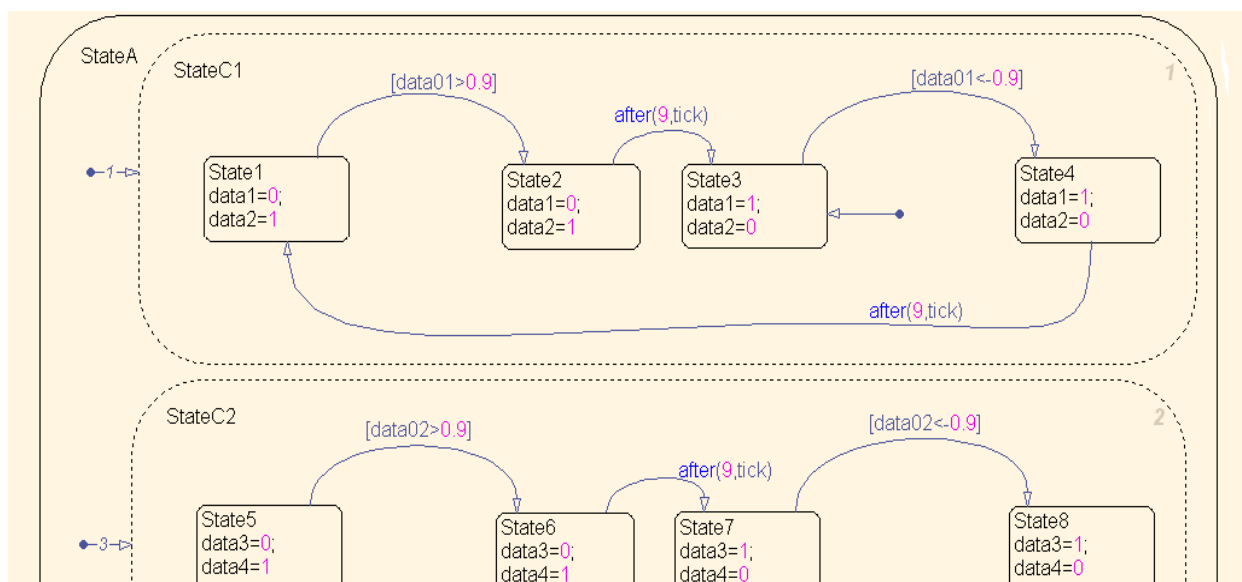


Figure 3 – Stateflow diagrams representing the functioning of Petri nets

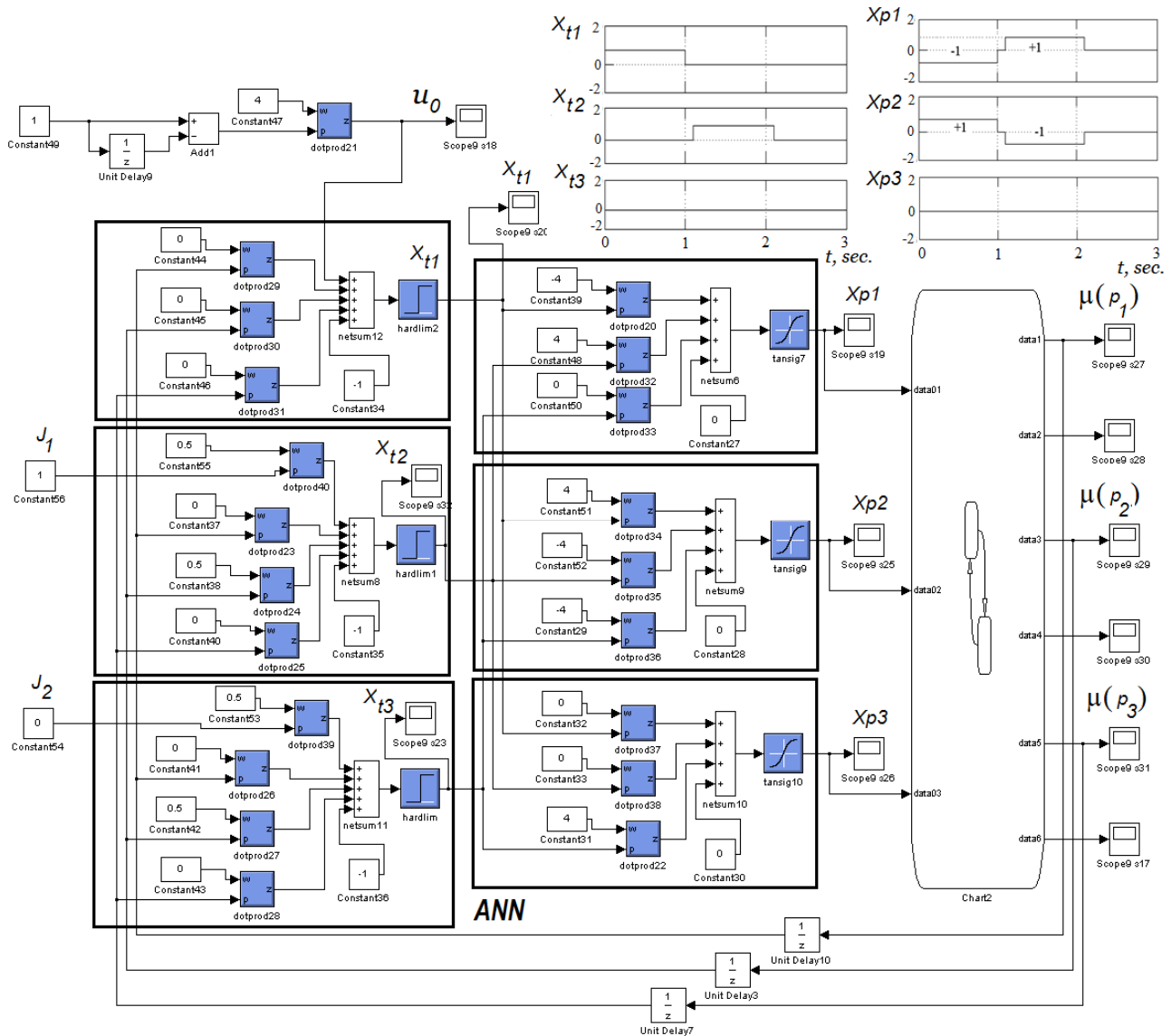


Figure 4 – Block diagram of the neural network implementing the synthesis of Petri net

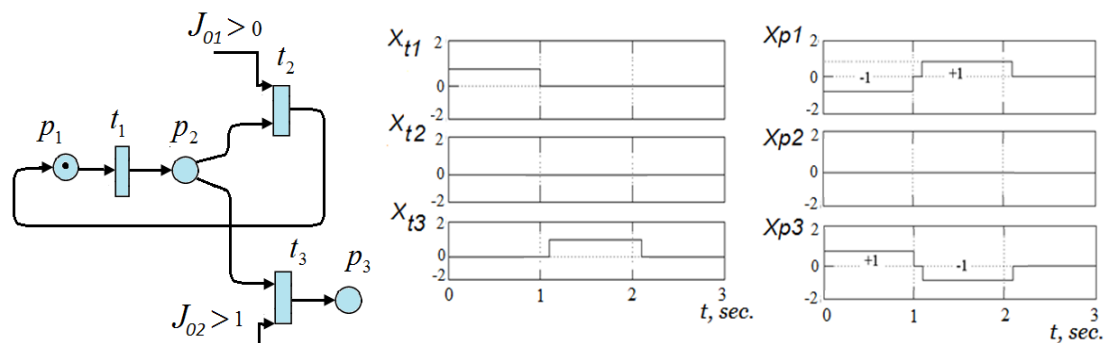


Figure 5 – The timing diagrams of the functioning of the synthesized Petri net

Analyzing the temporal characteristics we can conclude about the principal suitability of the considered system presented in Figures 4 and 5. This system is able to represent the functioning of various Petri nets. Tuning the parameters of an artificial neural network of such system

is linked with the automatic synthesis of Petri net. Thus, the change in the coefficients of interneuronal connections when tuning the network is linked with the change in the synthesized Petri net.

6 DISCUSSION

The automatic synthesis of Petri nets we can imagine in two stages. At the first stage the system realizes the selection of a certain algorithm and the corresponding Petri net among the possible variants. At the second stage the system realizes the correction of the selected algorithm and Petri net.

These two stages can be realized using an artificial neural network and its tuning up. For this case the system of joint functioning of the artificial neural network with Petri nets was designed (Figure 5). In order to compose the data of Petri nets into one common network we can use the output signals of the artificial neural network. These signals are required to form the corresponding incidence matrix.

If the generated Petri net is unsatisfactory then we must indicate which net element has the error. Then the system has to start the corresponding change in the coefficients of the interneural connections in the neural network.

The mathematical description of the change in the coefficients for interneuronal connections during training of the network was presented in the scientific work as one of the attempts to implement the enumeration of possible variants of connections in Petri net [14]. In this case the formed Petri net is a peculiar visual reflection of the coefficient set of interneural connections in the artificial neural network.

CONCLUSIONS

The scientific novelty of the results. The problem linked to with the development of the system for the joint functioning of the artificial neural network and Petri nets was solved in the present work. The developed system is necessary for providing of the automatic synthesis of some algorithms. Thus the corresponding method of automatic synthesis of Petri net has got the further development. In this paper we have presented the algorithm of change in the coefficients of interneural connections of the network during the synthesis of Petri net.

The practical significance of the results. The completed scientific research has confirmed the suitability of the developed system of the joint functioning of the artificial neural network and Petri nets for providing the automatic synthesis for some algorithms. Due to the method of automatic synthesis of Petri nets we can solve the problem of automatized tuning for models of the complicated control systems providing the coordination of various transients.

The prospects for further research. Further development of the scientific direction must be directly related to the formation of training methods of neural networks for the synthesis of Petri nets. It is also necessary to develop more complex models of systems showing the joint functioning of an artificial neural network and Petri nets.

ACKNOWLEDGMENTS

The scientific work was carried out at the department of technological processes automation and robotic systems.

The authors thank the management of the Odessa Refinery Plant too. The Odessa Refinery Plant has given the experimental data for designing models of the coordinating automatic control system.

REFERENCES

1. Yang Z., Blanke M. A unified approach to controllability analysis for hybrid control systems, *Nonlinear Analysis: Hybrid Systems*, 2007, Volume 1, Issue 2, pp. 212–222. DOI: 10.1016/j.nahs.2006.08.002
2. Drighiciu, M. A. Hybrid Petri nets: A framework for hybrid systems modeling, *2017 International Conference on Electromechanical and Power Systems (SIELMEN)*, IEEE, 2017, pp. 020–025. DOI: 10.1109/SIELMEN.2017.8123285
3. Filimonov A. B., Filimonov N. B. O problematike sinteza koordiniruyuschih sistem avtomaticheskogo upravleniya, *Izvestiya SfedU, Engineering sciences*, 2012, Vol. 3, pp. 172–180. ISSN 1999–9429 2311–3103
4. Peterson, J. L. Petri net theory and the modeling of systems / Prentice Hall PTR, 1981, 290 p. ISBN 0–13–661983–5
5. He D. W., Strege B., Tolle H., Kusiak A. Decomposition in automatic generation of Petri nets for manufacturing system control and scheduling, *International Journal of Production Research*, 2000, Volume 38, Issue 6, pp. 1437–1457. DOI: 10.1080/002075400188942
6. Ndiaye M. A., Petin J. F., Camerini J., Georges J. P. Performance assessment of industrial control system during pre-sales uncertain context using automatic Colored Petri Nets model generation, *2016 International Conference on Control, Decision and Information Technologies (CoDIT)*, IEEE, 2016, pp. 671–676. DOI: 10.1109/CoDIT.2016.7593643
7. Durmuş M. S., Yıldırım U., Söylemez M. T. Automatic generation of Petri Net supervisors for railway interlocking design, *2012 2nd Australian Control Conference*, IEEE, 2012, pp. 180–185. ISBN: 978-1-922107-63-3
8. Gurskiy A. A., Dubna S. M. Nastroyka neuronnoi seti pri avtomaticheskoy sinteze setei Petri, *Automation of technological and business processes*, 2018, No. 1, pp. 22–32. DOI: <https://doi.org/10.15673/atbp.v10i1.877>
9. Ahson, S. I. Petri net models of fuzzy neural networks, *IEEE Transactions on systems, man, and cybernetics*, 1995, Volume 38, Issue 6, pp. 926–932. ISSN: 2168-2909
10. Chow T. W., Li J. Y. Higher-order Petri net models based on artificial neural networks, *Artificial Intelligence*, 1997, Volume 92, Issues 1–2, pp. 289–300. DOI: [https://doi.org/10.1016/S0004-3702\(96\)00048-3](https://doi.org/10.1016/S0004-3702(96)00048-3)
11. Hanna M. M., Buck A., Smith R. Fuzzy Petri nets with neural networks to model products quality from a CNC-milling machining centre, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 1996, Vol. 26, Issue 5, pp. 638–645. ISSN: 1558-2426. DOI: 10.1109/3468.531910
12. Boichuk L. M. Sintez koordiniruyushih sistem avtomaticheskogo upravleniya. Moscow, Energoatomizdat, 1991, 160 p. ISBN 5–283–01521–1
13. Gurskiy A. A., Goncharenko A. E., Dubna S. M. Algorithms for tuning of the coordinating automatic control systems [Text], *Radio electronics, computer science, control*, 2020,

УДК 681.513

АВТОМАТИЧНИЙ СИНТЕЗ МЕРЕЖ ПЕТРІ НА ОСНОВІ ФУНКЦІОНУВАННЯ ШТУЧНОЇ НЕЙРОННОЇ МЕРЕЖІ

Гурський О. О. – канд. техн. наук, доцент кафедри автоматизації технологічних процесів і робототехнічних систем інституту комп'ютерних систем і технологій «Індустрія 4.0» ім. П. Н. Платонова Одеської національної академії харчових технологій, Одеса, Україна.

Денисенко А. В. – старший викладач кафедри інформаційних систем Одеського національного політехнічного університету, Одеса, Україна.

Дубна С. М. – старший викладач кафедри автоматизації технологічних процесів і робототехнічних систем інституту комп'ютерних систем і технологій «Індустрія 4.0» ім. П. Н. Платонова Одеської національної академії харчових технологій, Одеса, Україна.

АНОТАЦІЯ

Актуальність. Вирішена актуальна задача, що пов'язана з розробкою методів автоматичного синтезу мереж Петрі для настроювання координувальних систем автоматичного управління. Важливість розробки даних методів обумовлена розвитком інтелектуальних систем, що забезпечують автоматизацію трудомістких процесів, в окремому випадку це настроювання певного класу складних систем управління.

Мета роботи – мінімізація часу та автоматизація процесу настроювання багаторівневих координувальних систем автоматичного управління.

Метод. Запропоновано принцип автоматичного синтезу мереж Петрі та певних алгоритмів настроювання складних систем управління на основі функціонування штучної нейронної мережі. Представлений математичний опис методу зміни коефіцієнтів міжнейронних зв'язків мережі при синтезі мереж Петрі.

Результати. У програмному середовищі Matlab/Simulink 2012a були проведені експерименти, пов'язані зі спільним функціонуванням штучної нейронної мережі і мереж Петрі. Функціонування мереж Петрі в середовищі Matlab/Simulink було представлено за допомогою Statflow діаграм.

У результаті експериментів були отримані часові характеристики функціонування штучної нейронної мережі, яка забезпечує композицію мереж Петрі. На основі часових характеристик була встановлена принципова придатність застосування штучної нейронної мережі для забезпечення автоматичної композиції мереж Петрі.

Висновки. У даній роботі було вирішено задачу, яка пов'язана з розробкою системи спільного функціонування нейронної мережі і мереж Петрі для формування алгоритмів і послідовних обчислень. Тим самим, одержали подальший розвиток методика автоматичного синтезу мереж Петрі та методика розробки певних алгоритмів на основі функціонування нейронної мережі.

КЛЮЧОВІ СЛОВА: штучна нейронна мережа, мережі Петрі, автоматичний синтез, координувальна система, алгоритм настроювання.

УДК 681.513

АВТОМАТИЧЕСКИЙ СИНТЕЗ СЕТЕЙ ПЕТРИ НА ОСНОВЕ ФУНКЦИОНИРОВАНИЯ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

Гурский А. А. – канд. техн. наук, доцент кафедры автоматизации технологических процессов и робототехнических систем института компьютерных систем и технологий «Индустрия 4.0» им. П. Н. Платонова Одесской национальной академии пищевых технологий, Одесса, Украина.

Денисенко А. В. – старший преподаватель кафедры информационных систем Одесского национального политехнического университета, Одесса, Украина.

Дубна С. М. – старший преподаватель кафедры автоматизации технологических процессов и робототехнических систем института компьютерных систем и технологий «Индустрия 4.0» им. П. Н. Платонова Одесской национальной академии пищевых технологий, Одесса, Украина.

АННОТАЦИЯ

Актуальность. Решена актуальная задача, связанная с разработкой методов автоматического синтеза сетей Петри для настройки координирующих систем автоматического управления. Важность разработки данных методов обусловлена развитием интеллектуальных систем, обеспечивающих автоматизацию трудоемких процессов, в частном случае это настройка определенного класса сложных систем управления.

Цель работы – минимизация времени и автоматизация процесса настройки многоуровневых координирующих систем автоматического управления.

Метод. Предложен принцип автоматического синтеза сетей Петри и определенных алгоритмов настройки сложных систем управления на основе функционирования искусственной нейронной сети. Представлено математическое описание метода изменения коэффициентов межнейронных связей сети при синтезе сетей Петри.

Результаты. В программной среде Matlab\Simulink 2012a были проведены эксперименты, связанные с совместным функционированием искусственной нейронной сети и сетей Петри. Функционирование сетей Петри в среде Matlab\Simulink было представлено с помощью Statflow диаграмм.

В результате экспериментов были получены временные характеристики функционирования искусственной нейронной сети, обеспечивающей композицию сетей Петри. На основе временных характеристик была установлена принципиальная пригодность применения искусственной нейронной сети для обеспечения автоматической композиции сетей Петри.

Выводы. В настоящей работе была решена задача, связанная с разработкой системы совместного функционирования нейронной сети и сетей Петри для формирования алгоритмов и последовательных вычислений. Тем самым, получили дальнейшее развитие методика автоматического синтеза сетей Петри и методика разработки определенных алгоритмов на основе функционирования нейронной сети.

КЛЮЧЕВЫЕ СЛОВА: искусственная нейронная сеть, сети Петри, автоматический синтез, координирующая система, алгоритм настройки.

ЛІТЕРАТУРА / LITERATURA

1. Yang Z. A unified approach to controllability analysis for hybrid control systems / Z. Yang, M. Blanke // *Nonlinear Analysis: Hybrid Systems*. – 2007. – Volume 1, Issue 2. – P. 212–222. DOI: 10.1016/j.nahs.2006.08.002
2. Drighiciu M. A. Hybrid Petri nets: A framework for hybrid systems modeling / M. A. Drighiciu // 2017 International Conference on Electromechanical and Power Systems (SIELMEN) – IEEE, 2017. – P. 020–025. DOI: 10.1109/SIELMEN.2017.8123285
3. Филимонов А. Б. О проблематике синтеза координирующих систем автоматического управления / А. Б. Филимонов, Н. Б. Филимонов // *Известия Южного федерального университета. Технические науки*. – 2012. – Т. 128, № 3. – С. 172–180.
4. Peterson J. L. Petri net theory and the modeling of systems / J. L. Peterson. – Prentice Hall PTR, 1981. – 290 p. ISBN 0–13–661983–5
5. Decomposition in automatic generation of Petri nets for manufacturing system control and scheduling / [D. W. He, B. Strege, H. Tolle, A. Kusiak] // *International Journal of Production Research* – 2000. – Volume 38, Issue 6. – P. 1437–1457. DOI: 10.1080/002075400188942
6. Performance assessment of industrial control system during pre-sales uncertain context using automatic Colored Petri Nets model generation / M. A. Ndiaye, J. F. Petin, J. Camerini, J. P. Georges // 2016 International Conference on Control, Decision and Information Technologies (CoDIT) – IEEE, 2016. – P. 671–676. DOI: 10.1109/CoDIT.2016.7593643
7. Durmuş M. S. Automatic generation of Petri Net supervisors for railway interlocking design / M. S. Durmuş, U. Yildirim, M. T. Söylemez // 2012 2nd Australian Control Conference. – IEEE, 2012. – P. 180–185. ISBN: 978-1-922107-63-3
8. Гурский А. А. Настройка нейронной сети при автоматическом синтезе сетей Петри / А. А. Гурский, С. М. Дубна // *Автоматизация технологических и бизнес-процессов*. – Одесса: 2018. – Т. 10, № 1. – С. 22–32.
9. Ahson, S. I. Petri net models of fuzzy neural networks / S. I. Ahson // *IEEE Transactions on systems, man, and cybernetics* – 1995. – Volume 38, Issue 6. – P. 926–932. ISSN: 2168-2909
10. Chow T. W. Higher-order Petri net models based on artificial neural networks / T. W. Chow, J. Y. Li // *Artificial Intelligence* – 1997. – Volume 92, Issues 1–2. – P. 289–300. DOI: [https://doi.org/10.1016/S0004-3702\(96\)00048-3](https://doi.org/10.1016/S0004-3702(96)00048-3)
11. Hanna, M. M. Fuzzy Petri nets with neural networks to model products quality from a CNC-milling machining centre / M. M. Hanna, A. Buck, R. Smith // *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* – 1996. – Volume 26, Issue 5. – P. 638–645. ISSN: 1558-2426. DOI: 10.1109/3468.531910
12. Бойчук Л. М. Синтез координирующих систем автоматического управления / Л. М. Бойчук. – М. : «Энергоатомиздат», 1991. – 160 с.
13. Gurskiy, A. A. Algorithms for tuning of the coordinating automatic control systems / A. A. Gurskiy, A. E. Goncharenko, S. M. Dubna // *Radio electronics, computer science, control*. – 2020. – No. 1. – P. 190–199. DOI: <https://doi.org/10.15588/1607-3274-2020-1-19>
14. Sutton R. S. Reinforcement learning: An introduction / R. S. Sutton, A. G. Barto. – MIT press, 2018. – 322 p. ISBN 0–262–19398–1

СИНТЕЗ ТА ВИКОРИСТАННЯ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ З ЙМОВІРНІСНИМ КОДУВАННЯМ СТРУКТУРИ

Леощенко С. Д. – аспірант кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя Україна.

Олійник А. О. – канд. техн. наук, доцент кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна.

Субботін С. О. – д-р техн. наук, професор, завідувач кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна.

Гофман Є. О. – старший науковий співробітник науково-дослідної частини Національного університету «Запорізька політехніка», Запоріжжя Україна.

Ільяшенко М. Б. – канд. техн. наук, доцент кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка», Запоріжжя Україна.

АНОТАЦІЯ

Актуальність. Розглянуто задачу кодування інформації моделей на базі штучних нейронних мереж для подальшого пересилання та використання таких моделей. Об'єктом дослідження є процес кодування штучних нейронних мереж з використанням ймовірнісних структур даних.

Мета роботи полягає у розробці методу кодування нейронних мереж для зменшенні ресурсоемності процесу нейроevolюційного синтезу моделей.

Метод. Запропоновано метод кодування нейронних мереж на основі ймовірнісних структур даних. На початку метод використовує основні принципи підходу прямого кодування інформації про мережу та базуючись на секвенуванні кодує матрицю міжнейронних зв'язків у вигляді біополімерів. Потім використовуються ймовірнісні структури даних для більш компактного представлення початкової матриці. Для цього використовуються хеш-функції, початкова матриця проходить через процес хешування, що дозволяє значно знизити вимоги до ресурсів пам'яті. Метод дозволяє скоротити витрати пам'яті при пересилці штучних нейронних мереж, що значно розширює практичне використання таких моделей, запобігаючи різкому зменшенню точності їх роботи.

Результати. Розроблений метод реалізовано та досліджено при вирішенні задачі класифікації стану південнонімецьких кредиторів. Використання розробленого методу дозволило збільшити швидкість синтезу нейромоделі на 15–17,6%, в залежності від використовуваних обчислювальних ресурсів. Також метод дозволив скоротити долю пересилок інформації на 8%, що також свідчить про прискорення та більш раціональне використання ресурсів.

Висновки. Проведені експерименти підтвердили працездатність запропонованого математичного забезпечення і дозволяють рекомендувати його для використання на практиці при кодуванні моделей на основі штучних нейронних мереж для подальшого вирішення задач діагностування, прогнозування, оцінювання та розпізнавання образів. Перспективи подальших досліджень можуть полягати в попередній обробці даних для більш жорсткого контролю процесу кодування з метою мінімізації втрат якості роботи моделей на основі нейронних мереж.

КЛЮЧОВІ СЛОВА: нейроevolюція, кодування, ймовірнісні структури даних, нейронні мережі, генетичний алгоритм.

АБРЕВІАТУРИ

NEAT – Neuroevolution of augmenting topologies;
RAM – Random Access Memory;
ГНМ – глибокі нейронні мережі;
ДНК – дезоксирибонуклеїнова кислота;
ЙСД – ймовірнісні структури даних;
НЕС – нейроevolюційний синтез;
ПМГА – паралельний модифікований генетичний алгоритм;
РНК – рибонуклеїнова кислота;
РНМ – рекурентні нейронні мережі;
ШНМ – штучна нейронна мережа.

НОМЕНКЛАТУРА

$\varepsilon_{estimation}$ – помилка оцінки;
 δ – ймовірність для значення із Count-min sketch;
 $cardinality_{hashfunction}$ – потужність хеш-функції;
 $depth_{sketch}$ – глибина (висота) ескізу;

G – популяція або генерація штучних нейронних мереж;
 n – кількість вхідних ознак, що характеризують екземпляри вибірки;
 N_i – множина нейронів на вході мережі;
 N_{i_l} – нейрон на вході мережі;
 N_o – множина нейронів на виході мережі;
 N_{o_p} – нейрон на виході мережі;
 N_h – множина нейронів прихованого шару мережі;
 N_{h_r} – нейрон прихованого шару мережі;
 NN – штучна нейронна мережа;
 NN_{param} – мета-параметри (інформація) про мережу;
 NN_{struct} – структура мережі;
 l – кількість нейронів на вході мережі;

m – кількість залежних (категоріальних) ознак екземплярів вибірки;

p – кількість нейронів на виході мережі;

q – кількість міжнейронних зв'язків у мережі;

r – кількість нейронів у прихованому шарі мережі;

Sample – вибірка даних;

t – розмір популяції;

w – множина міжнейронних зв'язків;

w_q – міжнейронний зв'язок мережі;

$width_{sketch}$ – ширина ескізу;

x_n – незалежна ознака екземпляру вибірки;

X – множина незалежних змінних – ознак;

y_m – значення залежної змінної (ознака) екземпляру вибірки;

$У_{очік}$ – реальне значення залежної змінної;

$У_{моделі}$ – значення залежної змінної отримане завдяки моделі;

$У_{моделі'}$ – значення залежної змінної отримане завдяки моделі після кодування та пересилки;

Y – множина значень залежних змінних.

ВСТУП

У наш час для синтезу моделей на основі нейронних мереж широкого використання набули еволюційні методи [1–9]. Такий підхід не вимагає наявності експертних знань про структуру моделей, дозволяє працювати з недиференційованими функціями активації, забезпечує змогу виходу з локальних екстремумів, а також можливості розв'язання завдань як параметричного, так і структурного синтезу моделей на основі нейронних мереж [10–14]. Важливим етапом нейроеволюційного синтезу діагностичних та розпізнавальних моделей є кодування інформації, тобто подання інформації про структуру моделі та її параметри для можливості застосування еволюційного пошуку. Проблема кодування інформації про нейронні мережі виникає при послідовному та паралельному нейроеволюційному синтезі, а також при збереженні та передачі ШНМ для подальшого їх опрацювання (рис. 1).

Усі існуючі методи кодування ШНМ можна розділити на прямі методи кодування та непрямі методи кодування [15, 16].

Обидві групи методів кодування мають свої переваги та недоліки. Пряме кодування не враховує тісний зв'язок між генним складом і індивідуальною продуктивністю, а тільки доводить, що його кодувача конструкція може сприяти ефективній еволюції нейронної мережі. А непряме кодування вимагає розробки набору правил кодування і декодування для трансформації послідовностей генів і окремих фенотипів, тому необхідно більше розуміти генетичні та еволюційні механізми біології. Спосіб кодування генів організмів надзвичайно ефективний, і дуже коротка послідовність генів може керувати складними фенотипами індивідуумів. Багато сучасних досліджень непрямого кодування знаходяться на дослі-

дницькій стадії, але з безперервним розвитком досліджень механізмів біологічної еволюції дослідження нейроеволюційних методів, заснованих на обох напрямках кодування, все ще мають великий потенціал.

Крім того, ряд труднощів із вибором типу кодування пов'язані із сучасними топологіями ШНМ [15–21]. Так, класичні методи прямого кодування майже неможливо застосовувати для РНМ [21], де окрім прямих зв'язків між нейронами є ще і зворотні. При кодуванні ГНМ [10] виникають проблеми як при використанні прямих методів, так і при застосуванні непрямих методів кодування, оскільки в такому випадку необхідно не тільки закодувати приховані нейрони, а і зберегти структуру нейромоделі, забезпечуючи розподіл за шарами.

Саме тому актуальною є науково-прикладна задача розробки нових методів кодування нейромоделей, які дозволяють кодувати інформацію про різні топології ШНМ не витрачаючи при цьому великої кількості комп'ютерних ресурсів.

Об'єкт дослідження – процес кодування ШНМ з використанням ЙСД.

Існуючі методи кодування ШНМ цілком або ж окремих мета-параметрів мережі характеризуються значною ресурсоемністю та є достатньо складними для імплементації еволюційних операторів. Крім того, при використанні існуючих [15, 16] часто виникає проблема із кодуванням інформації про РНМ або ГНМ.

Предмет дослідження – метод кодування ШНМ з використанням ЙСД.

На сьогоднішній день існує два основні підходи до кодування ШНМ. Проте більшість таких методів відрізняються такими недоліками: ресурсоемність та складність інтерпретування. Тому, у роботі запропоновано підхід, що базується на поданні інформації про ШНМ та кодуванні її з використанням ЙСД.

Мета роботи – розробка методу кодування нейронних мереж для зменшенні ресурсоемності процесу нейроеволюційного синтезу моделей.

1 ПОСТАНОВКА ПРОБЛЕМИ

Нехай задана вибірка вхідних даних $Sample = \langle X, Y \rangle$, де $X = \{x_1, x_2, \dots, x_n\}$ – множина незалежних змінних – ознак, $Y = \{y_1, y_2, \dots, y_m\}$ – множина значень залежних змінних, n та m – кількість вхідних ознак, що характеризують екземпляри вибірки.

Спираючись на це представимо проблему наступним чином: для синтезу ШНМ (NN) необхідно визначити множину нейронів $N = \{N_i, N_o, N_h\}$, що складається із підмножин вхідних $N_i = \{N_{i_1}, N_{i_2}, \dots, N_{i_l}\}, l = 1, 2, \dots, |N_i|$, вихідних $N_o = \{N_{o_1}, N_{o_2}, \dots, N_{o_p}\}, p = 1, 2, \dots, |N_o|$, прихованих нейронів $N_h = \{N_{h_1}, N_{h_2}, \dots, N_{h_r}\}, r = 1, 2, \dots, |N_h|$ та множини ваг зв'язків між нейронами $w = \{w_q\}$. Визначи-

вши значення елементів множин, можна вважати синтез ШНМ – завершеним.

Для цього будемо використовувати НЕС на базі ПМГА. В цьому випадку будемо оперувати популяціями ШНМ ($G = \{NN_1, NN_2, NN_3, \dots, NN_t\}$). Тоді постає проблема, як представити структуру (NN_{struct}) та інформацію про ШНМ (NN_{param}), щоб переслати її з головного ядра на задіяні потоки $NN = \{NN_{struct}, NN_{param}\} = \{N, w, NN_{param}\}$.

2 ОГЛЯД ЛІТЕРАТУРИ

У класичній генетиці, зазвичай виділяється генотип і фенотип. Генотип – це генетичне представлення істоти, а фенотип – актуалізоване фізичне представлення істоти [21, 22]. Еволюційні алгоритми завжди сильно відображають біологію, нейроеволюція нічим не відрізняється в цьому плані [10–12].

Питання кодування походить від питання про те, як представляти ШНМ генетично під час роботи. Спосіб, яким кодують ШНМ, визначає шлях, за яким метод буде обробляти ключові еволюційні процеси: відбір, мутацію і кросовер (або рекомбінацію). Будь-яке кодування потрапляє в одну з двох категорій, пряме або непряме.

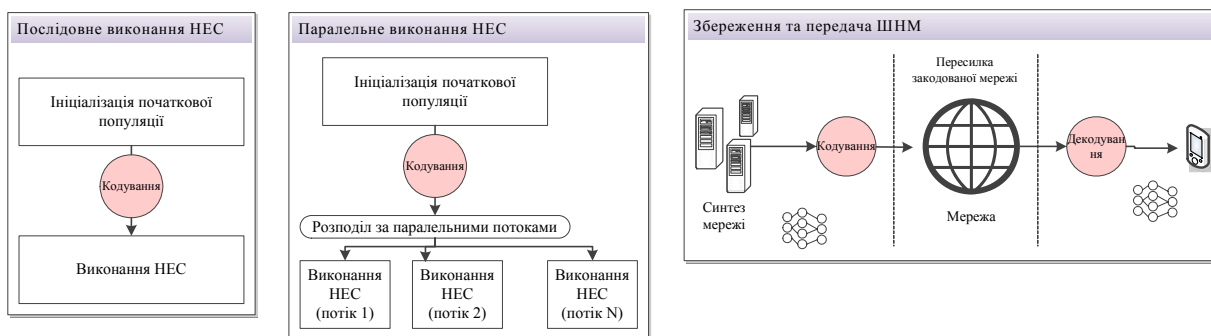


Рисунок 1 – Кейси використання кодування ШНМ

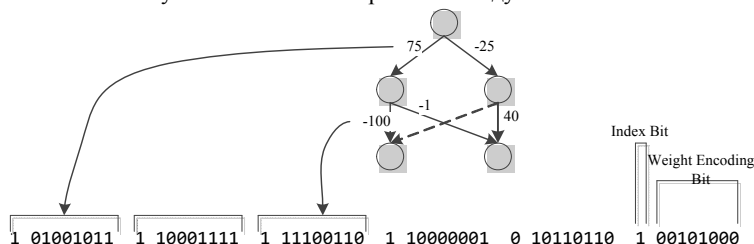


Рисунок 2 – Приклад прямого кодування ШНМ

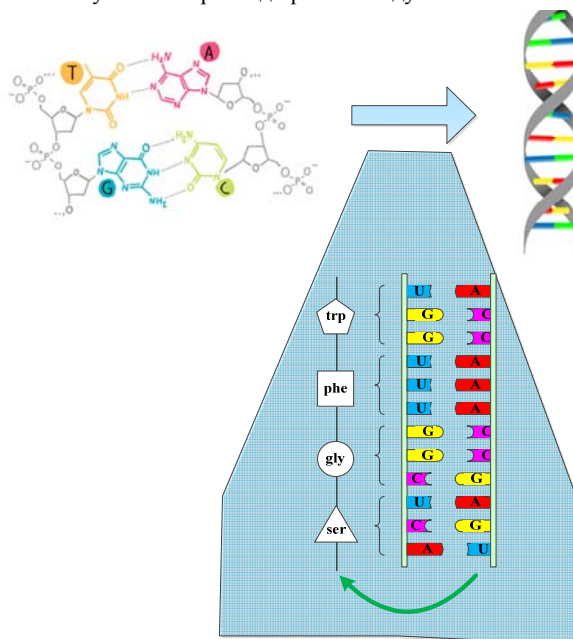


Рисунок 3 – Процес трансляції ДНК

Пряме кодування оперує хромосомами, що представляють деяке лінійне представлення ШНМ, в якому в явному вигляді вказані всі нейрони, ваги і зв'язку ШНМ [10, 12]. Таким чином, завжди можна побудувати взаємно-однозначну відповідність між структурними елементами ШНМ (нейронами, зв'язками, вагами та ін.), тобто фенотипом, і відповідними ділянками хромосоми, тобто генотипом.

Цей спосіб представлення нейронної мережі є найбільш простим і інтуїтивним, а також дозволяє застосовувати до отриманих хромосом вже наявний апарат генетичного пошуку (наприклад, оператори кросингверу і мутації). З найбільш очевидних мінухів такої схеми кодування можна відзначити «розпухання» генотипу при збільшенні кількості нейронів і зв'язків ШНМ і, як наслідок, низьку ефективність за рахунок значного збільшення простору пошуку [10].

В прикладі The Index Bit використовується для вказівки того, чи існує з'єднання чи ні. Weight Encoding Bits кодують ваги в двійковій формі. Ряд дослідників запропонували стратегію кодування, яка кодує ваги в дійсних числах, а також визначені деякі мутаційні оператори, які підходять для кодування.

Непряме кодування використовує більш «біологічний» принцип: в генотипі кодується не сам фенотип, але правила його побудови (умовно кажучи, якась програма) [13]. При декодуванні генотипу ці правила застосовуються в певній послідовності (найчастіше, рекурсивно і, найчастіше, застосовність правил залежить від поточного контексту), в результаті чого і будується нейронна мережа.

При використанні непрямих методів кодування генетичне представлення (а, відповідно, і простір пошуку для генетичних алгоритмів) виходить більш компактним, а сам генотип дозволяє кодувати модульні структури, що дає в певних умовах переваги в адаптивності отриманих результатів [14]. Натомість же отримуємо практичну неможливість простежити, які зміни в генотипі призвели до заданих змін у фенотипі, а також безліч труднощів з підбором генетичних операторів, збіжністю і продуктивністю.

Історично, пряме кодування було досліджено раніше і глибше, проте ряд недоліків цього підходу змушують дослідників все більш пильно придивлятися до непрямих методів кодування. Однак за своєю суттю непрямі методи досить складні для аналізу. Наприклад, одна і та ж мутація правила, розташованого на початку «програми», надає колосальний ефект, а застосована до «кінцевих» правил – ефекту не має зовсім, а в результаті – генетичний пошук має сильну тенденцію до передчасної збіжності. Підбір операторів кросингверу також є нетривіальним завданням, тому що використання стандартних бінарних операторів, як правило, призводить до частой появи нежиттєздатних рішень.

Існує також низка інших методик нейроеволюції. Розглянемо найбільш популярні з наведенням коротких описів кожної з них [10, 14]:

– Боєрс і Куйпер (Boers and Kuiper) – використання контексто-залежних І-систем;

– Ділаерт і Бір (Dellaert and Beer) – підхід, аналогічний Кангелоссі і Елману, але з використанням випадкових булевих нейромереж (random boolean networks);

– Харп, Самад і Гуха (Harp, Samad and Guha) – позонне пряме кодування структури;

– Груау (Gruau) – використання граматичного дерева для завдання інструкцій при діленні клітин (чимось схоже на Кангелоссі, Парісі і Нолфі);

– Вааріо (Vaario) – ріст клітин задається L-системами.

В результаті непряме кодування зазвичай більш компактне. З іншого боку, встановлення правил для непрямих кодування може призвести до сильного зміщення в просторі пошуку, тому набагато важче створити непряме кодування без суттєвих знань про те, як буде використовуватися кодування.

3 МАТЕРІАЛИ ТА МЕТОДИ

З теорії генетики [21–23] відомо, що секвенування біополімерів (білків і нуклеїнових кислот – ДНК і РНК) – це визначення їх амінокислотної або нуклеотидної послідовності [21–23]. В результаті секвенування отримують формальний опис первинної структури лінійної макромолекули у вигляді послідовності мономерів в текстовому вигляді [21, 22]. Розміри секвенуємих ділянок ДНК зазвичай не перевищують 100 пар нуклеотидів і 1000 пар нуклеотидів при секвенуванні по Сенгеру. В результаті секвенування ділянок ДНК, що перекриваються, отримують послідовності ділянок генів, цілих генів, тотальної мРНК або повних геномів організмів.

У цій роботі новий метод кодування інформації про ШНМ пропонується організувати базуючись на схожому принципі. Для кодування зв'язків у генотипі особини представимо інформацію про ваги міжнейронних зв'язків нейромоделі. При цьому кожен ген буде містити інформацію про індекси початкового та кінцевого нейрону зв'язку, а також його вагу. У випадку, коли метод працює з рекурентними нейронними мережами, додається додаткова комірка з вагою зворотного зв'язку, а його індекс визначається номером вихідного нейрону.

Введемо правила індексування нейронів:

1) оскільки кількість входів і виходів мережі є фіксованою величиною, то індекси відповідних нейронів є постійними і приймають значення в інтервалі $[0; N_i - 1]$ для вхідних нейронів, та $[N_i; N_i + N_o - 1]$ – для вихідних, де N_i і N_o – кількість входів і виходів мережі, відповідно. Видалення вхідних і вихідних нейронів є неможливим;

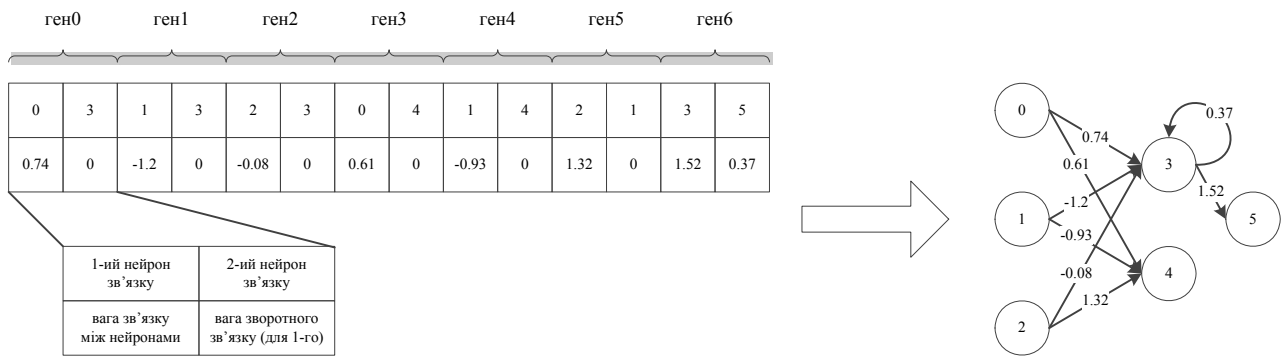


Рисунок 4 – Приклад кодування мережі

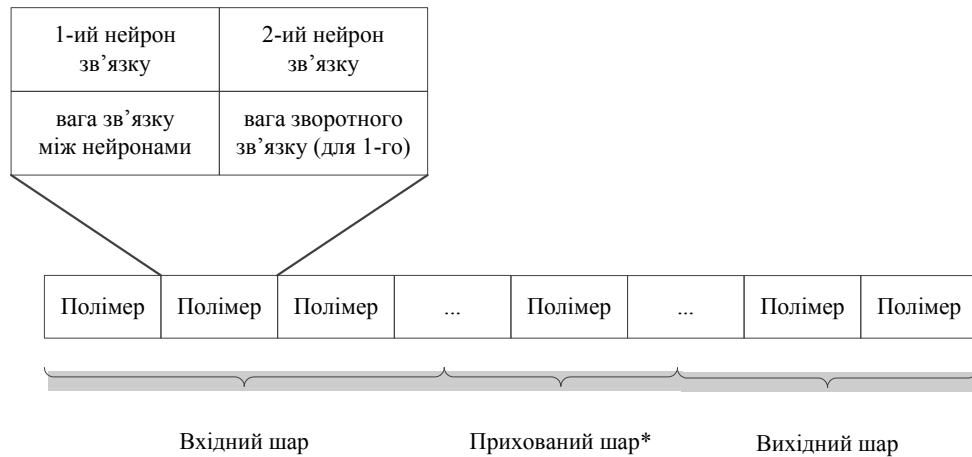


Рисунок 5 – Секвенування послідовностей за шарами ШНМ

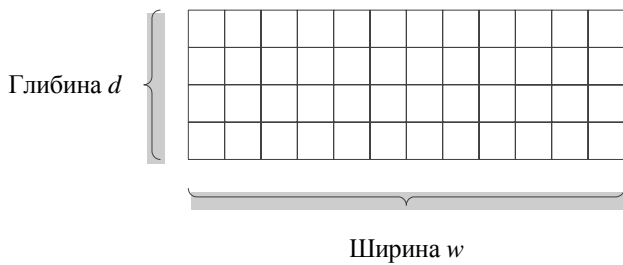


Рисунок 6 – Count-min sketch з глибиною d та шириною w

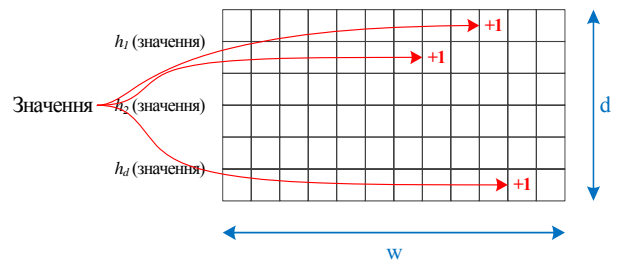


Рисунок 7 – Розподіл значень у структурі

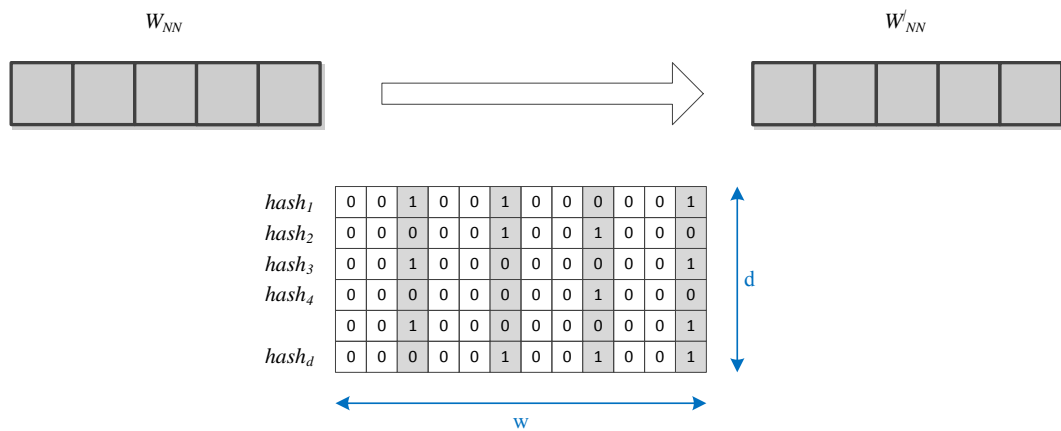


Рисунок 8 – Кодування та декодування інформації про ШНМ з використанням Count-min sketch

2) нові нейрони, що з'являються в результаті мутацій, отримують мінімальний можливий індекс. Наприклад, якщо особина представляє мережу з трьома входами, трьома виходами і не містить прихованих нейронів, то новому нейрону в цій мережі буде присвоєно індекс «5», наступному, що з'явився – «6» і т. д.;

3) індекси нейронів в мережі не можуть містити пропущені значення, тобто не може бути ШНМ з нейронами, що мають, наприклад, індекси N_0, N_1, N_2, N_5, N_6 . Якщо такий випадок виникає, наприклад, після видалення з мережі нейрона з індексом 4, індекси нейронів, що залишилися коригуються таким чином: $N_5 \rightarrow N_4$, $N_6 \rightarrow N_5$, при цьому змінюються дані у зв'язках, що належать до цих нейронів.

Тож отримаємо дворядковий список, де кожні чотири комірки (по дві з рядка) зберігають інформацію про нейрон (рис. 4).

Додатково слід відзначити, що друге правило забезпечує певне впорядкування шарів – секвенування послідовностей в списках (рис. 5).

На наступному етапі кодування використовуються ймовірнісні структури даних. Робота [24] пропонує використання модифікації фільтра Блума. Проте такий підхід не дозволяє в подальшому закодувати зворотні зв'язки, а використання такого підходу під час мутацій з видаленням певних нейронів вимагає запровадження додаткової матриці – другого фільтра Блума для підрахунку в зворотньому порядку [25]. Тож кількість обрахунків збільшується. З огляду на це більш доцільним є використання структури: Count–min sketch [26, 27]. Count–min sketch – це ймовірнісна структура даних, яка представляє собою таблицю частот подій в потоці даних. Вона використовує хеш-функції для зіставлення подій з частотами, але на відміну від хеш-таблиці використовує тільки сублінійний простір, за рахунок перерахунку деяких подій через колізій [25–27].

Структура Count–min sketch може представлятися у вигляді ескізу із глибиною d та шириною w (рис. 6).

Тоді представимо розподіл значень у масиві Count–min sketch, виходячи з принципу, зображеному на рисунку 7. Такий розподіл можна вважати ескізом.

Загалом матимемо представлення із розподілом за наступним правилом (1) та (2).

$$\varepsilon_{estimation} \leq 2 \cdot \max cardinalit_{y_{hashfunction}} / width_{sketch} \quad (1)$$

$$\delta = 1 - 0.5^{depth_{sketch}} \quad (2)$$

Варто відзначити, що ширина ескізу обмежує величину помилки (1), а висота (глибина) контролює ймовірність того, що оцінка проб'є цю межу (2).

Таким чином представимо масив даних про ШНМ у вигляді Count–min sketch та декодуємо його після цього (рис. 8).

Таким чином запропонований метод кодування розпочинається із початкового представлення матриці, що будується з так званих полімерів – комірок, що містять інформацію про вхідні-вихідні зв'язки між нейронами та ваги таких зв'язків, додатково для РНМ комірки будуть зберігати ваги зворотних зв'язків.

Далі визначаються хеш-функції (за замовченням рекомендується брати 4 хеш-функції) і створюється матриця для їх виведення, як показано на рис. 8.

Після цього для кожного елементу потоку даних про ШНМ обчислюються хеш-виходи і збільшується відповідний лічильник в матриці.

Таким чином, збільшуючи відповідні відліки в матриці, отримаємо оновлену матрицю.

У деяких випадках через хеш-колізію можлива ситуація, коли отримана частота елементу є дещо більшою, ніж очікувана. Точність кодування буде залежати від того, наскільки унікальними є хеш-функції, що повертають значення елементу. Крім того, чим більше хеш-функція, тим точнішою буде частота.

У цьому випадку ймовірнісна структура даних Count–min sketch дозволяє обчислити частоту потоків великих даних в суб-лінійному просторі, з часовою складністю, що оцінюється $O(1)$, тобто з постійною часовою складовою, що не залежить від параметрів нейромережевої моделі.

4 ЕКСПЕРИМЕНТИ

Виконано експериментальне дослідження розробленого методу шляхом його порівняння з відомими аналогами. Слід зазначити, що найбільш схожим та часто використовуваним методом кодування інформації при нейроеволюційному синтезі є метод NEAT [10–12]. Тож в подальших експериментальних дослідженнях будемо спиратися на нього.

На сьогоднішній день також є паралельні інтерпретації методу NEAT. Тому під час експериментального дослідження будемо аналізувати швидкість та долю пересилок під час паралельного виконання синтезу авторського методу [28–31] з новим видом кодування та NEAT.

Наступним етапом експериментального дослідження буде використання нового методу для кодування попередньо синтезованої мережі. Після кодування така мережа буде пересилатися за допомогою мережі Інтернет на іншу робочу станцію де буде протестована точність отриманої мережі.

Для першого етапу тестування буде використовуватися апаратне забезпечення кафедри програмних засобів Національного університету «Запорізька політехніка»: процесор Xeon E5-2660 v4 (14 ядер), оперативна пам'ять 4x16 ГБ DDR4.

У якості вхідних даних буде використовуватися вибірка даних про Південнонімецький кредитний стан (South German Credit Data Set) [32, 33]. Оновлений набір даних [34–36] містять корекцію і деяку довідкову інформацію, засновану на поданні відкритих даних Ludwig-Maximilians-Universität München (2010) тих же

даних і декількох інших німецькомовних ресурсів. Характеристики вибірки даних наведено у табл. 1.

Таблиця 1 – Загальна характеристика вибірки даних

Criterion	Characteristic
Data Set Characteristics:	Multivariate
Attribute Characteristics:	Integer, Real
Number of Instances:	1000
Number of Attributes	21

5 РЕЗУЛЬТАТИ

У таблиці 2 наведено результати порівняння виконання ПМГА із застосуванням компактного кодування та паралельної версії NEAT – Parallel NEAT.

Таблиця 2 – Результати експерименту

ПМГА					
Кількість ядер	1	2	4	8	16
Час синтезу, с	2361,870	1226,841	638,669	345,845	200,759
Доля пересилок	0,000	0,039	0,082	0,171	0,360
Час пересилок, с	0,000	47,691	52,136	59,288	72,273
Parallel NEAT					
Кількість ядер	1	2	4	8	16
Час синтезу, с	2716,510	1415,454	739,186	402,626	235,997
Доля пересилок	0,000	0,042	0,088	0,186	0,390
Час пересилок, с	0,000	59,608	65,370	74,773	92,039

Оскільки комп'ютерна система дозволяє змінювати кількість задіяних ядер – це бралось до уваги.

Зміни та різницю у часі пересилок при виконанні методів зображено у вигляді графіків, що представлені на рис. 9.

Експериментальні результати другого етапу наведено у табл. 3. Тут порівнюються результати синтезований ШНМ до та після пересилки з кодуванням. Засвідчимо, що швидкість передачі – не важлива.

Для повного відображення роботи із тестовою вибіркою використаємо нарізну діаграму, де один ряд значень це значення роботи моделі до пересилки, а дзеркальний ряд – значення після передачі та запуску мережі на іншій робочій станції. Нарізна діаграма зображена на рис. 10.

В такій діаграмі прогалини між дзеркальними стовбцями свідчать про певні неточності в роботі мережі, що була запущена після закодованої пересилки та декодування на новій робочій станції.

Таблиця 3 – Результати експерименту

№	Уочік	Умоделі	Умоделі'
1	63,40	62,16	62,16
2	60,60	59,36	59,36
3	63,84	62,56	62,56
4	64,25	63,02	63,02
5	58,48	57,28	57,28
6	61,79	60,53	60,53
7	63,72	62,46	62,46
8	61,97	60,77	60,77
9	58,25	57,02	57,02
...
100	59,63	59,03	59,03

6 ОБГОВОРЕННЯ

Експериментальна перевірка засвідчила, що ПМГА відпрацював швидше, хоча різниця у швидкості була не велика (табл. 2). Більш того, така різниця може бути пояснена саме меншою долею пересилок. Тобто під час роботи методу, менша кількість інформації пересилалася між потоками системи, а отже менше навантажувала пам'ять та систему загалом.

Проте, варто зазначити, що незважаючи ні на що, час пересилок зменшився, а отже інформація, що передавалася займала менше простору у пам'яті (рис. 9). Це додатково засвідчують графіки отримані на основі розподілу часу синтезу.

Тож перший етап експериментального дослідження довів результативність методу не в повній мірі.

Для другої частини експерименту буде використано попереднє апаратне забезпечення та додаткова робоча станція в наступній конфігурації: станція Acer Swift 3 під управлінням Intel Core i5-8250U із тактовою частотою 1,6–3,4 ГГц, RAM 16 ГБ (DDR4) та твердотільним накопичувачем на 128 ГБ. Ці дві системи будуть передавати закодовану ШНМ через мережу Інтернет. Засвідчимо, що швидкість передачі – не важлива.

Порівняємо помилку роботи ШНМ на тестовій частині вибірки. Як видно із табл. 3 у більшості випадків результати залишилися однаковими. Загалом же середня помилка збільшилася із 0,01811 до 0,1901. Додатково можемо у цьому пересвідчитися на нарізній діаграмі.

З нарізної діаграми (рис. 10) видно, що в деяких випадках модель отримана після декодування інформації після пересилки має певні неточності – це місця розривів стовбців на діаграмі. Проте таких випадків загалом достатньо мало. Таку ситуацію можливо отримати, коли після декодування були втрачені певні зв'язки.

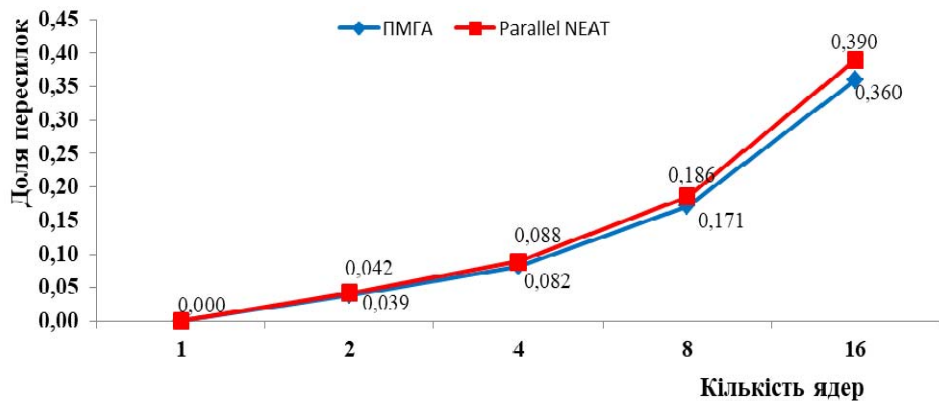


Рисунок 9 – Графіки розподілу пересилок під час паралельного синтезу ШНМ

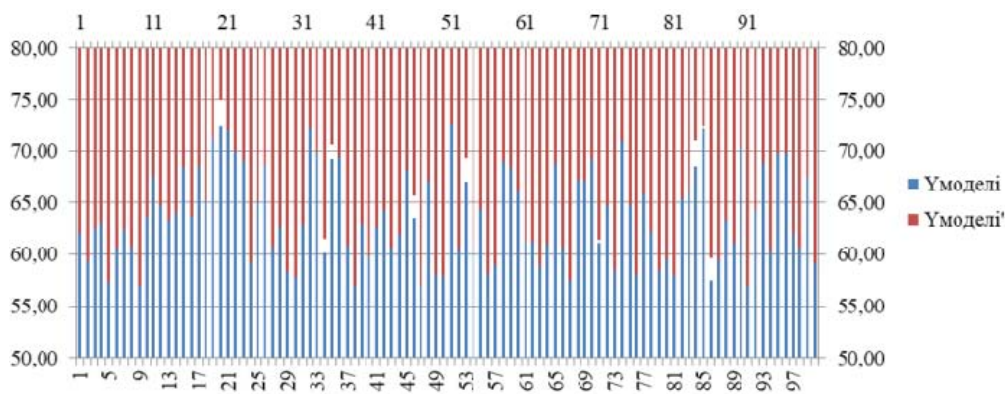


Рисунок 10 – Нарізна діаграма порівняння точності роботи моделей

Допустити такі незначні відхилення при втраті окремих нейронів в прихованому шарі – не можливо. Отже скоріше за все були втрачені саме зв'язки, які відповідали за певні взаємозв'язки у даних.

При таких результатах особливо цікаво було б відслідкувати зміни в роботі при дослідженні даних із певним періодом. До таких даних можливо віднести певну фінансову інформацію, коли наприклад, коливання курсу валют або вартості пакетів паперів має певний період, але все ж таки не повторюється повністю.

Аналізуючи загальні результати експериментів можна прийти до висновку, що запропонований метод компактного кодування не є досить ефективним для використання в паралельних системах, адже не самостійно не зменшує саме долю пересилок інформації, а лише частково пришвидшує цей процес. Більш того це можна пояснити виключно меншим об'ємом використаного ресурсу пам'яті.

Проте значно кращі та важливіші результати було отримано саме під час використання методу для кодування інформації про ШНМ для пересилки з однієї робочої станції на іншу. Це значно розширює саме практичні межі використання ШНМ. Адже таким чином можна виконати високо ітеративний процес синтезу на обчислювальному сервері, а результуючу мережу переслати для використання на конкретну робочу станцію. Беручи до уваги, що такою робочою

станцією може бути, наприклад, мобільний пристрій такий підхід забезпечує високу швидкість передачі ШНМ та подальшого декодування та практичного імплементування такої моделі.

Подальшу роботу можна вести у напрямі саме жорсткого контролю процесу синтезу та початкового компактного кодування, щоби подальші втрати в точності були контрольованими та не критичними.

ВИСНОВКИ

Вирішено актуальну науково-прикладну проблему компактного кодування нейромоделей в процесі нейроволюційного синтезу.

Наукова новизна полягає в тому що розроблено метод кодування нейронних мереж на основі ймовірнісних структур даних, який використовує принципи підходу прямого кодування. У запропонованому методі використовуються ЙСД для компактного зберігання інформації про ШНМ. Зберігання інформації про міжнейронні зв'язки підвищує логічну прозорість методу та робить можливим кодування інформації про РНМ та ГНМ. Використання ЙСД Count-min sketch дозволяє більш стисло та компактно кодувати матрицю початкових даних про ШНМ та прискорити процес передачі цих даних.

Практична цінність полягає в тому, що розв'язано практичні завдання кодування ШНМ, які в подальшому можуть використовуватися для діагнос-

тування, прогнозування, оцінювання та розпізнавання образів. Результати експериментів показали, що запропоновані методи кодування дозволяють більш компактно кодувати інформацію про ШНМ для її подальшої передачі на робочі станції для використання у якості моделі для діагностування, прогнозування, оцінювання та моделювання.

Майбутні напрями дослідження та розробки полягають у розробці більш жорстких процесів контролю синтезу та початкового кодування, щоби подальші втрати в точності були контрольованими та не критичними. Додатковим напрямом може стати попередня обробка вхідних даних для отримання більш явних зв'язків між даними та контролю відповідних міжнейронних зв'язків.

ПОДЯКА

Робота була проведена за підтримки держбюджетного науково-дослідного проекту державного бюджету Національного університету «Запорізька політехніка» «Інтелектуальні методи та програмне забезпечення для діагностики та неруйнівного контролю якості військової та цивільної техніки» (номер державної реєстрації 0119U100360).

ЛІТЕРАТУРА / LITERATURA

1. Kanaan M. T-Minus AI: Humanity's Countdown to Artificial Intelligence and the New Pursuit of Global Power / M. Kanaan. – Dallas : BenBella Books, 2019. – 270 p.
2. Taulli T. Artificial Intelligence Basics: A Non-Technical Introduction / T. Taulli. – New York : Apress, 2019. – 199 p.
3. Rothman D. Artificial Intelligence By Example: Acquire advanced AI, machine learning, and deep learning design skills / D. Rothman. – Birmingham : Packt Publishing, 2020. – 578 p.
4. Chang A.C. Intelligence-Based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare / A.C. Chang. – Cambridge : Academic Press, 2020. – 534 p.
5. Ponteves de P. AI Crash Course: A fun and hands-on introduction to machine learning, reinforcement learning, deep learning, and artificial intelligence with Python / P. de Ponteves. – Birmingham : Packt Publishing, 2019. – 360 p.
6. Artasanchez A. Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3.x / A. Artasanchez, P. Joshi. – Birmingham : Packt Publishing, 2020. – 618 p.
7. Additional training of neuro-fuzzy diagnostic models / [A. Oliinyk, S. Subbotin, S. Leoshchenko et al.] // Radio Electronics, Computer Science, Control. – 2018. – № 3. – P. 113–119. DOI: 10.15588/1607-3274-2018-3-12.
8. Using Modern Architectures of Recurrent Neural Networks for Technical Diagnosis of Complex Systems / [S. Leoshchenko, A. Oliinyk, S. Subbotin, T. Zaiko] // 2018 International Scientific-Practical Conference Problems of Informatics. Science and Technology (PIC S&T), Kharkiv, 9–12 October 2018 : proceedings. – Kharkiv : IEEE, 2018. – P. 411–416. DOI: 10.1109/INFOCOMMST.2018.8632015
9. Method of Artificial Neural Network Synthesis for Using in Integrated CAD / [S. Leoshchenko, A. Oliinyk, S. Subbotin et al.] // 2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Polyana, 26 February – 2 March 2019 : proceedings. – Lviv : IEEE, 2019. – P. 1–6. DOI: 10.1109/CADSM.2019.8779248
10. Iba H. Evolutionary Approach to Machine Learning and Deep Neural Networks: Neuro-Evolution and Gene Regulatory Networks / H. Iba. – New York : Springer, 2018. – 258 p.
11. Omelianenko I. Hands-On Neuroevolution with Python: Build high-performing artificial neural network architectures using neuroevolution-based algorithms / I. Omelianenko. – Birmingham : Packt Publishing, 2019. – 368 p.
12. Bergel A. Agile Artificial Intelligence in Pharo: Implementing Neural Networks, Genetic Algorithms, and Neuroevolution / A. Bergel. – New York : Apress, 2020. – 407 p.
13. Blokdyk G. Neuroevolution of augmenting topologies: Second Edition / G. Blokdyk. – Ohio : 5STARCOoks, 2018. – 128 p.
14. Lockett A.J. General-Purpose Optimization Through Information Maximization (Natural Computing Series) / A. J. Lockett. – New York : Springer, 2020. – 579 p.
15. Rouhiainen L. Artificial Intelligence: 101 Things You Must Know Today About Our Future / L. Rouhiainen. – Scotts Valley : CreateSpace Independent Publishing Platform, 2018. – 300 p.
16. Koul A. Practical Deep Learning for Cloud, Mobile, and Edge: Real-World AI & Computer-Vision Projects Using Python, Keras & TensorFlow / A. Koul, S. Ganju, M. Kasam. – Newton : O'Reilly Media, 2019. – 620 p.
17. Singh A. Mobile Deep Learning with TensorFlow Lite, ML Kit and Flutter: Build scalable real-world projects to implement end-to-end neural networks on Android and iOS / A. Singh, R. Bhadani. – Birmingham : Packt Publishing, 2020. – 380 p.
18. Davies J. The Internet of Things: From Data to Insight / J. Davies, C. Fortuna. – Hoboken : Wiley, 2020. – 240 p.
19. Zheng N. Learning in Energy-Efficient Neuromorphic Computing: Algorithm and Architecture Co-Design / N. Zheng, P. Mazumder. – New York : Wiley-IEEE Press, 2019. – 296 p.
20. Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis (Springer-Briefs in Computer Science) / [F.M. Bianchi, E. Maiorino, M.C. Kampffmeyer, A. Rizzi, R. Jenssen]. – New York : Springer, 2017. – 81 p.
21. Ozkan L. RNA Sequencing: Principles and Data Analysis / L. Ozkan. – Traverse City : Independently published, 2020. – 118 p.
22. Robinson T.R. Genetics For Dummies / T.R. Robinson, L. Spock. – New York : For Dummies, 2020. – 400 p.
23. Tan T.W. Beginners Guide to Bioinformatics for High Throughput Sequencing / T.W. Tan, E. Lee. – Singapore : World Scientific Publishing Co Pte Ltd, 2018. – 300 p.
24. Weightless: Lossy Weight Encoding For Deep Neural Network Compression / [B. Reagen, U. Gupta, R. Adolf, M. Mitzenmacher, A. Rush et al.] // International Conference on Machine Learning (ICML 2018), Stockholm, 10–15 July : proceedings. – Stockholm, PMLR, 2018. – P. 1–10.
25. Gakhov A. Probabilistic Data Structures and Algorithms for Big Data Applications / A. Gakhov. – Madison : Books on Demand, 2019. – 220 p.

26. Knebl H. Algorithms and Data Structures: Foundations and Probabilistic Methods for Design and Analysis / H. Knebl. – New York : Springer, 2020. – 360 p.
27. Introduction to Algorithms / [T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein]. – Cambridge : The MIT Press, 2009. – 1292 p.
28. Modification and parallelization of genetic algorithm for synthesis of artificial neural networks / [S. D. Leoshchenko, A. O. Oliinyk, S. A. Subbotin et al.] // Radio Electronics, Computer Science, Control. – 2019. – № 4. – P. 68–82. DOI: 10.15588/1607-3274-2018-3-12.
29. Synthesis of artificial neural networks using a modified genetic algorithm / [S. Leoshchenko, A. Oliinyk, S. Subbotin, N. Gorobii et al.] // 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018), Lviv, 28–30 October, 2018 : proceedings. – Lviv : CEUR WS, 2018. – P. 1–13.
30. Implementation of Selective Pressure Mechanism to Optimize Memory Consumption in the Synthesis of Neuromodels for Medical Diagnostics / [S. Leoshchenko, A. Oliinyk, S. Subbotin et al.] // 2nd International Workshop on Informatics and Data-Driven Medicine (IDDM 2019), Lviv, 11–13 November, 2019 : proceedings. – Lviv : CEUR WS, 2019. – P. 109–120.
31. Leoshchenko S. Adaptive Mechanisms for Parallelization of the Genetic Method of Neural Network Synthesis / S. Leoshchenko, A. Oliinyk, S. Subbotin // 10th International Conference on Advanced Computer Information Technologies (ACIT 2020), Deggendorf, 16–18 November : proceedings. – Ternopil: IEEE, 2020. – P. 446–450. DOI: 10.1109/ACIT49673.2020.9208905.
32. South German Credit Data Set [Electronic resource]. – Access mode: <https://archive.ics.uci.edu/ml/datasets/South+German+Credit>
33. Grömping U. South German Credit Data: Correcting a Widely Used Data Set / U. Grömping // Reports in Mathematics, Physics and Chemistry, Department II. – Berlin : Beuth University of Applied Sciences Berlin, 2019. – 14 p.
34. Feature Selection Based on Parallel Stochastic Computing / [A. Oliinyk, S. Subbotin, V. Lovkin et al.] // 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT'2018), Lviv, 11–14 September 2018 : proceedings. – Lviv : IEEE, 2018. – P. 347–351. DOI: 10.1109/STC-CSIT.2018.8526729
35. Development of the indicator set of the features informativeness estimation for recognition and diagnostic model synthesis / [A. Oliinyk, S. Subbotin, V. Lovkin et al.] // 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET 2018), Slavsko, 20–24 February 2018 : proceedings. – Lviv : IEEE, 2018. – P. 903–908. DOI: 10.1109/TCSET.2018.8336342.
36. Methods of semantic proximity extraction between the lexical units in infocommunication systems / [S. Leoshchenko, A. Oliinyk, S. Subbotin, T. Zaiko] // International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, 10–13 October 2017 : proceedings. – Kharkiv : IEEE, 2017. – P. 7–12. DOI: 10.1109/INFOCOMMST.2017.8246137.

Стаття надійшла до редакції 14.04.2021.
Після доробки 29.04.2021.

УДК 004.896

СИНТЕЗ И ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ С ВЕРОЯТНОСТНЫМ КОДИРОВАНИЕМ СТРУКТУРЫ

Леошенко С. Д. – аспирант кафедры программных средств Национального университета «Запорожская политехника», Запорожье Украина.

Олейник А. А. – канд. техн. наук, доцент, доцент кафедры программных средств Национального университета «Запорожская Политехника», Запорожье Украина.

Субботин С. А. – д-р техн. наук, профессор, заведующий кафедрой программных средств Национального университета «Запорожская политехника», Запорожье Украина.

Гофман Е. А. – старший научный сотрудник научно-исследовательской части Национального университета «Запорожская политехника», Запорожье Украина.

Ильяшенко М. Б. – канд. техн. наук, доцент, доцент кафедры компьютерных систем и сетей Национального университета «Запорожская политехника», Запорожье Украина.

АННОТАЦИЯ

Актуальность. Рассмотрена задача кодирования информации моделей на базе искусственных нейронных сетей для дальнейшей пересылки и использования таких моделей. Объектом исследования является процесс кодирования искусственных нейронных сетей с использованием вероятностных структур данных.

Цель работы заключается в разработке метода кодирования нейронных сетей для уменьшения ресурсоемкости процесса нейроэволюционного синтеза моделей.

Метод. Предложен метод кодирования нейронных сетей на основе вероятностных структур данных. В начале метод использует основные принципы подхода прямого кодирования информации о сети и, базируясь на секвенировании, кодирует матрицу межнейронных связей в виде биополимеров. Затем используются вероятностные структуры данных для более компактного представления исходной матрицы. Для этого используются хэш-функции, начальная матрица проходит через процесс хэширования, что позволяет значительно снизить требования к ресурсам памяти. Метод позволяет сократить затраты памяти при пересылке искусственных нейронных сетей, что значительно расширяет практическое использование таких моделей, предотвращая резкое уменьшение точности их работы.

Результаты. Разработанный метод реализован и исследован при решении задачи классификации состояния южногерманских кредиторов. Использование разработанного метода позволило увеличить скорость синтеза нейромоделей на 15–

17,6%, в зависимости от используемых вычислительных ресурсов. Также метод позволил сократить долю пересылок информации на 8%, что также свидетельствует об ускорении и более рациональном использовании ресурсов.

Выводы. Проведенные эксперименты подтвердили работоспособность предложенного математического обеспечения и позволяют рекомендовать его для использования на практике, при кодировании моделей на основе искусственных нейронных сетей, для дальнейшего решения задач диагностирования, прогнозирования, оценивания и распознавания образов. Перспективы дальнейших исследований могут состоять в предварительной обработке данных для более жесткого контроля процесса кодирования с целью минимизации потерь качества работы моделей, на основе нейронных сетей.

КЛЮЧЕВЫЕ СЛОВА: нейроэволюция, кодирование, вероятностные структуры данных, нейронные сети, генетический алгоритм.

UDC 004.896

SYNTHESIS AND USAGE OF NEURAL NETWORK MODELS WITH PROBABILISTIC STRUCTURE CODING

Leoshchenko S. D. – PhD student of the Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

Oliinyk A. O. – PhD., Associate Professor, Associate Professor of the Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

Subbotin S. A. – Dr. Sc., Professor, Head of the Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

Gofman Ye. O. – PhD, Senior Researcher of the Research Unit, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine

Ilyashenko M. B. – PhD, Associate Professor, Associate Professor of the Department of Computer Systems and networks, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

ABSTRACT

Context. The problem of encoding information of models based on artificial neural networks for further transmission and use of such models is considered. The object of research is the process of coding artificial neural networks using probabilistic data structures.

Objective of this work is to develop a method for coding neural networks to reduce the resource intensity of the process of neuroevolutionary model synthesis.

Method. A method for encoding neural networks based on probabilistic data structures is proposed. At the beginning, the method uses the basic principles of the approach of direct encoding of network information and, based on sequencing, encodes a matrix of interneuronal connections in the form of biopolymers. Then, probabilistic data structures are used to represent the original matrix more compactly. For this purpose, hash functions are used, the initial matrix goes through the hashing process, which significantly reduces the requirements for memory resources. The method allows to reduce memory costs when sending artificial neural networks, which significantly expands the practical use of such models, preventing a sharp decrease in the accuracy of their operation.

Results. The developed method is implemented and investigated in solving the problem of classification of the state of South German creditors. The use of the developed method allowed increasing the rate of neuromodel synthesis by 15–17.6%, depending on the computing resources used. The method also reduced the share of information transfers by 8%, which also indicates faster and more efficient use of resources.

Conclusions. The conducted experiments confirmed the efficiency of the proposed mathematical software and allow us to recommend it for use in practice, when encoding models based on artificial neural networks, for further solving problems of diagnostics, forecasting, evaluation and pattern recognition. Prospects for further research may consist in pre-processing data for more strict control of the encoding process in order to minimize the loss of quality of models based on neural networks.

KEYWORDS: neuroevolution, coding, probabilistic data structures, neural networks, genetic algorithm.

REFERENCES

1. Kanaan M. T-Minus AI: Humanity's Countdown to Artificial Intelligence and the New Pursuit of Global Power. Dallas, BenBella Books, 2019, 270 p.
2. Taulli T. Artificial Intelligence Basics: A Non-Technical Introduction. New York, Apress, 2019, 199 p.
3. Rothman D. Artificial Intelligence By Example: Acquire advanced AI, machine learning, and deep learning design skills. Birmingham, Packt Publishing, 2020, 578 p.
4. Chang A.C. Intelligence-Based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare. Cambridge, Academic Press, 2020, 534 p.
5. Ponteves de P. AI Crash Course: A fun and hands-on introduction to machine learning, reinforcement learning, deep learning, and artificial intelligence with Python. Birmingham, Packt Publishing, 2019, 360 p.
6. Artasanchez A., Joshi P. Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3.x. Birmingham, Packt Publishing, 2020, 618 p.
7. Oliinyk A., Subbotin S., Leoshchenko S., Ilyashenko M., Myronova N., Mastinovskiy Y. Additional training of neuro-fuzzy diagnostic models, *Radio Electronics, Computer Science, Control*, 2018, № 3, pp. 113–119. DOI: 10.15588/1607-3274-2018-3-12.
8. Leoshchenko S., Oliinyk A., Subbotin S., Zaiko T. Using Modern Architectures of Recurrent Neural Networks for Technical Diagnosis of Complex Systems, *2018 International Scientific-Practical Conference Problems of Informatics. Science and Technology (PIC S&T), Kharkiv, 9–12 October 2018, proceedings*. Kharkiv, IEEE, 2018, pp. 411–416. DOI: 10.1109/INFOCOMMST.2018.8632015
9. Leoshchenko S., Oliinyk A., Subbotin S., Shylo S., Shkarupko V. Method of Artificial Neural Network Synthesis for

- Using in Integrated CAD, *15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Polyana, 26 February – 2 March 2019, proceedings*. Lviv, IEEE 2019, pp. 1–6. DOI: 10.1109/CADSM.2019.8779248
10. Iba H. *Evolutionary Approach to Machine Learning and Deep Neural Networks: Neuro-Evolution and Gene Regulatory Networks*, New York, Springer, 2018, 258 p.
 11. Omelianenko I. *Hands-On Neuroevolution with Python: Build high-performing artificial neural network architectures using neuroevolution-based algorithms*, Birmingham, Packt Publishing, 2019, 368 p.
 12. Bergel A. *Agile Artificial Intelligence in Pharo: Implementing Neural Networks, Genetic Algorithms, and Neuroevolution*, New York, Apress, 2020, 407 p.
 13. Blokdyk G. *Neuroevolution of augmenting topologies: Second Edition*, Ohio, 5STARCOoks, 2018, 128 p.
 14. Lockett A.J. *General-Purpose Optimization Through Information Maximization (Natural Computing Series)*. New York, Springer, 2020, 579 p.
 15. Rouhiainen L. *Artificial Intelligence: 101 Things You Must Know Today About Our Future*, Scotts Valley, CreateSpace Independent Publishing Platform, 2018, 300 p.
 16. Koul A., Ganju S., Kasam M. *Practical Deep Learning for Cloud, Mobile, and Edge: Real-World AI & Computer-Vision Projects Using Python, Keras & TensorFlow*, Newton, O'Reilly Media, 2019, 620 p.
 17. Singh A., Bhadani R. *Mobile Deep Learning with TensorFlow Lite, ML Kit and Flutter: Build scalable real-world projects to implement end-to-end neural networks on Android and iOS*, Birmingham, Packt Publishing, 2020, 380 p.
 18. Davies J., Fortuna C. *The Internet of Things: From Data to Insight*, Hoboken, Wiley, 2020, 240 p.
 19. Zheng N., Mazumder P. *Learning in Energy-Efficient Neuromorphic Computing: Algorithm and Architecture Co-Design*, New York, Wiley-IEEE Press, 2019, 296 p.
 20. Bianchi F.M., Maiorino E., Kampffmeyer M.C., et al. *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis (SpringerBriefs in Computer Science)*, New York, Springer, 2017, 81 p.
 21. Ozkan L. *RNA Sequencing: Principles and Data Analysis*, Traverse City, Independently published, 2020, 118 p.
 22. Robinson T.R., Spock L. *Genetics For Dummies*, New York, For Dummies, 2020, 400 p.
 23. Tan T.W., Lee E. *Beginners Guide to Bioinformatics for High Throughput Sequencing*, Singapore, World Scientific Publishing Co Pte Ltd, 2018, 300 p.
 24. Reagen B., Gupta U., Adolf R., Mitzenmacher M., et al. *Weightless: Lossy Weight Encoding For Deep Neural Network Compression, International Conference on Machine Learning (ICML 2018), Stockholm, 10–15 July, proceedings*. Stockholm, PMLR, 2018, pp. 1–10.
 25. Gakhov A. *Probabilistic Data Structures and Algorithms for Big Data Applications*, Madison : Books on Demand, 2019, 220 p.
 26. Knebl H. *Algorithms and Data Structures: Foundations and Probabilistic Methods for Design and Analysis*, New York, Springer, 2020, 360 p.
 27. Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. *Introduction to Algorithms*, Cambridge, The MIT Press, 2009, 1292 p.
 28. Leoshchenko S. D., Oliinyk A. O., Subbotin S. A., Lytvyn V. A., Shkaruplyo V.V. *Modification and parallelization of genetic algorithm for synthesis of artificial neural networks, Radio Electronics, Computer Science, Control, 2019, № 4, pp. 68–82. DOI: 10.15588/1607-3274-2018-3-12.*
 29. Leoshchenko S., Oliinyk A., Subbotin S., Gorobii N., Zaiko T. *Synthesis of artificial neural networks using a modified genetic algorithm, 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018), Lviv, 28–30 October, 2018 : proceedings*. Lviv, CEUR WS, 2018, pp. 1–13.
 30. Leoshchenko S., Oliinyk A., Subbotin S., Gorobii N., Zaiko T. *Implementation of Selective Pressure Mechanism to Optimize Memory Consumption in the Synthesis of Neuro-models for Medical Diagnostics, 2nd International Workshop on Informatics and Data-Driven Medicine (IDDM 2019), Lviv, 11–13 November, 2019. proceedings*. Lviv, CEUR WS, 2019, pp. 109–120.
 31. Leoshchenko S., Oliinyk A., Subbotin S. *Adaptive Mechanisms for Parallelization of the Genetic Method of Neural Network Synthesis, 10th International Conference on Advanced Computer Information Technologies (ACIT 2020), Deggendorf, 16–18 November, proceedings*. Ternopil, IEEE, 2020, pp. 446–450, DOI: 10.1109/ACIT49673.2020.9208905.
 32. South German Credit Data Set [Electronic resource]. Access mode: <https://archive.ics.uci.edu/ml/datasets/South+German+Credit>
 33. Grömping U. *South German Credit Data: Correcting a Widely Used Data Set, Reports in Mathematics, Physics and Chemistry*, Department II, Berlin, Beuth University of Applied Sciences Berlin, 2019, 14 p.
 34. Oliinyk A., Subbotin S., Lovkin V., Leoshchenko S., Zaiko T. *Feature Selection Based on Parallel Stochastic Computing, 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT'2018), Lviv, 11–14 September 2018, proceedings*. Lviv, IEEE, 2018, P. 347–351. DOI: 10.1109/STC-CSIT.2018.8526729
 35. Oliinyk A., Subbotin S., Lovkin V., Leoshchenko S., Zaiko T. *Development of the indicator set of the features informativeness estimation for recognition and diagnostic model synthesis, 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET 2018) Slavsko, 20–24 February 2018, proceedings*. Lviv, IEEE, 2018, P. 903–908. DOI: 10.1109/TCSET.2018.8336342.
 36. Leoshchenko S., Oliinyk A., Subbotin S., Zaiko T. *Methods of semantic proximity extraction between the lexical units in infocommunication systems, 2017 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T) Kharkiv, 10–13 October 2017, proceedings*. Kharkiv, IEEE, 2017, pp. 7–12. DOI: 10.1109/INFOCOMMST.2017.8246137.

TREE-BASED SEMANTIC ANALYSIS METHOD FOR NATURAL LANGUAGE PHRASE TO FORMAL QUERY CONVERSION

Litvin A. A. – Postgraduate student of department of Microprocessor Technology, V.M. Glushkov Institute of Cybernetics, Kyiv, Ukraine.

Velychko V. Yu. – PhD, Senior researcher of department of Microprocessor Technology, V.M. Glushkov Institute of Cybernetics, Kyiv, Ukraine.

Kaverynskyi V. V. – PhD, Senior researcher of department of Abrasion- and Corrosion-Resistant Powder Construction Materials, I.M. Frantsevich Institute for Problems of Material Science, Kyiv, Ukraine.

ABSTRACT

Context. This work is devoted to the problem of natural language interface construction for ontological graph databases. The focus here is on the methods for the conversion of natural language phrases into formal queries in SPARQL and CYPHER query languages.

Objective. The goals of the work are the creation of a semantic analysis method for the input natural language phrases semantic type determination and obtaining meaningful entities from them for query template variables initialization, construction of flexible query templates for the types, development of program implementation of the proposed technique.

Method. A tree-based method was developed for semantic determination of a user's phrase type and obtaining a set of terms from it to put them into certain places of the most suitable formal query template. The proposed technique solves the tasks of the phrase type determination (and this is the criterion of the formal query template selection) and obtaining meaningful terms, which are to initialize variables of the chosen template. In the current work only interrogative and incentive user's phrases are considered i.e. ones that clearly propose the system to answer or to do something. It is assumed that the considered dialog or reference system uses a graph ontological database, which directly impacts the formal query patterns – the resulting queries are destined to be in SPARQL or Cypher query languages. The semantic analysis examples considered in this work are aimed primarily at inflective languages, especially, Ukrainian and Russian, but the basic principles could be suitable to most of the other languages.

Results. The developed method of natural language phrase to a formal query in SPARQL and CYPHER conversion has been implemented in software for Ukrainian and Norwegian languages using narrow subjected ontologies and tested against formal performance criteria.

Conclusions. The proposed method allows the dialog system fast and with minimum number of steps to select the most suitable query template and extract informative entities from a natural language phrase given the huge phrase variability in inflective languages. Carried out experiments have shown high precision and reliability of the constructed system and its potential for practical usage and further development.

KEYWORDS: natural language processing, graph data base, semantic analysis, formal query, decision tree, ontology.

ABBREVIATIONS

SPARQL is a query language to the data presented by the RDF model;

RDF is a Resource Description Framework;

OWL is a Web Ontology Language;

CYPHER is a query language for graph databases accepted in Neo4j;

LODQA is a Linked Open Data Question Answering.

NOMENCLATURE

A_C is an accuracy criterion;

A is a subject model on which the text is interpreting;

a_i is a possible word t_{ij} meaning;

C_L is a left context of the grammar unit;

C_R is a right context of the grammar unit;

D is a dictionary of a natural language in the alphabet X ;

F_I is a complex criterion to estimate precision and recall;

F is a formalization for the set of the words in the alphabet X ;

F_n is a number of false negatives;

F_p is a number of false positives;

L is a formalization for a natural language on the given alphabet;

P_r is a precision criterion;

P is a formalization for the aggregate of a natural language grammar rules;

p_i is one of the grammar rules of a natural language;

R is a recall criterion;

R_E is a formalization for the aggregate of a natural language grammar relationships;

R_{p_i} is a formalization for a binding of grammar rule and relationship;

S is a dictionary of words definitions for the language L ;

T is a currently considered natural language text;

T_p is a number of true positives;

T_n is a number of true negatives;

t_i is a sentence of a currently considered natural language text;

t_{ij} is a grammar unit of a sentence;

X is a formalization for an alphabet of some natural language;

γ is a relationship that defines meanings and types of words in a dictionary of the language;

Π is a predicates signature;

π^{kr} is an atomic predicate;
 τ_i is a word t_{ij} possible type;
 φ is a relationship of the text interpretation on the subject model.

INTRODUCTION

Nowadays dialog and reference program systems are becoming wide-spread and convenient. They help to automate frequent typical question answering, to obtain relevant information about an actual problem or to perform a necessary sequence of action. They can reduce the burden for human consultants in the routine of answering typical, many times asked and bored questions, and also being asynchronous a program system can serve several clients at the same time. It is obvious that for many cases the most comfortable and friendly for a user is a natural language interface of such system, which means that the user enters his question in a natural language. Hence, in the most of situations such dialog system is, actually, natural language interface of a database.

The problem of natural language database interfaces building is not new. Since computers have become wide-spread in the most aspects of life and work and appear common for large numbers of people even far from computer science and calculation tasks the problem of user's friendly interfaces creation become actual. And one of the directions of such interfaces is a natural language interface for databases that is highly desirable for persons who not familiar with programming and formal query languages. In contrast, asking the system using natural language seems rather more natural and convenient for an ordinary person. Despite the long existence of the problem, the development of such interfaces and new approaches to their construction continues [1], and there are certain reasons for this. Due to the complexity of the task of interpreting the semantics of queries in a natural language, the development process moved with varying success, experiencing ups and downs. The processing of an incoming request by a previously prepared semantic model may not be processed correctly for other requests. New attempts to create natural language interfaces for databases are constantly being made. Nevertheless, far from all of them turn out to be quite successful and indeed make a contribution to the state of the problem. However, every new research in this area is valuable because it brings new ideas and provides a better understanding of what really works and what does not.

The object of study is a natural language interface for ontological graph databases.

The subject of study is the development of methods for natural language of inflective type conversion to formal queries for graph databases.

The purpose of the work is to create a semantic analysis method for the input natural language phrases semantic type determination and obtaining meaningful entities from them for query template variables initialization, to const the flexible query templates for the needed

semantic types and to develop and test a program implementation of the proposed technique.

1 PROBLEM STATEMENT

The main purpose of this work is to develop a program system that is able to return as its output formal queries in SPARQL or Cypher query languages for graph databases using as input interrogative and imperative phrases in a natural language of inflective type. The method used for this task should be easily interpretable, be able to be easily corrected, supplemented and adapted, be fast and not resource intensive.

To solve this problem the following tasks are to be done:

1. To develop a tree-based method for semantic analysis of a natural language phrase;
2. To construct the scheme of the decision tree used in the developed system for determination of semantic type of the input natural language phrase;
3. To create flexible templates of formal queries for SPARQL and Cypher corresponding to the required semantic types;
4. To come out with the method for obtaining meaningful entities from the phrase to initialize templates variables;
5. To develop a program realization of the business logic separate from the tree and templates files;
6. To integrate the natural language to formal query conversion modulus as a service into a multi-agent dialog system;
7. To carry-out experiments for testing the developed system performance.

To test the quality of the developed system results the following criteria are to be used: accuracy, precision, recall and F_1 -score, which are common for such kinds of systems [2]. These metrics are calculated as follows:

$$A_C = \frac{T_p + T_n}{T_p + F_p + F_n + T_n}, \quad (1)$$

$$P_r = \frac{T_p}{T_p + F_p}, \quad (2)$$

$$R = \frac{T_p}{T_p + F_n}, \quad (3)$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R}. \quad (4)$$

Formal mathematic problem characteristic is given here according to [3]:

Let X is an alphabet of some natural language and $F(X)$ is the set of the words in X . $L \subseteq F(X)$ is a natural language on the given alphabet and its sentences are constructed according to the grammar rules $P = \{p_i : i = 1, \dots, m\}$. The grammar rules define the relationships $R_E = \{R_{p_i} : p_i \in P\}$ each from them is corresponding to a grammar rule. Let $T \in L(X)$ is a natural language L text

and $t_i \in T$ are sentences of the text T , $i = 1, 2, \dots, n$. Every sentence of the text T has its structure $t_i = t_{i1}t_{i2} \dots t_{im}$, where t_{ij} are grammar units that form a sentence. $C_L(t_{ij}) = t_{i1} \dots t_{ij-1}$ is the left context of the word t_{ij} and $C_R(t_{ij}) = t_{ij+1} \dots t_{im}$ is the right context of it in the sentence t_i . S is a dictionary of the language $L(X)$ where the definitions of words t_{ij} are. $\gamma \subseteq T \times S$ is a relationship that defines meanings and types of words in S . $A = (D, \Pi)$ is the subject model on which the text T is interpreting. $\varphi \subseteq T \times A$ is the relationship of the text T interpretation on the model $A = (D, \Pi)$. Predicates signature $\Pi = \{\pi^{kl}, \dots, \pi^{kr}\}$ includes atomic predicates to build more complicated formulas. Every atomic predicate has its type. Relationship γ is evaluated as:

$$\gamma(t_{ij}) = \{(a_1, \tau_1), (a_2, \tau_2), \dots, (a_s, \tau_s)\} \quad (5)$$

The relationship φ if the model $A = (D, \Pi)$ is defined can be determinate as following:

$$\begin{aligned} \varphi(t_i) = & \{ \varphi(\gamma(t_{i1})\gamma(C_R(t_{i1}))), \varphi(\gamma(C_R(t_{i2}))\gamma(t_{i2})\gamma(C_R(t_{i2}))), \dots, \\ & \varphi(\gamma(C_L(t_{in}))\gamma(t_{in})) \}, \\ & \varphi(\gamma(t_{ij})) = \gamma(\varphi(t_{ij})), \\ & \varphi(\gamma(C_L(t_{ij}))) = C_L(\gamma(\varphi(t_{ij}))), \\ & \varphi(\gamma(\pi_r^k(p_1, \dots, p_k))) = \\ & \gamma(\varphi(\pi_r^k))(\varphi(\gamma(p_1), \dots, \gamma(p_k))). \end{aligned} \quad (6)$$

It should be noticed, that here $\gamma(\varphi(\pi_r^k))$ is a predicate coordinated with arguments $\gamma(p_1), \dots, \gamma(p_k)$.

Thus, the main formal goal of a semantic analysis technique development is to construct a method for γ and φ implementation.

2 REVIEW OF THE LITERATURE

In [4] proposed a transformation system that includes three main components: the first component converts a natural language query into a query tree, the second component interactively checks the conversion, turning to the user, and the third component converts the query tree into SQL. The main disadvantage of this approach is a user's checking the constructed query tree, which makes such system not convenient for a user not familiar with formal queries and databases.

Methods based on statistical machine learning and neural networks have begun to appear in recent years. A methodology is proposed in [5] implements reinforcement training. This approach suggests using a deep neural network to translate natural language phrases into corresponding SQL queries. This method takes advantage of the structure of SQL queries to significantly reduce the output range of generated queries. Despite the huge data set in the training sample, the accuracy of the model was not high: the accuracy of execution was 59.4 %, the accuracy © Litvin A. A., Velychko V. Yu., Kaverynskyi V. V., 2021
 DOI 10.15588/1607-3274-2021-2-11

racy of the logical form was 48.3 %. Thus, for the real implementation of such system, an extensive collection of training sets will be required, which is difficult to provide.

As an example of converting natural language queries into formal queries in SPARQL can be considered the LODQA system presented in [6]. It parses a natural language phrase and creating a graphical representation of the request, which is called a pseudo graphic template. The pseudo graphic template is a graph pattern for finding the target graph of RDF subgraphs that match it. Natural language term can be normalized for more than one RDF term due to ambiguity. Therefore, from one pseudo graphic template, more than one linked template can be obtained by normalization. In order to take into account the structural inconsistency between the attached pseudo graphic template and the actual structure in the target data set, it tries to generate SPARQL queries for all possible structural variations. The considered LODQA system is focused on the English language only. Detailed features of its functioning are not given in [6], limiting only to a general description and analysis of work examples.

Another approach to natural language to SPARQL conversion is considered in [7] PAROT framework. PAROT adopts an approach that generates the most likely triple from a user query. The triple is then validated by the lexicon. It relies on a dependency parser to process user's queries to user triples. The user triples are then converted to ontology triples by the lexicon. The triples generated by the lexicon are used to construct SPARQL query that fetches the answers from the underlying ontology. Testing the PAROT framework by the authors of [7] showed that for simple questions it demonstrates about 81 – 82 % precision, about 43 – 56 % for complex questions, and for a specific thematic dataset (geography) precision was up to 88 %. But even the authors pay attention to some weaknesses of PAROT: it has low precision and recall when processing aggregation based questions.

In spite most of natural language to formal query conversion systems deal with English, works is also proceeding for other languages. For example, in [8] proposed a method to generate SPARQL queries from Korean natural language queries.

Taking into account the success in machine translation systems development using neural networks achieved in recent years it is no wonder that approach is trying to be applied to translation from a natural language to a formal query language such as SPARQL. For now, there are only a few examples of such approach, for example [9]. Some tricks to avoid or at least minimize criticality of typical machine translation mistakes are taken into account there. For instance, to train the model they use not SPARQL queries as they are but previously converted them into special sequences where language symbols and constructions are encoded as constant symbol sequences. Some constant query structure elements were omitted or abbreviated. Thus the translations result in this method is a specific sequence that is actually an instruction, by which it is possible to build a sufficient SPARQL query. Despite the authors of [9] claimed a rather good accuracy of their

model its behavior on others datasets related to different ontologies is still a question.

An example of natural language to SPARQL conversion open source realization is FREyA [10]. It is available on GIT-hub [11]. FREyA is an interactive Natural Language Interface for querying ontologies. It uses syntactic parsing in combination with the ontology-based lookup in order to interpret the question and involves the user if necessary. The user's choices are used for training the system in order to improve its performance over time. It deals with English language. In [11] some examples of natural language questions are given can be converted to SPARQL using FREyA. It seems that the query formation by FREyA very depends on the ontology data, and its configuration needs to be tuned to the certain ontology.

Conversion systems for natural language questions to Cypher (a query language for graph database Neo4j) are rather less developed, but works in this area are also proceeding. Analyzing posts in topical internet forums it seems to be that development of such tool is highly desirable. Here are a few examples of such works: [2, 12]. The system proposed in [12] is rather primitive. The queries need to be of a pre-defined structure; actually it needs a file with ready natural language sentences where some words are replaced by placeholders and matching Cypher templates. The described approach has its advantages and disadvantages. The main advantage is simplicity, which guarantees that the result will be just obtained or not obtained without appearing of strange situations when wrong or not completely correct result springs up. But obviously, there are a lot of drawbacks. First of all, for a real big system, a large number of phrases templates is needed which involves all possible users' questions considering their variety.

The flexibility of the query template approach can be increased if rigid phrases templates will be replaced with semantic analysis of an input phrase which also recognizes the entities (words) that are to be substituted into the query template [13]. In this work is considered the problem of selection of a correct query template and singling out corresponding entities basing on the user's phrase.

3 MATERIALS AND METHODS

The only input information coming from the user is a phrase in Ukrainian language. It should be a single sentence. If there are several sentences in the user's reply, each of them will be separately considered by the system. For this purpose the input text is tokenized to a list of sentences. Each sentence is then tokenized to a list of words. These operations are conforming so called preliminary graphematic analysis. Then these words list are cleaned from senseless and not informative words (interjections, emotion expressions, senseless introductive words). For this task lists of such words are pre-prepared. Each of the words from the input phrase is lemmatized and comparing with words from this lists. If it matches this word is to omit from the sequent processing. Obtained in such way input material than used for semantic analysis.

© Litvin A. A., Velychko V. Yu., Kavrynskyi V. V., 2021
DOI 10.15588/1607-3274-2021-2-11

The first goal of the semantic analysis is to determine the semantic type of the input phrase, which is φ from the formal problem statement, i.e. what kind of information the user wants to receive. The considered here system deals only with interrogative and imperative sentences. Narrative sentences where nosing is implicitly asking are out of consideration in this work.

The criteria of the semantic type determine are facts of the presence of certain words and words sequences (including preposition) in certain forms. The main complicity is that there are a great number of possible semantic intents φ and many of them could be represented in plenty of different ways (A). Hence, runtime enumerating all possible options is a long and unproductive way. In the basis of the proposed here solution for is a tree approach when analysis is going step by step and on each step only a few factors are considered that allows one to exclude many of others variants so they do not need sequent consideration. The determination algorithm is proceeding until all the necessary and sufficient conditions for the certain semantic type appear observed. The method also allows at the same process to find input entities to be substituted into the query template. Thus, this determinative tree is actually γ from the formal problem statement.

Let us consider the proposed determination method in more detail.

Even in inflective language, at least in Ukrainian and Russian words are not going completely random but some patterns of their sequence exist, which significantly simplifies the analysis. For instance, most types of questions, excluding so called general questions, begin with question words (“що” – “what”, “коли” – “when”, “як” – “how”, etc.), which are the most crucial factor in the process of the semantic type determination. Imperative sentences begin with a verb in imperative form, or with appeal followed by such verb. Preposition in Ukrainian and Russian languages can go only before nouns, pronouns and name groups (the linked group of nouns and adjectives that in aggregate describe one entity). Thus, the tree algorithm analyses not just presence of a certain word but it goes through the words of the phrase as a sequence.

The algorithm could be described like a frame moving through the sequence of words. The size of this frame is from one word to several and is equal to a number of words in consideration during the current condition analysis. The tree is a-ary, so there is no limitation of outgoing links from its node to the nodes of the lower level. The tree, its conditions, should be constructed in such manner that only one possible way must exist to the lower level from a node; i. e. algorithm should go through the tree without dividing. If not only one from the conditions options of the node are suitable, one that is the most suitable must be selected (but such situations are reasons to improve the conditions, and may be add another level of conditions). More suitable means the following:

– if the number of checked words (size of frame) in condition A is bigger than it in condition B and both conditions are matching the current situation, the preference is to be given to condition A (with bigger frame size);

– if condition A matches only by word characteristic (verb in the certain form, noun in some case, etc) and condition B matches by the text (at least in lemma form), the preference is to be given to condition B (text matching);

– if condition A matches only by part of speech (just verb, or noun, or adjective etc.) and condition B matches also by word form characteristic (case, tense, verb form etc.), the preference is to be given to condition B (word form matching).

On each step, only one of the conditions should be chosen, which determines the conditions checked in the following step. If there is no matching condition the only result is that the answer can't be obtained, which can appear if the conditions system in the tree is not full. Such situation is perfectly acceptable if the tree is specialized for a narrow subject area and number of existing query templates is rather restricted. Such approach makes the system easier and faster by early screening out the questions for which there are no answers in the database. Each following step could restrict the number of possible query templates or just check the next necessary condition. At the end, only one of the possible templates should remain and all the necessary and sufficient conditions for it should be observed. If in some case more than one query template is remain (that is not desirable) several formal queries are to be formed, but that is evidence of the initial phrase nebulosity or imperfection of the determination system. For some cases the frame size is flexible, it is necessary when there expected a possibility of a words group with some characteristics but of not determined length (name group, homogeneous parts of the sentence, parts of full name, date, etc.). The frame shift for the following step depends on the actual frame size of the selected condition, so each time only new words appear in consideration, crossover shifts of the frame are not provided in this method. To determine the semantic type of the user's phrase (γ) and the corresponding query template is important, but not enough. A set of entities substituted to the query template is also needed. Selection of those entities is performed after the template determination, because they depend on it. For this purpose in the end node of the solution tree are also indicated positions of the way in the tree (by level) from where corresponding variables of the template should be initialized. If in some position a number of entities are present (described above cases of flexible size frames), a list will be obtained to initialize the template variable. If some of template variables are initialized as list, the condition row of the template during variables values substitution will be repeated for the each value, if this variable in the template is marked as "allow list", otherwise its value shall be made by joining the list members. The first option ("allow list") is better for homogeneous parts of the sentence and the second one – for name groups, which are stored in the ontology (graph database) in such joined way. In some cases of narrow subjected ontologies some of the entities are not to be substituted from the user's phrase but are predefined in the template, which makes it more reliable

in reasons of ontology answer obtaining possibility. Also, there could be some cases when obtained from the user's phrase entity is replaced by its synonym which presence in the ontology is guaranteed but of the user's word is not. If on the expected graph way position there is no entities to initialize the corresponding template variable (undesirable, degenerate, but also possible case), this condition string shall be omitted in the forming query, so the query becomes wider.

Here we limited to just a brief description of the proposed query template method without linguistic and semantic details. They very depend on the language, the subject area and the peculiarities of the ontology structure accepted in the system. It only should be noticed, that the full and complete system of semantic types of questions is not always necessary for most of the databases. It could be limited to only those types of queries types for which the answers in this ontology are expected and provided.

The common scheme of the provided natural language to formal query algorithm is illustrated by a UML activity diagram on Fig. 1.

Putting query templates and the template selection tree scheme into separate files allows a developer to adopt the system to certain ontology without touching the program code logics. The program code is in a python file to which a linguistic analyzer is bind to deal with a specific natural language.

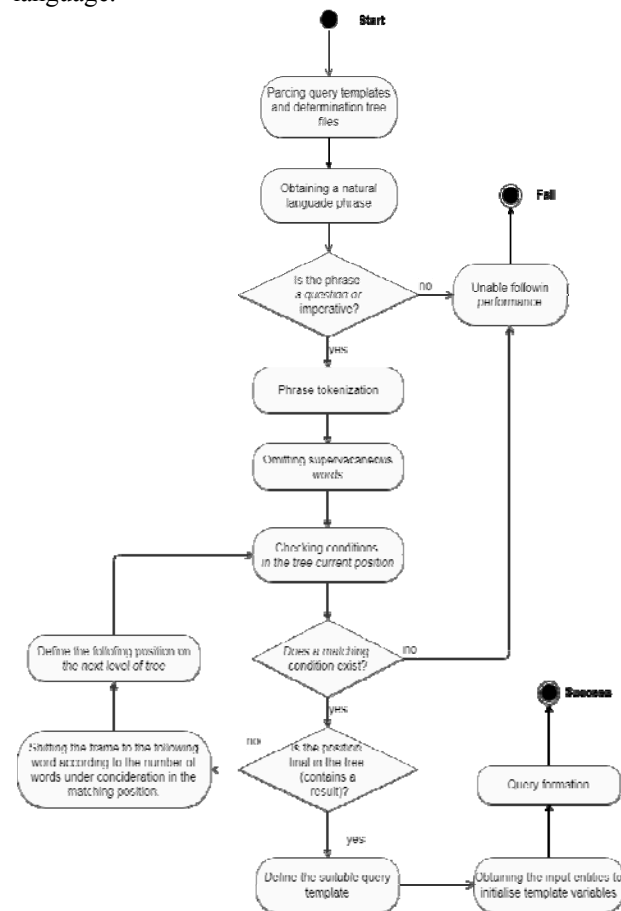


Figure 1 – UML activity diagram of the process of natural language phrase to a formal query conversion

The input information is a natural language user's phrase and the output is a SPARQL or Cypher formal query. The modulus is included to a program agent (web service) shell, so it could easily become a part of a natural language dialog or reference system.

4 EXPERIMENTS

To test the developed system we used the OWL ontology of letters written to the famous Ukrainian writer of XX century Oles Honchar. This ontology was semiautomatically constructed on the basis of the digest book of the letters [14]. Let us briefly describe the ontology structure. Now the working system is available by the following URL: <https://oles-gonchar-bot.herokuapp.com> This dialog web application involves not only the ontology for Oles Honchar letters but also several other ontologies devoted to the literary work of this writer. All ontologies are in OWL (RDF/XML) and SPARQL is used to query them. Neo4j graph database and Cypher queries are used in another developed by us dialog system which is a virtual financial consultant with a natural language interface (in this case – Norwegian). This application is available by the next URL: <http://178.128.245.158:8888/chatbot/>

For the system performance testing serial of experiments were carried out. As input were used natural language (Ukrainian) interrogative and imperative phrases. To estimate obtaining negative and positive results 30 of the phrases were directly related to the considered in the ontology subject (letters were written to Oles Honchar) and the answer on them supposed to be in the ontology, the others 30 of them were made specially to obtain a negative result. These phrases for a negative result obtaining were not completely meaningless or grammatically incorrect (because for such ones there is no a certain positive result). They actually could be divided into three groups: in the first group phrases are devoted to Oles Honchar letters but there is no a supposed answer to them (20 phrases), in the second one the phrases were about letters but not to Oles Honchar (10 phrases), and in the third one phrases were about Oles Honchar but not about letters (10 phrases). In total, a set of 60 test phrases was used in the testing all phrases were grammatically correct Ukrainian sentences. Not obtaining an answer to a phrase about a completely different subject is trivial, thus they were not used for the testing.

An answer of the system was considered as a true positive if the answer was informative and was given as it was supposed to be. True negative result is the absence of the requested information (but not wrong or not relevant information!) but only in the case when we expected it. As false negative we considered all results where a certain answer was expected but was obtained either wrong answer, not a complete answer, or no answer. False positive result may contain any information (complete or not, maybe even not completely correct) in the case when the absence of an answer was expected.

The obtained scour was used for the calculation of accuracy, precision, recall and F_1 criteria.

Also the time intervals of the phrase analysis and query formation process were measured. These times include just these process but not also the times of query execution, messages between agent and to the user sending, page rendering etc.

5 RESULTS

The scour for each type of the obtained during the testing results is shown in table 1. The values of formal estimation criteria are given in table 2.

Table 1 – Experimental results score

Type of result	Number of results
T_p	24
T_n	29
F_n	6
F_p	1

Table 2 – Evaluation criteria of the system performance results values

Criterion	Value
Accuracy	0.883
Precision	0.960
Recall	0.800
F_1	0.873

It should be noticed that one answer considered here as a false positive is actually not a completely correct answer, but the true negative result for it ought to be “no answer”. From the false negative results actually in four cases there were no answer and in two ones the answer was but incorrect.

The average time interval of the analysis and query formation process was 6 ± 3 ms. Any statistically valuable decadence of it from the phrase length and complicity was not observed.

6 DISCUSSION

Generally, according to the values of the criterion, the developed system seems to be rather qualitative and is on the level of claimed in [2, 7, 10]. Thus the proposed method may be workable and practically acceptable. But some of its peculiarities revealed in the experimental testing are to be discussed in more detail.

The system shows a high precision criterion value because it is not inclined to false positive results. The cause of this is the sensitivity of the tree-based method and then ontology querying system to not subjected information. It rather gives no results in this case than some results. Additional information, which we did not expect to receive, but suddenly received, simply cannot be taken with this approach to the construction of the ontology and queries to it. But sometimes seldom if the input information is rather similar to one stored in the ontology some answers (may be not very relevant) could be obtained.

Unlike the precision, recall criterion is not so high. This is mostly because the system tends to not find answers even if they are in the database. The reason for this is some incompleteness of the decision tree rules (some possible constructions or/and distinguishing words were not taken into account). Another cause may be that some

words which should to be ignored actually were not putted to the corresponding lists. This type of problem could be solved quite easily – by adding the necessary lacking rules and words lists. But the problem of wrong answers obtaining is some more complicated: it needs increasing of the semantic analysis profundity and precision for distinguishing more subtle aspects, also taking into account homonymy and correct interpretation of pronouns.

In the whole, the developed method has the potential to show a rather good accuracy and high performance. It is scalable and could be tuned both to small narrow subjected ontologies or wide and complicated ones. Moreover, it is rather simply correctible and the obtaining results could be easily explained and if needed it is clear what and where the corrections and additions have to be made. It is the same suitable for SPARQL and Cypher queries, and may by others query languages used in graph or relations databases.

For the moment it is elaborated only to some narrow subjected ontologies that need a rather limited variety of query templates and templates are constructed not for all the possible semantic types of questions. In the future, we are to make a more thoughtful and complicated decision tree to determine most of the semantic types of questions and imperative sentences.

As a main disadvantage of the proposed method is the need for manual creation of the query templates and the decision tree to determine the most suitable template, which is also very depend on the ontology structure. That may take a lot of effort. Thus, it seems to be suitable for the systems using narrow subjected ontologies that need to have high accuracy and stability of the answers and where databases are tend mostly growing in data size but not significantly change their structure.

CONCLUSIONS

A method is proposed for the conversion of natural language questions and imperative phrases to formal queries in SPARQL and Cypher query languages used in graph databases. The proposed technique assumes the presence of several query templates corresponding to each for a special semantic type of question. Meaningful entities extracted from the user's natural language phrase are substituted into the corresponding query template. For the most suitable query template selection a tree based semantic analysis method is proposed. The method assumes that the frame is shifting through the words list of the phrase considering on each step one or several words. These words are analyzed to match one of the conditions on the current tree position. The most matching condition determines the following position on the next level of the tree. The process proceeds until there will remain the only option of query template and all the sufficient conditions for the corresponding semantic type are proofed to be observed. Then depending on the selected query template input entities for it are taken from the given positions of the previous consideration.

The scientific novelty of the proposed approach is following: a technique of automatically conversion of a natural language phrase to formal queries in SPARQL and CYPHER query languages was further developed, the peculiarity of the presented method is using flexible query templates which selection and variables substitution is controlled by the tree-based semantic analysis, development and application of such methods is currently little elaborated; important feature of the proposed approach is its capability for inflective languages, especially, Ukrainian and Russian, for which such tools are poorly developed for the moment; were investigated the efficiency and possibilities for practical application of the proposed tree-based semantic analysis approach, standard criterions values were estimated, data on the assessment of this method were not previously known.

The practical value of the obtained results is that the proposed method seems to be useful in dialog and reference program systems which use graph databases either OWL/RDF or Neo4j based. The performance results of a program that implemented the proposed technique are clearly interpretable and explainable, which makes this approach highly customizable, fixable, and extensible. The tree-based method is rather scalable and could be suitable either for small narrow-subjected databases or big and complicated ones.

Quantitative indicators of the research results showed during the system testing are expressed by the following evaluation criteria values: $A_C = 0.883$, $P = 0.960$, $R = 0.800$, $F_1 = 0.873$. The obtained testing results seem rather reassuring and promising. Thus, the developed method is useful in constructing dialog and reference systems with a natural language interface. Also it shows the ability to easily corrections and editing if needed.

Prospects for further research are seemed as follows: investigation of different an more complicated graph data bases which represent more possible types of semantic relationships allocation explicitly and implicitly existing in them and constructing of new types of formal query templates for this purpose; further developing the proposed here semantic analysis technique for more (both general and subject oriented) semantic types determination, including mixed ones, and methods of meaningful concepts obtaining for them; development of methods that are able to as better as possible fit the terms obtained from an input natural language phrase to the corresponding values of the nodes in the graph database, which increases the possibility of an answer obtaining even for cases of long multiword terms in the nodes and expanding and narrowing an input term context regarding them.

ACKNOWLEDGEMENTS

Presented here investigation and results are the part of the research performed in the framework of the department project which is carried out in The V.M. Glushkov Institute of Cybernetics “Development of the methods for creating and using tools for working with knowledge

bases of information systems for research” (state registration number 0119U002226).

REFERENCES

1. Galitsky B. Developing Enterprise Chatbots. Learning Linguistic Structures. Berlin, Springer, 2019, 566 p.
2. Sun C. A Natural Language Interface for Querying Graph Databases: master's thesis ... master in computer science and engineering. USA, Massachusetts Institute of Technology, 2018, 69 p.
3. Palagin O. V., Krivij S. L., Bibikov D. S., Velichko V. Ju. ta in. Formal-logical approach to building analysis systems of knowledge in different domains, *Problems in programming*, 2010, No. 2–3, pp. 382–389.
4. Li F., Jagadish H. V. Understanding natural language queries over databases, *SIGMOD Record*, 2016, Vol. 45, pp. 6–13. DOI: 10.1145/2949741.2949744
5. Zhong V., Xiong G., Socher R. Seq2sql: generating structured queries from natural language using reinforcement learning, 2017 [Electronic resource]. Access mode: <https://arxiv.org/pdf/1709.00103.pdf>. arXiv: 1709.00103
6. Shaik S., Kanakam P., Hussain S. M., Suryanarayana D. Transforming natural language query to SPARQL for semantic information retrieval, *International Journal of Engineering Trends and Technology*, 2016, No. 7, pp. 347–350. DOI: 10.14445/22315381/IJETT-V41P263
7. Ochieng P. PAROT: Translating natural language to SPARQL, *Expert Systems with Applications*, 2020, No. 5, pp. 1–16. DOI: 10.1016/j.eswa.2021.114712
8. Jung H., Kim W. Automated conversion from natural language query to SPARQL query, *Journal of Intelligent Information Systems*, 2020, Vol. 55, pp. 501–520. DOI: 10.1007/s10844-019-00589-2
9. Yin X., Gromann D., Rudolph S. Neural machine translation from natural language to SPARQL, *Future Generation Computer Systems*, 2021, Vol. 117, pp. 510–519. DOI: 10.1016/j.future.2020.12.013
10. Damjanovic D., Agatonovic M., Cunningham H. FREYA: an interactive way of querying linked data using natural language, *The Semantic Web: ESWC 2011 Workshops*, 2011, pp. 125–138. DOI: 10.1007/978-3-642-25953-1_11
11. GIT-hub: FREYA documentation [Electronic resource]. Access mode: <https://github.com/nmvijay/freya>
12. GIT-hub Convert English sentences to Cypher queries documentation [Electronic resource]. Access mode: <https://github.com/gssrao/english2cypher>
13. Litvin A. A., Velychko V. Yu., Kaverynskyi V. V. Method of information obtaining from ontology on the basis of a natural language phrase analysis, *Problems in programming*, 2020, No 2–3, pp. 322–330. DOI: 10.15407/pp2020.02-03.322
14. Kiral' S. S. recenzenti, ta in. pid zagal'n. red. M. Stepanenka Listi do Olesja Gonchara. Kyiv, Sakcent Pljus, Vol. 1, 1946–1982, 2016, 736 p.

Received 09.03.2021.
Accepted 29.04.2021.

УДК 004.93

БАЗОВИЙ НА ДЕРЕВІ МЕТОД СЕМАНТИЧНОГО АНАЛІЗУ ФРАЗИ ПРИРОДНОЮ МОВОЮ ДЛЯ ПЕРЕТВОРЕННЯ ЇЇ У ФОРМАЛЬНИЙ ЗАПИТ

Литвин А. А. – аспірантка, аспірантка відділу мікропроцесорної техніки, Інститут кібернетики ім. В. М. Глушкова, Київ, Україна.

Величко В. Ю. – канд. техн. наук, доцент, старший науковий співробітник відділу мікропроцесорної техніки, Інститут кібернетики ім. В. М. Глушкова, Київ, Україна.

Каверінський В. В. – канд. техн. наук, старший науковий співробітник відділу зносо- і корозійностійких порошкових конструкційних матеріалів, Інститут проблем матеріалознавства ім. І. М. Францевича, Київ, Україна.

АНОТАЦІЯ

Актуальність. Ця робота присвячена проблемі побудови природномовного інтерфейсу для отримання інформації з графових баз даних. Основна увага приділяється методам перетворення фраз природною мовою у формальні запити на мовах запитів SPARQL та CYPRER.

Мета. Цілями роботи є створення методу семантичного аналізу типу вхідних природномовних фраз та виділення з них значущих сутностей для ініціалізації змінних шаблону запиту, побудова гнучких шаблонів запитів для відповідних семантичних типів фраз, розробка програмної реалізації запропонованого способу.

Метод. Розроблено метод, що базується на дереві прийняття рішень, для визначення семантичного типу фрази користувача і отримання з неї набору понять, для підстановки їх у певні місця найбільш підходящого шаблону формального запиту. Пропонована методика вирішує завдання визначення типу фрази (що безпосередньо пов'язано з критерієм вибору шаблону формального запиту) і отримання значущих понять, для ініціалізації змінних обраного шаблону. У поточній роботі розглядаються тільки питальні й наказові фрази користувача, тобто ті, які в явному вигляді пропонують системі дати відповідь. Передбачається, що розглянута діалогова або довідкова система використовує графову онтологічну базу даних, що безпосередньо впливає на формальні шаблони запитів – результуючі запити використовують SPARQL або Cypher. Приклади семантичного аналізу, розглянуті в цій роботі, відносяться переважно до мов флективного типу, а саме, української та російської, але основні принципи можуть бути придатними і для більшості інших мов.

Результати. Розроблений метод перетворення фрази на природній мові у формальний запит на SPARQL або CYPRER було програмно реалізовано для української та норвезької мов із використанням вузьких предметних онтологій та протестовано на відповідність формальним критеріям ефективності.

Висновки. Запропонований метод дозволяє діалоговій системі швидко та з мінімальною кількістю кроків вибрати найбільш підходящий шаблон запиту та витягти інформативні сутності із вхідної природномовної фрази, враховуючи величезну варіативність фраз у флективних мовах. Проведені експерименти показали високу точність та надійність побудованої системи та її потенціал для практичного використання та подальшого розвитку.

КЛЮЧОВІ СЛОВА: обробка природної мови, графова база даних, семантичний аналіз, формальний запит, дерево прийняття рішень, онтологія.

© Litvin A. A., Velychko V. Yu., Kaverynskyi V. V., 2021
DOI 10.15588/1607-3274-2021-2-11

ОСНОВАННЫЙ НА ДЕРЕВЕ МЕТОД СЕМАНТИЧЕСКОГО АНАЛИЗА ФРАЗЫ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ ДЛЯ ПРЕОБРАЗОВАНИЯ ЕЕ В ФОРМАЛЬНЫЙ ЗАПРОС

Литвин А. А. – аспирантка, аспирантка отдела микропроцессорной техники, Институт кибернетики им. В.М. Глушкова, Киев, Украина.

Величко В. Ю. – канд. техн. наук, доцент, старший научный сотрудник отдела микропроцессорной техники, Институт кибернетики им. В.М. Глушкова, Киев, Украина.

Каверинский В. В. – канд. техн. наук, старший научный сотрудник отдела износо- и коррозионно-стойких порошковых конструкционных материалов, Институт проблем материаловедения им. И.М. Францевича, Киев, Украина.

АННОТАЦИЯ

Актуальность. Данная работа посвящена проблеме построения интерфейса на естественном языке для онтологических графовых баз данных. Основное внимание уделяется методам преобразования фраз на естественном языке в формальные запросы на языках запросов SPARQL и Cypher.

Цель. Целями работы является создание метода семантического анализа типов входных естественных фраз и выделения из них значимых сущностей для инициализации переменных шаблона запроса, построение гибких шаблонов запросов для соответствующих семантических типов фраз, разработка программной реализации предложенного способа.

Метод. Разработан метод, основанный на дереве принятия решений, для определения семантического типа фразы пользователя и получения из нее набора понятий, для подстановки их в наиболее подходящие шаблоны формального запроса. Предлагаемая методика решает задачу определения типа фразы (непосредственно связано с критерием выбора шаблона формального запроса) и получения значимых понятий, для инициализации переменных выбранного шаблона. В текущей работе рассматриваются только вопросительные и повелительные фразы пользователя, то есть те, которые в явном виде предлагают системе дать ответ. Предполагается, что рассматриваемая диалоговая или справочная система использует графовую онтологическую базу данных, что непосредственно влияет на шаблоны формальных запросов – в получаемых в результате запросах используются SPARQL или Cypher. Примеры семантического анализа, рассмотренные в этой работе, относятся преимущественно к языкам флективного типа, а именно, украинскому и русскому, но основные принципы могут быть пригодными и для большинства других языков.

Результаты. Разработанный метод преобразования фразы на естественном языке в формальный запрос на SPARQL или Cypher был программно реализован для украинского и норвежского языков с использованием узких предметных онтологий и протестирован на соответствие формальным критериям эффективности.

Выводы. Предложенный метод позволяет диалоговой системе быстро и с минимальным количеством шагов выбрать наиболее подходящий шаблон запроса и извлечь информативные сущности их исходной естественной фразы, учитывая огромную вариативность фраз в флективных языках. Проведенные эксперименты показали высокую точность и надежность разработанной системы и ее потенциал для практического использования и развития.

КЛЮЧЕВЫЕ СЛОВА: обработка естественного языка, графовая база данных, семантический анализ, формальный запрос, дерево принятия решений, онтология.

ЛИТЕРАТУРА / LITERATURE

- Galitsky B. Developing Enterprise Chatbots. Learning Linguistic Structures / B. Galitsky – Berlin : Springer, 2019. – 566 p.
- Sun C. A Natural Language Interface for Querying Graph Databases: master's thesis ... master in computer science and engineering / C. Sun. – USA: Massachusetts Institute of Technology, 2018. – 69 p.
- Formal-logical approach to building analysis systems of knowledge in different domains / [O. V. Palagin, S. L. Krivij, D. S. Bibikov, V. Ju. Velichko ta in.] // Problems in progamming. – 2010. – № 2–3. – P. 382–389.
- Li F. Understanding natural language queries over databases / F. Li, H.V. Jagadish // SIGMOD Record. – 2016. – Vol. 45. – P. 6–13.
- Zhong V. Seq2sql: generating structured queries from natural language using reinforcement learning / V. Zhong, G. Xiong, R. Socher – 2017 [Electronic resource]. Access mode: <https://arxiv.org/pdf/1709.00103.pdf>
- Transforming natural language query to SPARQL for semantic information retrieval / [S. Shaik, P. Kanakam, S. M. Hussain, D. Suryanarayana] // International Journal of Engineering Trends and Technology. – 2016. – № 7. – P. 347–350.
- Ochieng P. PAROT: Translating natural language to SPARQL / P. Ochieng // Expert Systems with Applications. – 2020. – № 5. – P. 1–16.
- Jung H. Automated conversion from natural language query to SPARQL query / H. Jung, W. Kim // Journal of Intelligent Information Systems. – 2020. – Vol. 55. – P. 501–520.
- Yin X. Neural machine translation from natural language to SPARQL / X. Yin, D. Gromann, S. Rudolph // Future Generation Computer Systems. – 2021. – Vol. 117. – P. 510–519.
- Damljanovic D. FREyA: an interactive way of querying linked data using natural language / D. Damljanovic, M. Agatonovic, H. Cunningham // The Semantic Web: ESWC 2011 Workshops. – 2011. – P. 125–138.
- GIT-hub: FREyA documentation [Electronic resource]. Access mode: <https://github.com/nmvijay/freya>
- GIT-hub Convert English sentences to Cypher queries documentation [Electronic resource]. Access mode: <https://github.com/gssrao/english2cypher>
- Litvin A. A. Method of information obtaining from ontology on the basis of a natural language phrase analysis / A. A. Litvin, V. Yu. Velychko, V. V. Kaverynskyi // Problems in programming. – 2020. – № 2–3. – P. 322–330.
- Листи до Олеса Гончара / [рецензенти: С. С. Кіраль та ін.]. – Київ : Сакцент Плюс. Т. 1: 1946–1982 ; під загальн. ред. М. Степаненка. – 2016. – 736 с.

СТОХАСТИЧНА ПСЕВДОСПІНОВА НЕЙРОННА МЕРЕЖА З ТРИДАГОНАЛЬНИМИ СИНАПТИЧНИМИ ЗВ'ЯЗКАМИ

Пелешак Р. М. – д-р фіз.-мат. наук, професор кафедри інформаційних систем та мереж, Національний університет «Львівська політехніка», Львів, Україна.

Литвин В. В. – д-р техн. наук, професор, завідувач кафедри інформаційних систем та мереж, Національний університет «Львівська політехніка», Львів, Україна.

Черняк О. І. – аспірант кафедри математики, Дрогобицький державний педагогічний університет імені Івана Франка, Дрогобич, Україна.

Пелешак І. Р. – аспірант кафедри інформаційних систем та мереж, Національний університет «Львівська політехніка», Львів, Україна.

Дорошенко М. В. – канд. фіз.-мат. наук, доцент кафедри інформатики та інформаційних систем, Дрогобицький державний педагогічний університет імені Івана Франка, Дрогобич, Україна.

АНОТАЦІЯ

Актуальність. Для скорочення часу обчислювального ресурсу в задачах діагностування та розпізнавання спотворених образів на основі повнозв'язної стохастичної псевдоспінової нейронної мережі виникає необхідність прорідження синаптичних зв'язків між нейронами, що вирішується за допомогою методу діагоналізації матриці синаптичних зв'язків без втрати взаємодії між всіма нейронами в мережі.

Мета роботи. Створення архітектури стохастичної псевдоспінової нейромережі з розрідженими та діагональними синаптичними зв'язками без втрати взаємодії між всіма нейронами в мережі для зменшення часу її навчання.

Метод. У статті використовується метод Хаусхолдера, метод стиску вхідних образів на основі діагоналізації матриці синаптичних зв'язків та система комп'ютерної математики MATLAB для зведення повнозв'язної нейромережі до тридіагонального вигляду з прихованими синаптичними зв'язками між всіма нейронами.

Результати. Розроблено модель архітектури стохастичної нейромережі з розрідженими перенормованими синаптичними зв'язками, які враховують вилучені синаптичні зв'язки. На основі перетворення матриці синаптичних зв'язків повнозв'язної нейронної мережі до матриці Гессенберга з тридіагональними синаптичними зв'язками запропоновано перенормоване локальне правило Хебба. За допомогою системи комп'ютерної математики «WolframMathematica 11.3» розраховано в залежності від числа нейронів N відносний час налаштування синаптичних зв'язків (за одну ітерацію) у стохастичній псевдоспінової нейронній мережі з тридіагональною матрицею зв'язків, відносно часу налаштування синаптичних зв'язків (за одну ітерацію) у повнозв'язній синаптичній нейронній мережі.

Висновки. Встановлено, що зі збільшенням числа нейронів час налаштування синаптичних зв'язків (за одну ітерацію) у стохастичній псевдоспінової нейронній мережі з тридіагональною матрицею зв'язків, відносно часу налаштування синаптичних зв'язків (за одну ітерацію) у повнозв'язній синаптичній нейронній мережі, зменшується за гіперболічним законом. В залежності від напрямку псевдоспінів нейронів, запропоновано класифікацію перенормованої нейронної мережі із феромагнітною структурою, антиферомагнітною структурою та дипольним склом.

КЛЮЧОВІ СЛОВА: нейронна мережа, синаптичні зв'язки, матриця зв'язків, тридіагоналізація.

НОМЕНКЛАТУРА

$b_{\mu n}(t)$ – вагові коефіцієнти синаптичних зв'язків між вхідними сенсорними та псевдоспіновими нейронами у нейронній мережі Гессенберга;

E – функція енергії штучної нейромережі взаємодіючих псевдоспінових нейронів;

$f(h_n)$ – функція розподілу Фермі;

h_n – потенціал n -го стохастичного нейрона;

$\langle h_n(t) \rangle$ – локальне самоузгоджене поле, яке діє на

псевдоспіновий нейрон S_n^z зі сторони решти псевдоспінових нейронів;

$h_n^{ext}(t)$ – це зовнішнє поле, яке складається з інформаційного сигналу $\xi_{\mu n}$ і шуму $V_{noise}^n(t)$ у вхідному каналі;

k_λ – коефіцієнт, що описує час налаштування одного синаптичного зв'язку повнозв'язної матриці;

k_γ – коефіцієнт, що описує час налаштування одного синаптичного зв'язку матриці Гессенберга;

N – кількість нейронів;

$N_{\lambda_{ij}}$ – кількість синаптичних зв'язків повнозв'язної нейронної мережі;

$N_{\hat{\lambda}}$ – кількість синаптичних зв'язків повнозв'язної симетричної квадратної матриці;

$N_{\hat{\gamma}}$ – кількість явних синаптичних зв'язків, які залишаються при перетворенні матриці $\hat{\lambda}$ у матрицю $\hat{\gamma}$;

P_r – ймовірність переходу нейрона в стани ± 1 ;

S_n – стан n -го стохастичного нейрона;

$\langle S_n \rangle$ – очікуване значення стану n -го стохастичного нейрона;

T – аналог температури, який використовують для керування рівнем шуму та ступенем невизначеності перемикання;

t_λ – час налаштування синаптичних зв'язків λ_{ij} ;
 t_γ – час налаштування синаптичних зв'язків γ_{nk} ;
 α – коефіцієнт, який показує, у скільки разів зменшується кількість синаптичних зв'язків нейронної мережі після приведення її до тридіагонального вигляду
 $\hat{\gamma}$ – тридіагональна симетрична матриця зв'язків;
 γ_{nk} – синаптичні зв'язки нейромережі Гессенберга;
 $\delta_{mm'}$ – символ Кронекера;
 α_c – параметр завантаження нейромережі;
 $\hat{\lambda}$ – симетрична матриця повнозв'язних синаптичних зв'язків;
 ξ_i та ξ_j – образи i -го та j -го пікселів відповідно;
 $\{\xi_{mn}\}$ – набір, який представляє собою інформаційну копію нейронної мережі.

ВСТУП

Під час розв'язування задач розпізнавання спотворених образів за допомогою нейронних мереж виникають практичні питання мінімізації розміру нейромережі, архітектури синаптичних зв'язків між нейронами, зменшення часу навчання нейронної мережі, збільшення ємності P нейромережі ($P = \alpha_c N$, де α_c – параметр завантаження нейромережі, максимальне на даний час значення якого $\alpha_c \approx 0,14$ [1]) та підвищення ступеня узагальнення функціональної здатності нейромережі без втрати її продуктивності.

На даний час розмір нейромережі мінімізують за рахунок спрощення її структури [1]. Зокрема, є два підходи, щоб спростити структуру нейромережі: перший базується на «регуляризації», а другий – на вилученні із нейронної мережі синаптичних зв'язків.

У даній роботі спрощення структури повнозв'язної псевдоспінової нейронної мережі базується на приведенні симетричної матриці повнозв'язних синаптичних зв'язків до тридіагональної матриці Гессенберга [2]. У запропонованому підході зберігається взаємодія між всіма псевдоспіновими нейронами, на відміну від підходу вилучення синаптичних зв'язків. Ця взаємодія враховується у перенормованих зв'язках між найближчими сусідніми псевдоспіновими нейронами. При цьому підвищується ступінь узагальнення функціональної здатності нейромережі.

Об'єктом дослідження є процес розробки стохастичної псевдоспінової нейронної мережі з мінімізованою архітектурою синаптичних зв'язків.

Предметом дослідження є метод оцінки відносного часу налаштування синаптичних зв'язків (за одну ітерацію) у стохастичній псевдоспіновій нейронній мережі з тридіагональною матрицею синаптичних зв'язків у залежності від числа нейронів.

Метою даної роботи є створення архітектури стохастичної псевдоспінової нейромережі з мінімізованим числом синаптичних зв'язків для зменшення часу її навчання.

1 ПОСТАНОВКА ЗАДАЧІ

Нейрони штучної псевдоспінової повнозв'язної нейронної мережі розміщені в точках простору з координатами (рис. 1)

$$\vec{r}_i = m_i \vec{a}_1 + n_i \vec{a}_2 + p_i \vec{a}_3, \quad (1)$$

де $m_i, n_i, p_i \in Z$, $\vec{a}_1, \vec{a}_2, \vec{a}_3$ – базисні вектори.

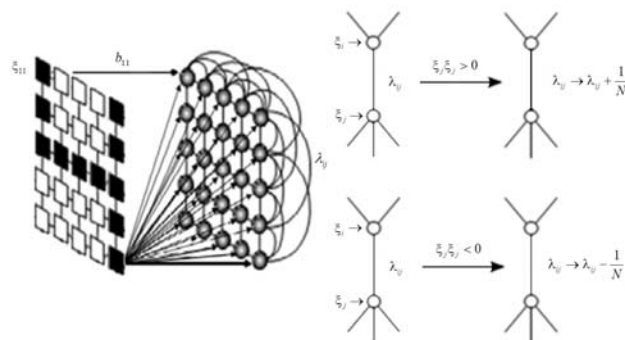


Рисунок 1 – Псевдоспінова нейронна мережа з повнозв'язними синаптичними зв'язками. Добуток $\xi_i \xi_j > 0$ вказує на підсилення синаптичного зв'язку між i -тим та j -тим нейронами, а $\xi_i \xi_j < 0$ – на його послаблення

Псевдоспінова нейромережа є найімовірнішою для конкретної реалізації динамічної нейронної мережі. Ця нейронна мережа складається з N стохастичних псевдоспінових нейронів, які взаємопов'язані синаптичними зв'язками λ_{ij} . Симетрична матриця повнозв'язних синаптичних зв'язків має наступний вигляд

$$\hat{\lambda} = \begin{pmatrix} 0 & \lambda_{12} & \lambda_{13} & \dots & \lambda_{1N} \\ \lambda_{12} & 0 & \lambda_{23} & \dots & \lambda_{2N} \\ \lambda_{13} & \lambda_{23} & 0 & \dots & \lambda_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{1N} & \lambda_{2N} & \lambda_{3N} & \dots & 0 \end{pmatrix}. \quad (2)$$

Для вирішення задачі мінімізації часу навчання нейронної мережі, пропонується звести симетричну матрицю повнозв'язних синаптичних зв'язків псевдоспінової нейронної мережі (2) до тридіагональної симетричної матриці $\hat{\gamma}$ (матриці Гессенберга [3]).

2 ОГЛЯД ЛІТЕРАТУРИ

У роботах [4, 5, 6] запропонована нова архітектура згорткових нейронних мереж, яка може бути застосована у випадку, коли вхідні дані не пов'язані між собою. Автори робіт [7, 8] досліджують задачу встанов-

лення зв'язку між нейронами мозку людини та мікротрубочками цитоскелета, зокрема у [7] описується енергія системи взаємодіючих псевдоспінів та розщеплення основного стану дипольної системи мікротрубочки на різні енергетичні стани.

Автор роботи [1] пропонує вилучення несуттєвих синаптичних зв'язків за рахунок матриці, оберненої до матриці Гессе, де виділяють несуттєві вагові синаптичні зв'язки багатозарового перцептрона і вилучають їх. У цьому підході функція вартості повинна бути двічі диференційованою по елементам вектора ваг синаптичних зв'язків.

Ще один спосіб вилучення ваг синаптичних зв'язків з нейронної мережі полягає в тому, що проводиться оцінка функції штрафу за складність

$$E_c(\bar{w}) = \sum_{i \in C_{total}} \frac{(w_i/w_0)^2}{1+(w_i/w_0)^2} \quad [9],$$

яка залежить виключно від самої моделі нейромережі. Оцінка базується на попередніх знаннях про модель нейромережі. Де w_0 – деякий наперед визначений параметр; w_i – вага i -го синаптичного зв'язку; C_{total} – множина всіх синаптичних зв'язків нейромережі. Якщо $|w_i| \gg w_0$ функція штрафу за складність (вартість) $E_c(\bar{w})$ для синаптичного зв'язку w_i досягає максимального значення і дорівнює 1. Це означає, що вага синаптичного зв'язку w_i має високу цінність для процесу навчання методом зворотнього поширення помилки. Таким чином функція штрафу дозволяє виявляти суттєві вагові синаптичні зв'язки між нейронами нейромережі, а не суттєві – вилучати.

У наступних роботах пропонується оптимізація структури нейронної мережі для таких задач, як обчислювання даних великомасштабних дата сетів [10] та автоматична розробка нейронних мереж на основі еволюційних алгоритмів [11].

Наступним методом мінімізації топології нейронної мережі (кількість нейронів, кількість синаптичних зв'язків між нейронами) є генетичний алгоритм (ГА) [12, 13].

Авторами роботи [14] запропоновано спрощення структури нейронної мережі за допомогою вилучення недіагональних компонентів синаптичних зв'язків методом діагоналізації матриці зв'язності.

3 МАТЕРІАЛИ І МЕТОДИ

Зведемо симетричну матрицю повнозв'язних синаптичних зв'язків псевдоспінової нейронної мережі (2) до тридіагональної симетричної матриці Гессенберга [3] $\hat{\gamma}$ за допомогою системи комп'ютерної математики Matlab з використанням функції *hess*, тобто $\hat{\gamma} = \text{hess}(\hat{\lambda})$ [15]

$$\hat{\gamma} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & & & \\ \gamma_{12} & \gamma_{22} & \gamma_{23} & & 0 \\ & \gamma_{23} & \gamma_{33} & & \\ \dots & \dots & \dots & \ddots & \dots \\ & 0 & & \gamma_{N-1,N-1} & \gamma_{N-1,N} \\ & & & \gamma_{N-1,N} & 0 \end{pmatrix}. \quad (3)$$

Тоді архітектура псевдоспінової нейронної мережі з повнозв'язними синаптичними зв'язками (рис. 1) набуває наступного вигляду (рис. 2). При такій трансформації архітектури синаптичних зв'язків взаємодія між всіма псевдоспіновими нейронами не зникає, а враховується у перенормованих зв'язках між найближчими сусідніми псевдоспіновими нейронами (суцільні криві на рис. 2), тобто $\gamma_{nk} = F(\lambda_{ij})$, де функція F описує закон перетворення матриці синаптичних зв'язків повнозв'язної нейронної мережі до тридіагональної матриці синаптичних зв'язків Гессенберга.

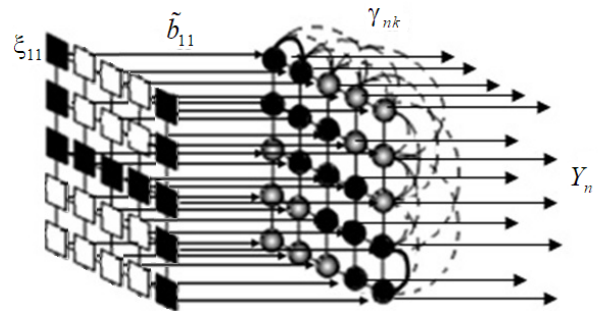


Рисунок 1 – Псевдоспінова нейронна мережа з розрідженими синаптичними зв'язками, отримана внаслідок приведення симетричної матриці повнозв'язних синаптичних зв'язків (2) до тридіагональної симетричної матриці (3)

Число синаптичних зв'язків повнозв'язної симетричної квадратної матриці (2), на головній діагоналі якої містяться нульові елементи ($\lambda_{ii} = 0$), визначається співвідношенням

$$N_{\hat{\lambda}} = N(N-1), \quad (4)$$

а число явних синаптичних зв'язків, які залишаються при приведенні матриці (2) до матриці Гессенберга (3), записується у вигляді

$$N_{\hat{\gamma}} = 3(N-1). \quad (5)$$

Таким чином, число явних синаптичних зв'язків нейронної мережі Гессенберга буде меншим у

$$\alpha = \frac{N_{\hat{\gamma}}}{N_{\hat{\lambda}}} = \frac{3}{N}. \quad (6)$$

разів.

Стан S_n ($n = 1, \dots, N$) n -го стохастичного нейрона, як і в моделі Хопфілда, є біполярним і приймає значення $S_n = \pm 1$. Перехід нейрона в стан $S_n = +1$ або $S_n = -1$ пов'язаний зі значенням потенціалу h_n не однозначно, а випадковим чином.

Імовірність переходу нейрона в стани ± 1 :

$$P_r(S_n(t+1) = 1) = f(h_n), \quad (7)$$

$$P_r(S_n(t+1) = -1) = f(-h_n) = 1 - f(h_n), \quad (8)$$

де $f(h_n) = 1 / (1 + \exp(-2\beta h_n))$ – функція розподілу Фермі; $0 < f(h_n) < 1$; $f(h_n) + f(-h_n) = 1$; $\beta = 1/T$, де T – це аналог температури, який використовують для керування рівнем шуму та ступенем невизначеності перемикачів. При цьому важливо відзначити, що T не описує фізичну температуру нейронної мережі (біологічної або штучної). Параметр T керує термальними флуктуаціями, які представляють ефект синаптичного шуму. У низькотемпературному інтервалі $T \rightarrow 0$ розподіл Фермі переходить в порогову функцію і поведінка мережі із стохастичних нейронів стає аналогічною поведінці мережі Хопфілда (без включення шуму), яка складена із звичайних біполярних нейронів.

Оскільки динаміка станів стохастичних нейронів є ймовірнісною, то потрібно цікавитись тільки середньою активністю, або очікуваними значеннями їх станів:

$$\langle S_n \rangle = (+1)\langle f(h_n) \rangle + (-1)\langle f(-h_n) \rangle. \quad (9)$$

В силу нелінійності функції розподілу Фермі усереднення зробимо у наближенні середнього поля [16]

$$\langle f(h_n) \rangle \cong f(\langle h_n \rangle). \quad (10)$$

У результаті отримаємо наступну замкнену систему рівнянь:

$$\langle S_n \rangle = (+1)f(\langle h_n(t) \rangle) - 1 \cdot (1 - f(\langle h_n(t) \rangle)) = 2f(\langle h_n(t) \rangle) - 1 = \tanh(\beta \langle h_n(t) \rangle). \quad (11)$$

де $\langle h_n(t) \rangle$ – локальне самоузгоджене поле, яке діє на псевдоспіновий нейрон S_n^z зі сторони решти псевдоспінових нейронів і знаходиться із рівняння релаксаційної динаміки [7]:

$$\frac{\partial \langle S_n^z(t) \rangle}{\partial t} = \langle h_n(t) \rangle, \quad (12)$$

$$\langle h_n(t) \rangle = -\frac{dE}{d\langle S_n^z(t) \rangle} = \sum_{\substack{k=1 \\ n \neq k}}^N \gamma_{nk} \langle S_k^z(t) \rangle + h_n^{ext}(t),$$

де
$$h_n^{ext}(t) = \sum_{\mu=1}^m b_{\mu n}(t) \xi_{\mu n} + V_{noise}^n(t). \quad (13)$$

Тут E – функція енергії штучної нейронної мережі взаємодіючих псевдоспінових нейронів, яка описується виразом, аналогічним до дипольної системи мікротрубочки цитоскелета нейрона [7]

$$E = -\frac{1}{2} \sum_{n=1}^N \sum_{\substack{k=1 \\ n \neq k}}^N \gamma_{nk} S_n^z S_k^z. \quad (14)$$

Мікротрубочки цитоскелета нейрона представляють собою інформаційно-білкові нанополімерні ланцюжки, які згорнуті в спіралі. Вони є найбільш прийнятними для нейронів мозку людини, що здійснюють «квантово-статистичні» обчислення [8, 17].

В залежності від знака елемента γ_{nk} матриці синаптичних зв'язків і характеру їх розподілу, штучні псевдоспінові нейронні мережі Гессенберга можна поділити на три типи:

– якщо всі елементи γ_{nk} матриці синаптичних зв'язків додатні ($\forall \gamma_{nk} > 0$), то в стані з мінімумом енергії $E = E_{\min}$ всі псевдоспінові нейрони мають однакові напрямки $S_n^z = +1$ (\uparrow) або $S_n^z = -1$ (\downarrow). При цьому енергія взаємодії будь-якої пари псевдоспінів $\delta E_{nk} = -\gamma_{nk} S_n^z S_k^z$ досягає свого мінімального значення $\delta E_{nk} = -\gamma_{nk}$. Такий стан нейронної мережі взаємодіючих псевдоспінових нейронів будемо називати «ферромагнітним» (рис. 3).

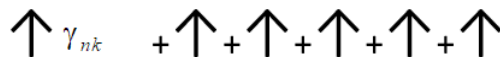


Рисунок 3 – «Ферромагнітний» стан нейронної мережі

– у випадку, коли всі елементи γ_{nk} матриці синаптичних зв'язків є від'ємні ($\forall \gamma_{nk} < 0$), то в стані з мінімумом енергії $E = E_{\min}$ половина псевдоспінових нейронів має один напрям $S_n^z = +1$ (\uparrow), а інша – протилежний $S_n^z = -1$ (\downarrow). У цьому випадку псевдоспінову нейронну мережу зручно поділити на дві псевдоспінові нейронні мережі з однаковими векторами псевдоспінових (дипольних) моментів. При цьому енергія взаємодії будь-якої пари псевдоспінових нейронів із різних поділених нейронних мереж ($S_n^z = +1, S_n^z = -1$)

$\delta E_{nk} = -\gamma_{nk} S_n^z S_k^z$ досягає свого мінімального значення $\delta E_{nk} = -\gamma_{nk}$. У цьому випадку псевдоспінова нейронна мережа називається «антиферромагнітною» (рис. 4).

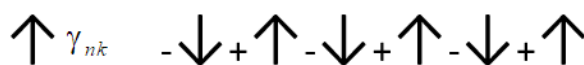


Рисунок 4 – «Антиферромагнітний» стан нейронної мережі

– якщо елементи матриці синаптичних зв'язків γ_{nk} між псевдоспіновими нейронами випадково приймають як додатні, так і від'ємні значення, то така псевдоспінова нейронна мережа має назву «дипольне скло» (рис. 5). Таке твердження випливає з аналізу структури твердого тіла з випадковим розміщенням дипольних моментів у вузлах кристалічної ґратки.

Відомо, що енергія фази «дипольного скла» в мікротрубочці цитоскелета володіє великою кількістю глобальних і локальних мінімумів в просторі конфігурацій дипольних моментів [7]. Тому навіть повністю неупорядкована дипольна нейромережа мікротрубочки цитоскелета (γ_{nk} розподілені випадково згідно

$$\gamma_n = \sum_{k=1}^z \gamma_{nk} = 0, I_n = \sum_{k=1}^z I_{nk} \neq 0) \text{ володіє пам'яттю.}$$

Оскільки γ_{nk} є випадковою величиною, то необхідно задати її статистичні властивості. Припускаємо, що випадкова величина γ_{nk} , задана на зв'язку між вузлами n та k , не залежить від того, які γ_{nk} знаходяться на інших зв'язках. У цьому випадку статистичні властивості γ_{nk} повністю визначаються функцією розподілу $f(\gamma_{nk})$ на даному зв'язку. Зазвичай припускають, що $f(\gamma_{nk})$ – Гаусова функція розподілу

$$f(\gamma_{nk}) = \frac{1}{I_{nk} \sqrt{2\pi}} \exp\left(-2\gamma_{nk}^2 I_{nk}^2\right).$$

У такому випадку вона задається двома параметрами: середнім значенням $\tilde{\gamma}_{nk}$ і дисперсією I_{nk} : $\langle \gamma_{nk} \rangle = \tilde{\gamma}_{nk}, I_{nk} = \langle \gamma_{nk}^2 \rangle - \langle \gamma_{nk} \rangle^2$.

Мінімальному значенню функції E (14) («атрактору») відповідає певна конфігурація дипольних моментів нейронів у нейромережі.

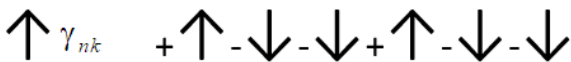


Рисунок 5 – Стан нейронної мережі «Дипольне скло»

Нехай необхідно записати в нейромережі Гессенберга (рис. 2) m різних образів-еталонів, кожен з яких характеризується своєю конфігурацією псевдоспінів $\{S_{\mu n}^z = \xi_{\mu n}\}$. Причому, різні конфігурації ортогональні, тобто

$$\frac{1}{N} \sum_{m=1}^N \xi_{\mu n} \xi_{\mu' n} = \delta_{\mu \mu'} \quad (15)$$

Інформаційний сигнал для запису (запам'ятовування) – це послідовність біт $\xi_1, \xi_2, \dots, \xi_\mu, \dots, \xi_{mN}$. Елемент ξ_μ з рівною ймовірністю приймає значення ± 1 . Тут довжина послідовності (mN) вибрана кратною числу нейронів. Перепозначимо елементи послідовності:

$$\begin{aligned} &\xi_{11}, \xi_{21}, \xi_{31}, \dots, \xi_{\mu 1}, \dots, \xi_{m1}, \\ &\xi_{12}, \xi_{22}, \xi_{32}, \dots, \xi_{\mu 2}, \dots, \xi_{m2}, \\ &\xi_{13}, \xi_{23}, \xi_{33}, \dots, \xi_{\mu 3}, \dots, \xi_{m3}, \\ &\dots\dots\dots \\ &\xi_{1N}, \xi_{2N}, \xi_{3N}, \dots, \xi_{\mu N}, \dots, \xi_{mN}. \end{aligned} \quad (16)$$

Таким чином, кожному нейрону $n = 1, \dots, N$ буде відповідати інформаційний набір $\{\xi_{\mu n}, \mu = 1, 2, 3, \dots, m\}$. Або при фіксованому μ набір $\{\xi_{\mu n}, n = 1, 2, 3, \dots, N\}$ представляє собою інформаційну копію мережі, тобто «репліку».

Для розрідження синаптичних зв'язків між сенсорними вхідними образами та обчислювальними нейронами зведемо матрицю синаптичних зв'язків між вхідними сенсорними образами та псевдоспіновими нейронами у нейронній мережі Гессенберга ($\hat{b}_{\mu n}$) до діагонального вигляду [14] і до симетричної форми. Для цього зробимо лінійне перетворення

$$(\hat{b}_{\mu n}) = \hat{U}^{-1} (b_{\mu n}) \hat{U}, \quad (17)$$

де \hat{U} – матриця, яка складається з власних базисних векторів \vec{u}_μ матриці $(\hat{b}_{\mu n})$, тобто $\hat{U} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_\mu, \dots, \vec{u}_N)$;

$$(\hat{b}_{\mu n}) \vec{u}_\mu = \beta_\mu \vec{u}_\mu. \quad (18)$$

У базисі з власних векторів \vec{u}_μ матриця лінійного перетворення $(\hat{b}_{\mu n})$ має діагональний вигляд, причому на головній діагоналі розташовані дійсні власні числа $\beta_\mu = \beta_n \delta_{\mu n}$ матриці $(\hat{b}_{\mu n})$.

Запис (запам'ятовування) інформації в нейромережі відбувається в результаті формування матриці синаптичних зв'язків $\hat{\gamma}$ між псевдоспіновими нейронами за допомогою перенормованого локального правила Хебба:

$$\gamma_{nk} = F \left(\lambda_{ij} = \sum_{\mu=1}^m \xi_{\mu i} \xi_{\mu j} \right). \quad (19)$$

Тоді із (11), з урахуванням (12) та (13), випливає, що середня активність станів нейронів (тобто середнє значення стану n -го нейрона) описується виразом

$$\langle S_n \rangle = \tanh \left[\beta \left(\sum_{k=1}^N \sum_{n \neq k} \gamma_{nk} \langle S_k^z(t) \rangle + \sum_{\mu=1}^m b_{\mu n}(t) \xi_{\mu n} + V_{noise}^n(t) \right) \right]$$

Для моделювання процесу еволюції системи псевдоспінів з дискретним часом на комп'ютерах нейромережа повинна здійснювати перетворення вхідної конфігурації $\langle S_{\mu n}^z \rangle = \xi_{\mu n}$ в дискретному часі так, щоб вихідна конфігурація $\langle S_{\mu n}^z \rangle$ була близькою до тієї картини-еталону, яка є правильною відповіддю, тобто $\langle S_{\mu n}^z(t+1) \rangle$ прийматиме значення

$$\operatorname{sgn} \left[\sum_{\substack{k=1 \\ n \neq k}}^N \gamma_{nk} \langle S_k^z(t) \rangle + \sum_{\mu=1}^m b_{\mu n}(t) \xi_{\mu n} + V_{noise}^n(t) \right], \quad (20)$$

де

$$\operatorname{sgn} \left[\sum_{\substack{k=1 \\ n \neq k}}^N \gamma_{nk} \langle S_k^z(t) \rangle + \sum_{\mu=1}^m b_{\mu n}(t) \xi_{\mu n} + V_{noise}^n(t) \right] = \begin{cases} 1, & \text{якщо } \sum_{\substack{k=1 \\ n \neq k}}^N \gamma_{nk} \langle S_k^z(t) \rangle + \sum_{\mu=1}^m b_{\mu n}(t) \xi_{\mu n} + V_{noise}^n(t) \geq 0, \\ -1, & \text{якщо } \sum_{\substack{k=1 \\ n \neq k}}^N \gamma_{nk} \langle S_k^z(t) \rangle + \sum_{\mu=1}^m b_{\mu n}(t) \xi_{\mu n} + V_{noise}^n(t) < 0. \end{cases}$$

Псевдоспінова нейромережа із законом еволюції (20) володіє пам'яттю, яка зберігає деякий заданий набір образів-еталонів, що намагається згадати один із них, якщо їй пред'являється який-небудь із цих образів, спотворений шумами.

У результаті завершення процесу еволюції в системі псевдоспінів вхідний образ асоціюється з одним із образів-еталонів, що запам'ятався раніше, і можна стверджувати, що нейромережа виступає в ролі розподіленої структури з асоціативною пам'яттю.

Відновлення інформації, що запам'яталась, здійснюється в процесі стохастичної паралельної динаміки з дискретним часом $t = 0, 1, 2, 3, \dots$

$$\langle S_{\mu n}^z(t+1) \rangle = \begin{cases} 1 & \text{із ймовірністю } p = f(h_n), \\ -1 & \text{із ймовірністю } q = 1 - p. \end{cases} \quad (21)$$

4 ЕКСПЕРИМЕНТИ

Час налаштування синаптичних зв'язків λ_{ij} (за одну ітерацію) у повнозв'язній нейронній мережі пропорційний числу синаптичних зв'язків N_λ , тобто

$$t_\lambda = k_\lambda N_\lambda. \quad (22)$$

Аналогічно, час налаштування синаптичних зв'язків γ_{nk} (за одну ітерацію) у нейронній мережі Гессенберга дорівнює

$$t_\gamma = k_\gamma N_\gamma. \quad (23)$$

Коефіцієнти k_λ , k_γ залежать від архітектури синаптичних зв'язків в нейронних мережах.

З урахуванням формул (4) та (5) отримаємо $t_\gamma / t_\lambda = 3k_\gamma(N-1) / k_\lambda N(N-1) = \alpha v$, де параметр α виражається формулою (6), а $v = k_\gamma / k_\lambda$.

Отже визначимо залежність відношення t_γ / t_λ від кількості стохастичних псевдоспінових нейронів N при $v = 1, 2, 3$.

Комп'ютерний експеримент реалізований у системі комп'ютерної математики «Wolfram Mathematica 11.3».

5 РЕЗУЛЬТАТИ

На рис. 6–8 наведені числові результати розрахунку залежності відношення t_γ / t_λ від кількості стохастичних псевдоспінових нейронів N при $v = 1, 2, 3$.

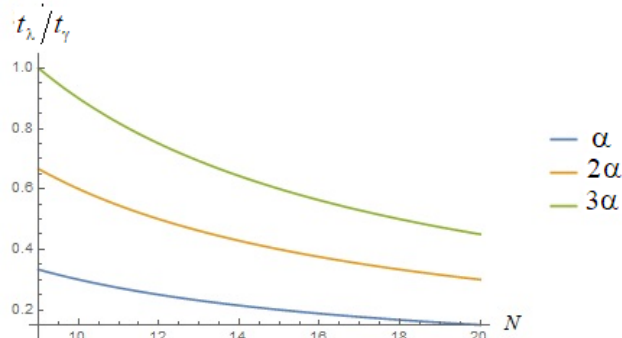


Рисунок 6 – Графік залежності відношення t_γ / t_λ від кількості стохастичних псевдоспінових нейронів при $N_{\max} = 20$

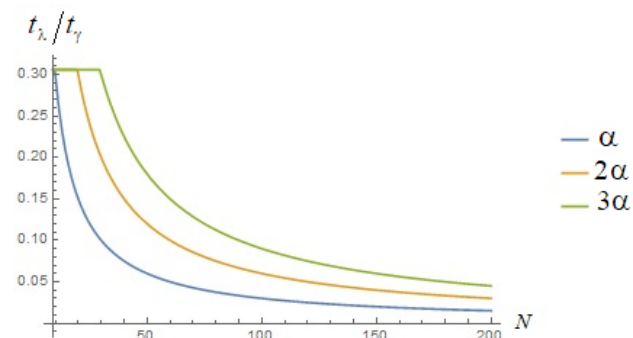
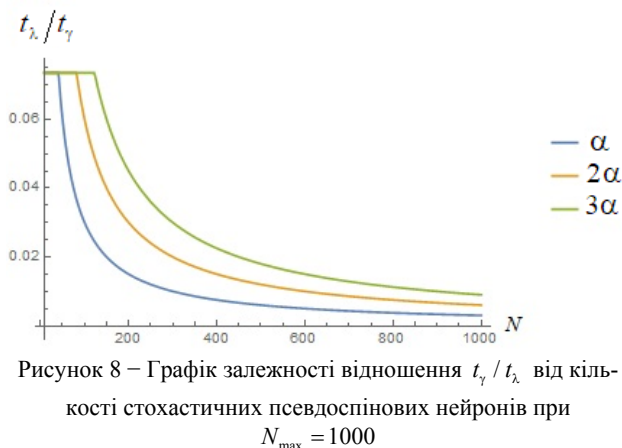


Рисунок 7 – Графік залежності відношення t_γ / t_λ від кількості стохастичних псевдоспінових нейронів при $N_{\max} = 200$



6 ОБГОВОРЕННЯ

Як видно з рис. 6–8 відношення t_y / t_x зменшується за гіперболічним законом при зростанні числа нейронів, отже час використаний для налаштування тридіагональної нейронної мережі відносно часу налаштування повнозв'язної нейронної мережі, буде зменшуватись зі збільшенням числа нейронів N , тобто для роботи тридіагональної нейронної мережі буде використана значно менша кількість обчислювального ресурсу, оскільки число синаптичних зв'язків буде

меншим у $\alpha = \frac{N_y}{N_x} = \frac{3}{N}$ разів. Зі збільшенням параметра α при однаковому значенні числа псевдоспінових нейронів N , відносний час t_y / t_x буде збільшуватись.

Зокрема при $N = 200$, при $\alpha = 1$, $t_y / t_x = 1,5\%$, тоді як при $\alpha = 3$, $t_y / t_x = 4,5\%$.

Порівняно з методами [1, 4–13] запропонований нами метод зменшення числа синаптичних зв'язків враховує повну інформацію повнозв'язної нейромережі, оскільки недиагональні компоненти повнозв'язної матриці вагових коефіцієнтів синаптичних зв'язків містяться у головних компонентах тридіагональної матриці синаптичних зв'язків. Крім цього, цей метод вимагає менший обчислювальний ресурс порівняно з методом генетичного алгоритму.

ВИСНОВКИ

Наукова новизна отриманих результатів полягає в тому, що вперше:

- розроблено модель архітектури стохастичної нейромережі з розрідженими синаптичними зв'язками, яка враховує взаємодію між всіма псевдоспіновими нейронами у перенормованих зв'язках між найближчими сусідніми псевдоспіновими нейронами, що описуються тридіагональною матрицею синаптичних зв'язків Гессенберга;

- запропоновано спосіб розрідження синаптичних зв'язків між сенсорними та обчислювальними нейронами на основі діагоналізації повнозв'язної матриці синаптичних зв'язків, що приводить до ефекту змен-

шення часу налаштування вагових коефіцієнтів синаптичних зв'язків у нейронній мережі;

- запропоновано перенормоване локальне правило Хебба на основі перетворення повнозв'язної матриці синаптичних зв'язків між нейронами до тридіагональної матриці перенормованих синаптичних зв'язків (матриці Гессенберга);

- встановлено, що зі збільшенням числа нейронів час налаштування синаптичних зв'язків (за одну ітерацію) у стохастичній псевдоспіновій нейронній мережі з тридіагональною матрицею зв'язків, відносно часу налаштування синаптичних зв'язків (за одну ітерацію) у повнозв'язній синаптичній нейронній мережі, зменшується за гіперболічним законом;

- запропоновано класифікацію нейронної мережі Гессенберга із феромагнітною, антиферомагнітною структурами та дипольним склом в залежності від напрямку псевдоспінів нейронів.

Результати дослідження можуть бути безпосередньо **практично використані** для:

- зменшення обчислювальних складностей за рахунок діагоналізації матриці синаптичних зв'язків (відсутність недиагональних компонент синаптичних зв'язків) між сенсорними і обчислювальними нейронами без втрати інформації, що міститься в недиагональних компонентах і зберігається в головних перенормованих тридіагональних компонентах матриці синаптичних зв'язків;

- шифрування інформації з підвищеним ступенем криптостійкості за рахунок перенормування повнозв'язної матриці синаптичних зв'язків до тридіагональної, оскільки ключ шифрування в тридіагональній нейромережі має більш складну будову;

- підвищення узагальненої функціональної здатності нейронної мережі;

- розпізнавання промоторів ДНК, оскільки аналогом гомологічного індексу, який визначає близькість реальних промоторів ДНК до консенсус послідовності є енергія стану нейромережі.

Перспективи подальших досліджень полягають у тому, щоб дослідити ефективність запропонованого методу при розпізнаванні образів, шифруванні інформації та апроксимації складних нелінійних функцій.

ЛІТЕРАТУРА / LITERATURA

1. Haykin S. Neural Networks and Learning Machines / S. Haykin. – Hamilton: Pearson, 2008. – 936 p.
2. Datta B. N. Numerical Linear Algebra and Applications / B. N. Datta. – Illinois: Second Ed. Society for Industrial and Applied Mathematics (SIAM), 2010. – 554 p. DOI: 10.1137/1.9780898717655
3. Richard L. Numerical Analysis / L. Richard, J. Burden, F. Douglas. – Boston : Cengage Learning, 2010. – 888 p.
4. Корнієнко О. В. Згорткова нейромережа для виявлення шахрайських операцій з кредитними картками / О. В. Корнієнко, С. О. Субботін // Automation of technological and business processes. – 2019. – Том. 11, № 3. – С. 65–74. DOI: 10.15673/atbp.v11i3.1503
5. Ismoilov N. A Comparison of Regularization Techniques in Deep Neural Networks / N. Ismoilov, S. Jang // Symmetry. – 2018. – Vol. 10, Issue 11. – P. 648. DOI:10.3390/sym10110648

6. Deep Neural Networks Regularization for Structured Output Prediction / [S. Belharbi, R. Hérault, C. Chatelain et al.] // *Neurocomputing*. – 2018. – Vol. 281. – P. 169–177. DOI: 10.1016/j.neucom.2017.12.002
7. Слядников Е. Е. Физическая модель и ассоциативная память дипольной системы микротрубочки цитоскелета / Е. Е. Слядников // *Журнал технической физики*. – 2007. – Том. 77, № 7. – С. 77–86.
8. Penrose R. *Shadows of the Mind: A Search for the Missing Science of Consciousness* / R. Penrose. – Oxford : Oxford University Press, 1996. – 480 p.
9. Weigened A. S. Generalization by weight-elimination with application to forecasting / A. S. Weigened, D. E. Rumelhart, B. A. Huberman // *Advances in neural information processing systems*. – 1991. – Vol. 3. – P. 875–882. DOI: 10.1109/ijcnn.1991.155287
10. A structure optimization algorithm of neural networks for large-scale data sets / [J. Yang, J. Ma, M. Berryman et al.] // *Fuzzy Systems: First International Conference, China, 6–11 July 2014: proceedings*. – Beijing : IEEE, 2014. – P. 956–961. DOI: 10.1109/fuzz-ieee.2014.6891662
11. *Neural Architecture Optimization* / [R. Luo, F. Tian, T. Qin et al.] // *Neural Information Processing Systems: Thirty-second Conference, Canada, 3 – 8 December 2018: proceedings*. – Montreal: NIPS, 2018. – P. 1–12.
12. Genetic Algorithm Based Deep Learning Neural Network Structure and Hyperparameter Optimization / [S. Lee, J. Kim, H. Kang et al.] // *Appl. Sci.* – 2021. – Vol. 11, Issue 2. – P. 744. DOI: 10.3390/app11020744
13. Wu W. Neural network structure optimization based on improved genetic algorithm / W. Wu // *International Conference on Advanced Computational Intelligence: Fifth International Conference, China, 18–20 October 2012: proceedings*. – Nanjing: ICACI, 2012. – P. 893–895. DOI: 10.1109/ICACI.2012.6463299
14. Lytvyn V. The compression of the input images in neural network that using method diagonalization the matrices of synaptic weight connections / V. Lytvyn, I. Peleshchak, R. Peleshchak // *Conference on Advanced Information and Communication Technologies: Second International Conference, Ukraine, 4–7 July 2017: proceedings*. – Lviv: AICT, 2017. – P. 66–70. DOI: 10.1109/AIACI.2017.8020067
15. Attaway S. *Matlab: A Practical Introduction to Programming and Problem Solving* / S. Attaway. – Oxford : Butterworth-Heinemann, 2019. – 626 p. DOI: 10.1016/b978-0-12-815479-3.00003-9
16. Ежов А. Нейрокомпьютинг и его применения в экономике и бизнесе / А. Ежов, А. Шумский. – Москва : МИФИ, 1998. – 224 с.
17. Hameroff S. R. Quantum coherence in microtubules: A neural basis for emergent consciousness / S. R. Hameroff // *Journal of consciousness studies*. – 1994. – Vol. 1, Issue 1. – P. 91–118.

Received 08.04.2021.

Accepted 21.05.2021.

УДК 004.8, 004.032.26

СТОХАСТИЧЕСКАЯ ПСЕВДОСПИНОВАЯ НЕЙРОННАЯ СЕТЬ С ТРЕХДИАГОНАЛЬНЫМИ СИНАПТИЧЕСКИМИ СВЯЗЯМИ

Пелешак Р. М. – д-р физ.-мат. наук, профессор кафедры информационных систем и сетей, Национальный университет «Львовская политехника», Украина.

Литвин В. В. – д-р техн. наук, профессор, заведующий кафедрой информационных систем и сетей, Национальный университет «Львовская политехника», Украина.

Черняк О. И. – аспирантка кафедры математики, Дрогобычский государственный педагогический университет имени Ивана Франко, Дрогобыч, Украина.

Пелешак И. Р. – аспирант кафедры информационных систем и сетей, Национальный университет «Львовская политехника», Украина.

Дорошенко Н. В. – канд. физ.-мат. наук, доцент кафедры информатики и информационных систем, Дрогобычский государственный педагогический университет имени Ивана Франко, Дрогобыч, Украина.

АННОТАЦИЯ

Актуальность. Для сокращения времени вычислительного ресурса в задачах диагностирования и распознавания искаженных образов на основе полносвязной стохастической псевдоспиновой нейронной сети возникает необходимость прореживания синаптических связей между нейронами, которое решается с помощью метода диагонализации матрицы синаптических связей без потери взаимодействия между всеми нейронами в сети.

Цель работы. Создание архитектуры стохастической псевдоспиновой нейросети с разреженными и диагональными синаптическими связями без потери взаимодействия между всеми нейронами в слое для уменьшения времени ее обучения.

Метод. В статье используется метод Хаусхолдера, метод сжатия входных образов на основе диагонализации матрицы синаптических связей и система компьютерной математики MATLAB для приведения полносвязной нейросети к трехдиагональному виду со скрытыми синаптическими связями между всеми нейронами.

Результаты. Разработана модель архитектуры стохастической нейросети с разреженными перенормированными синаптическими связями, которые учитывают изъятые синаптические связи. На основе преобразования матрицы синаптических связей полносвязной нейронной сети к матрице Гессенберга с трехдиагональными синаптическими связями предложено перенормированное локальное правило Хебба. С помощью системы компьютерной математики «WolframMathematica 11.3» рассчитано в зависимости от числа нейронов N относительное время настройки синаптических связей (за одну итерацию) в стохастической псевдоспиновой нейронной сети с трехдиагональной матрицей связей, относительно времени настройки синаптических связей (за одну итерацию) в полносвязной синаптической нейронной сети.

Выводы. Установлено, что с увеличением числа нейронов время настройки синаптических связей (за одну итерацию) в стохастической псевдоспиновой нейронной сети с трехдиагональной матрицей связей, относительно времени настройки синаптических связей (за одну итерацию) в полносвязной синаптической нейронной сети, уменьшается за гиперболическим законом. В зависимости от направления псевдоспинов нейронов, предложена классификация перенормированной нейронной сети с ферромагнитной структурой, антиферромагнитной структурой и дипольным стеклом.

КЛЮЧЕВЫЕ СЛОВА: нейронная сеть, синаптические связи, матрица связей, тридиагонализация.

UDC 004.8, 004.032.26

STOCHASTIC PSEUDOSPIN NEURAL NETWORK WITH TRIDIAGONAL SYNAPTIC CONNECTIONS

Peleshchak R. M. – Doctor of Physical and Mathematical Sciences, Professor of the Department of Information Systems and Networks, Lviv Polytechnic National University, Lviv, Ukraine.

Lytvyn V. V. – Dr. Sc., Professor, Head of the Department of Information Systems and Networks, Lviv Polytechnic National University, Ukraine.

Cherniak O. I. – Postgraduate student of the Department Of Mathematics, Drohobych Ivan Franko State Pedagogical University, Drohobych, Ukraine.

Peleshchak I. R. – Postgraduate student of the Department Of Information Systems And Networks, Lviv Polytechnic National University, Lviv, Ukraine.

Doroshenko M. V. – PhD, Associate Professor of the Department of informatics and information systems, Drohobych Ivan Franko State Pedagogical University, Drohobych, Ukraine.

ABSTRACT

Context. To reduce the computational resource time in the problems of diagnosing and recognizing distorted images based on a fully connected stochastic pseudospin neural network, it becomes necessary to thin out synaptic connections between neurons, which is solved using the method of diagonalizing the matrix of synaptic connections without losing interaction between all neurons in the network.

Objective. To create an architecture of a stochastic pseudo-spin neural network with diagonal synaptic connections without losing the interaction between all the neurons in the layer to reduce its learning time.

Method. The paper uses the Householder method, the method of compressing input images based on the diagonalization of the matrix of synaptic connections and the computer mathematics system MATLAB for converting a fully connected neural network into a tridiagonal form with hidden synaptic connections between all neurons.

Results. We developed a model of a stochastic neural network architecture with sparse renormalized synaptic connections that take into account deleted synaptic connections. Based on the transformation of the synaptic connection matrix of a fully connected neural network into a Hessenberg matrix with tridiagonal synaptic connections, we proposed a renormalized local Hebb rule. Using the computer mathematics system “WolframMathematica 11.3”, we calculated, as a function of the number of neurons N , the relative tuning time of synaptic connections (per iteration) in a stochastic pseudospin neural network with a tridiagonal connection Matrix, relative to the tuning time of synaptic connections (per iteration) in a fully connected synaptic neural network.

Conclusions. We found that with an increase in the number of neurons, the tuning time of synaptic connections (per iteration) in a stochastic pseudospin neural network with a tridiagonal connection Matrix, relative to the tuning time of synaptic connections (per iteration) in a fully connected synaptic neural network, decreases according to a hyperbolic law. Depending on the direction of pseudospin neurons, we proposed a classification of a renormalized neural network with a ferromagnetic structure, an antiferromagnetic structure, and a dipole glass.

KEYWORDS: neural network, synaptic connections, connection matrix, tridiagonalization.

REFERENCES

1. Haykin S. *Neural Networks and Learning Machines*. Hamilton, Pearson, 2008, 936 p.
2. Datta B. N. *Numerical Linear Algebra and Applications*, Illinois: Second Ed. Society for Industrial and Applied Mathematics (SIAM), 2010, 554 p. DOI: 10.1137/1.9780898717655
3. Richard L., Burden J., Douglas F. *Numerical Analysis*. Boston, Cengage Learning, 2010, 888 p.
4. Korniienko O. V., Subbotin S. O. Zghorkova neiromerezhha dlia vyavleniia shakhrayskykh operatsii z kredytnymy kartkamy, *Automation of technological and business processes*, 2019, Vol. 11, Issue 3, pp. 65–74. DOI: 10.15673/atbp.v11i3.1503
5. Ismoilov N., Jang S. A Comparison of Regularization Techniques in Deep Neural Networks, *Symmetry*, 2018, Vol. 10, Issue 11, P. 648. DOI: 10.3390/sym10110648
6. Belharbi S., Herault R., Chatelain C. et al. Deep Neural Networks Regularization for Structured Output Prediction, *Neurocomputing*, 2018, Vol. 281, pp. 169–177. DOI: 10.1016/j.neucom.2017.12.002
7. Slyadnikov E. E. Fizicheskaya model i assotsiativnaya pamyat dipolnoy sistemyi mikrotrubochki tsitoskeleta, *Zhurnal tehnicheskoy fiziki*, 2007, Vol. 77, Issue 7, pp. 77–86.
8. Penrose R. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford, Oxford University Press, 1996, 480 p.
9. Weigened A. S., Rumelhart D. E., Huberman B. A. Generalization by weight-elimination with application to forecasting, *Advances in neural information processing systems*, 1991, Vol. 3, pp. 875–882. DOI: 10.1109/ijcnn.1991.155287
10. Yang J., Ma J., Beryman M. et al. A structure optimization algorithm of neural networks for large-scale data sets, *Fuzzy Systems: First International Conference, China, 6–11 July 2014: proceedings*. Beijing, IEEE, 2014, pp. 956–961. DOI: 10.1109/fuzz-ieee.2014.6891662
11. Luo R., Tian F., Qin T. et al. Neural Architecture Optimization, *Neural Information Processing Systems: Thirty-second Conference, Canada, 3–8 December 2018: proceedings*. Montreal, NIPS, 2018, pp. 1–12.
12. Lee S., Kim J., Kang H. et al. Genetic Algorithm Based Deep Learning Neural Network Structure and Hyperparameter Optimization, *Appl. Sci*, 2021, Vol. 11, Issue 2, P. 744. DOI: 10.3390/app11020744
13. Wu W. Neural network structure optimization based on improved genetic algorithm, *International Conference on Advanced Computational Intelligence: Fifth International Conference, China, 18–20 October 2012: proceedings*. Nanjing: ICACI, 2012, pp. 893–895. DOI: 10.1109/ICACI.2012.6463299
14. Lytvyn V., Peleshchak I., Peleshchak R. The compression of the input images in neural network that using method diagonalization the matrices of synaptic weight connections, *Conference on Advanced Information and Communication Technologies: Second International Conference, Ukraine, 4–7 July 2017, proceedings*. Lviv, AICT, 2017, pp. 66–70. DOI: 10.1109/AICT.2017.8020067
15. Attaway S. *Matlab: A Practical Introduction to Programming and Problem Solving*. Oxford, Butterworth-Heinemann, 2019, 626 p. DOI: 10.1016/b978-0-12-815479-3.00003-9
16. Ejov A., Shumskiy A. Neyrokompjuting i ego primeneniya v ekonomike i biznese. Moscow, MIFI, 1998, 224 p.
17. Hameroff S. R. Quantum coherence in microtubules: A neural basis for emergent consciousness, *Journal of consciousness studies*, 1994, Vol. 1, Issue 1, pp. 91–118.

© Пелешак Р. М., Литвин В. В., Черняк О. І., Пелешак І. Р., Дорошенко М. В., 2021
DOI DOI 10.15588/1607-3274-2021-2-12

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

PROGRESSIVE INFORMATION TECHNOLOGIES

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

UDC 004.93

AUTOMATED PANSHARPENING INFORMATION TECHNOLOGY OF SATELLITE IMAGES

Hnatushenko V. V. – Dr. Sc., Professor, Head of the Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

Kashtan V. Yu. – PhD, Associate Professor of the Department of Information Technology and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine.

ABSTRACT

Context. Nowadays, information technologies are widely used in digital image processing. The task of joint processing of satellite image obtained by different space systems that have different spatial differences is important. The already known pansharpening methods to improve the quality of the resulting image, there are new scientific problems associated with increasing the requirements for high-resolution image processing and the development of automated technology for processing the satellite data for further thematic analysis. Most spatial resolution techniques result in artifacts. Our work explores the major remote sensing data fusion techniques at pixel level and reviews the concept, principals, limitations and advantages for each technique with the program implementation of research.

Objective. The goal of the work is analyze the effectiveness of the traditional pan-sharpening methods like the Brovey, the wavelet-transform, the GIHS, the HCT and the combined pansharpening method for satellite images of high-spatial resolution.

Method. In this paper we propose the information technology for pansharpening high spatial resolution images with automation of choosing the best method based on the analysis of quantitative and qualitative evolutions. The method involves the scaling multispectral image to the size of the panchromatic image; using histogram equalization to adjust the primary images by aligning the integral areas of the sections with different brightness; conversion of primary images after the spectral correction on traditional pansharpening methods; analyze the effectiveness of the results obtained for conducts their detailed comparative visual and quantitative evaluation. The technology allows determining the best method of pansharpening by analyzing quantitative metrics: the NDVI index, the RMSE and the ERGAS. The NDVI index for the methods Brovey and HPF indicate color distortion in comparison with the reference data. This is due to the fact that the Brovey and HPF methods are based on the fusion of three channel images and do not include the information contained in the near infrared range. The RMSE and the ERGAS show the superiority of the combined HSV-HCT-Wavelet method over the conventional and state-of-art image resolution enhancement techniques of high resolution satellite images. This is achieved, in particular, by preliminary processing of primary images, data processing localized spectral bases, optimized performance information, and the information contained in the infrared image.

Results. The software implementing proposed method is developed. The experiments to study the properties of the proposed algorithm are conducted. Experimental evaluation performed on eight-primary satellite images of high spatial resolution obtained WorldView-2 satellite. The experimental results show that a synthesized high spatial resolution image with high information content is achieved with the complex use of fusion methods, which makes it possible to increase the spatial resolution of the original multichannel image without color distortions.

Conclusions. The experiments confirmed the effectiveness of the proposed automated information technology for pansharpening high-resolution satellite images with the development of a graphical interface to obtain a new synthesized image. In addition, the proposed technology will effectively carry out further recognition and real-time monitoring infrastructure.

KEYWORDS: image fusion, satellite image, pansharpening, automating, WorldView, high resolution imaging, metanalysis.

ABBREVIATIONS

RGB is a multispectral image;

PAN is a panchromatic image;

NIR is a Near Infrared band;

GIHS is a generalized intensity-saturation-tone;

HPF is a high-frequency filter;

HCT is a hyperspherical conversion;

NDVI is a normalized vegetation index;

GUI is a graphical user interface.

NOMENCLATURE

F_i is an i -th channel in the image after the pansharpening;
 P is a panchromatic image;
 RGB_i^\uparrow is an i -th image channel RGB after scaling to size P ;
 S is a synthesized image;
 N is a total number of RGB channel;
 $*$ is a convolution operator;
 g is a high-frequency filter;
 P' is a detail coefficients of the software;
 NIR is a near infrared light intensity value;
 RED is a red reflection, which corresponds respectively to channel 8 for images obtained by satellite WorldView-2.

INTRODUCTION

In recent years satellites imaging systems have been developing rapidly. Nowadays, these satellites allow to obtain data with a spatial resolution of half a meter or less to monitoring the state of forests, sea areas, shelves, etc. Images of high spatial resolution are required to use of special methods of their processing [1].

Satellite imagery of high spatial resolutions is a source of geospatial information. One of the problems of using satellite images obtained by various space systems is the joint processing of a multispectral image with low spatial resolution (channels: Coastal, Blue, Green, Yellow, Red, Red Edge, NIR1, NIR2) and a panchromatic image (PAN) of high spatial resolution, which is monochrome. The PANs have a smaller pixel size and are fixed in a sufficiently wide spectral range of radiation intensity and have a significantly higher spatial disparity than the RGBs obtained in narrow spectral intervals [2]. The current level of requirements for the reliability of image interpretation makes it necessary to overcome the physical limitations of available satellite sensors, in particular, by fusion high spatial and spectral disparity indicators in one graphical object.

1 PROBLEM STATEMENT

The Data Fusion approach means that the resulting synthesized data combines the properties of both types of source data and carries more information than a simple combination of information sources considered separately from each other. These information sources contain heterogeneous data represented by different formats, structures and implemented on different types of platforms.

Eight-channel images from the WorldView-2 satellite are used as input data. The panchrome wavelength of this satellite was expanded from the visible to the near infrared range. An image can be expressed in matrix format by the formula (1):

$$A_{n,b} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{b,1} & \cdots & a_{b,n} \end{pmatrix}, \quad (1)$$

where n represents the number of the pixels and b is the number of bands. Considering each WorldView-2 band as a vector, the above matrix can be simplified by the formula (2):

$$A_b = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_8 \end{pmatrix}. \quad (2)$$

The main requirements to satellite images for their processing are as follows:

1. PAN and RGB images should be of the same territory and taken in the same season.
2. Images should be done with the imaging systems of the same type and the spectral bands of the imaging systems are the same.
3. PAN and RGB images should be of the same or close resolution, otherwise the images will contain different information about objects.

Therefore, there is a need to investigate the effectiveness of existing techniques for pansharpening high spatial difference satellite imagery with automation of choosing the best method based on the analysis of quantitative and qualitative evolutions for synthesized images.

2 REVIEW OF THE LITERATURE

Over the past two decades, a large number of methods for pansharpening satellite imagery have been developed, which can be divided into: component replacement, multiscale analysis and combined methods [3–6]. Component replacement methods are the conversion of a RGB with the replacement of some components from the software and the conversion of the result to the primary view. IHS (intensity – brightness, hue – color tone, saturation – saturation) is a method in which the I-component of the RGB is replaced with software with high. This transformation can be applied to only three images; and for images with four channels, a modified GIHS method [7, 8] is used. Rahmani et al. proposed a modified IHS pansharpening method which provides good quality visual high resolution multispectral image [9]. This method allows an image adaptive coefficient for IHS was found. It obtained a more accurate spectral resolution for primary image. Work [10] proposed a pansharpening with multiscale normalized nonlocal means filter. This filter computes each pixel value as the weighted average of all pixels over a sliding window, and the final smooth fused image is obtained. All these approaches are fast and easy to implement but not efficient for the latest image generation.

Efficient algorithm is a multispectral image conversion algorithm that uses independent component analysis (PCA) method, in which the first main component (RS1) is replaced with PAN image [11]. Brovey transform and Gram-Schmidt conversions are also common methods for replacing components. They make it possible to obtain a

new synthesized image with good visual quality, but disrupt spectral quality [12, 13]. The main disadvantage of component replacement methods is the appearance of spectral distortions. Multi-scale analysis methods provide images that will have a high spatial resolution (due to the software), a high spectral resolution (due to the RGB) and contain information that is not present in the primary data (due to images obtained using microwave shooting and active location). An example of such methods is the Laplace pyramid method [14] and the A-transform method [15]. In [16] authors suggested method which combines the Independent Component Analysis and optimization wavelet transform. The wavelet decomposition coefficients are determined as a result of solving the optimization problem by the criterion of maximizing the entropy. Wavelet transformations allow you to save spectral information of the image, but do not take into account a priori and expert information.

3 MATERIALS AND METHODS

The work developed an automated technology for comparing and the choosing the best method of pansharpening satellite images of high spatial. Satellite imagery obtained by satellite WorldView-2 has a spatial resolution of 1.84 m in the multispectral channel and 46 cm in the panchromatic channel, respectively. The scheme of the proposed algorithm is presented in Fig. 1 and consists of the following main stages:

1. Loading of primary input data different spatial resolution images of the same scene.
2. Preprocessing the snapshot. In this step, the RGB was scaled to the size of the panchromatic image, and the histogram was equalized to the multispectral image. This made it possible to adjust the primary images by aligning the integral areas of the sections with different brightness proposed [17] by the algorithm in operation (3):

$$v(x,y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} \cdot P_{ij} . \quad (3)$$

3. Analysis of pansharpening methods. In the proposed algorithm, the following methods are analyzed at

this stage: generalized intensity – saturation – tone, the “Brovey” method, the high-pass filter method, wavelet transformations and combined methods. These techniques are easy to process and can be applied with any number of selected input spectral channels.

3.1 The generalized intensity-saturation-tone (GIHS) method is a modified IHS method in which the number of channels can be selected. Channel I contains information about bright component of image, others – information about spectral components. The method can be represented according to [18] the following expressions (4):

$$F_i = RGB_{i\uparrow} + (P - S), \quad S = \frac{1}{N} \sum_{i=1}^N RGB_{i\uparrow} . \quad (4)$$

3.2 Method “Brovey” [19]. This method is based on formation of new spectral components of software in accordance with the following expression (5):

$$F_i = \frac{P}{S} RGB_{i\uparrow}, \quad S = \frac{1}{N} \sum_{i=1}^N RGB_{i\uparrow} . \quad (5)$$

3.3 High-frequency filter (HPF) method [20] increases image information content based on transfer of high-frequency components of panchromatic channel to spectral channels in accordance with the following expression (6):

$$F_i = \frac{[P * g] + RGB_{i\uparrow}}{2} . \quad (6)$$

3.4 Hyperspherical conversion (HCT) is a transformation between a native color space with Cartesian coordinates into a hyperspherical color space that effectively works with any number of RGB input channels [21] in the sample (7):

$$F_i^2 = \sum_{i=1}^N RGB_i^2, \quad P^2 = (Pan)^2 . \quad (7)$$

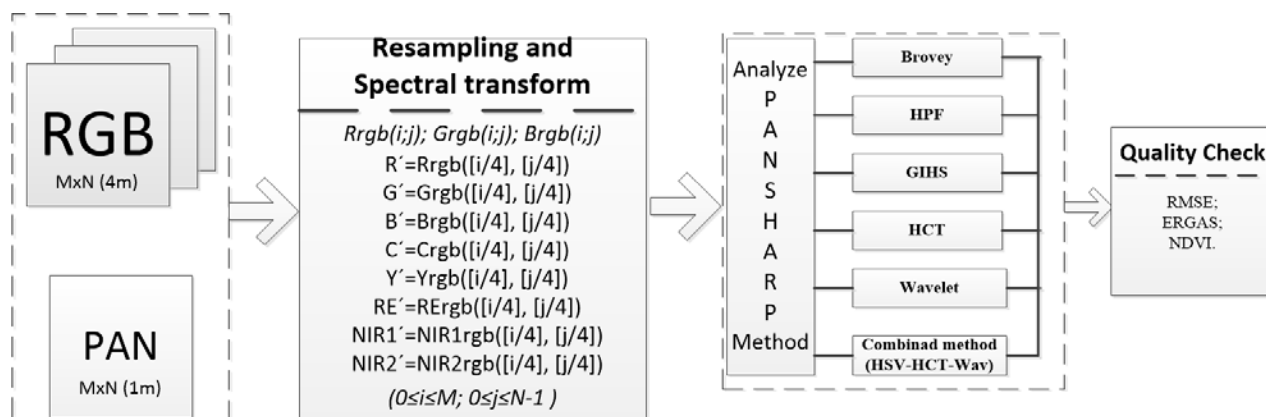


Figure 1 – Algorithm scheme

3.5 Methods on the basis of wavelet transform are implemented by convolution of signal with several bandpass filters and decimation of result [22]. To perform the reverse operation (synthesis), an operation is performed to interpolate subband signals, filter and add them. The synthesis uses the same filters as the analysis, but the coefficients are presented in reverse order by the formula (8):

$$F_i = RGB_{i\uparrow} + \frac{RGB_{i\uparrow}}{\sum_{i=1}^N RGB_{i\uparrow}} \cdot P' \quad (8)$$

3.6 Combined methods combine the advantages of existing pansharpening methods with the minimization of color distortions and taking into account physical mechanisms for capturing species information [5, 6]. In [18] proposes an HSV conversion method, hyperspherical conversion, and a packet-type HSV-HCT transformation, which improves the spatial capability of the primary digital image and avoids spectral distortion in local regions.

4. The next stage is the analysis of the qualitative and quantitative characteristics of the pansharpening methods for satellite imagery considered in the work in order to determine their state and significance.

4.1 The RMSE criterion determines the deviation of the position of the point on the shot relative to its position in space [23].

4.2 ERGAS is a quality index that calculates the “number” of spectral distortions of a synthesized image [21]. An ERGAS value higher than three corresponds to the resulting low-quality image. If ERGAS is less than three, then the image quality is considered satisfactory.

4.3 One of the criteria for checking spectral properties of synthesized images is calculation of spectral index, in particular normalized vegetation index (NDVI) [24] was calculated in the sample (9):

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (9)$$

4 EXPERIMENTS

Research of influence of fusion methods was performed on the primary satellite images by WorldView-2 of 400x400 pixels, images are shown in Fig. 2.

We have selected the whole scene, which constitutes buildings, trees, grass, and roads. Fig. 2a shows the original multispectral image, Fig. 2b – panchromatic image. After image fusion by pansharpening methods discussed above, images are shown in Fig. 3.



a



b

Figure 2 – Fragments of multispectral image: a – input RGB after resampling; b – input PAN

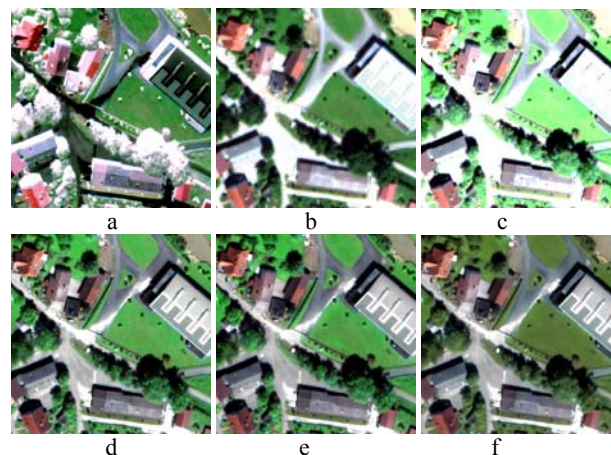


Figure 3 – Image fragments by methods: a – Brovey; b – HPF, c – GIHS; d – HCT; e – Wavelet; f – combined method

5 RESULTS

We created the graphical user interface for comparing and the choosing the best method of pansharpening satellite images of high spatial resolution. To use the GUI we have defined set of instructions:

Step 1: Click on Load Button. At first program propose to select panchromatic image, then multispectral

image. We need to be careful as the input data must be of the same scene and size.

Step 1.1: Choice the visible Bands for RGB image.

Step 2: Preprocessing the snapshot click on Scans Button and Resampling Button.

Step 2.1: Scans Button analyzed your primary input data: spectral and spatial resolution of the same scene.

Step 2.1: Resampling Button – scaled to the size of the panchromatic image and equalized the histogram to the multispectral image.

Step 3: Pansharpening Button analysis of the methods and show results. Then, we can automatically see the result in the Pansharpened image block for this you need click on Visible Button. All images are displayed in their stated blocks.

Step 4: Button Quantitative Metrics show the quantitative characteristics of the pansharpening methods for satellite imagery.

Result: Pansharpened Image and quantitative values for synthesized multispectral images.

Image quality assessment plays an important role in the processing of satellite images to determine their state and significance, especially when using image quality enhancement techniques. The first step of quality assessment is the visual analysis of the fused images. After the pansharpening of images according to the above methods, images were obtained that, even visually compared to the reference RGB, are more clear, but have significant color distortions (Figure 3). Figure 5a shows the multispectral image details containing two targets before and after pansharpening (Fig. 5b, Fig. 5c, Fig. 5d, Fig. 5e, Fig. 5f). The

synthesized multispectral image by the Brovey method has many artifacts due to colors. Image is too noisy and colors are not well synthesized as a whole and locally. Green trees are not green enough. Shapes are not well defined; they are sometimes underlined by black lines. Too large bias is observed.

So, there is too much variance introduced by Brovey method, the resulting image are shown in Fig.5b. Fig. 5c illustrates the fused image of HPF method. The result has colors that are not synthesized well enough. Contours are not sharp enough. Small objects, are white and not colored. Other colors are missing in detail, thus preventing an accurate reading of small objects. GIHS – colors are not correctly synthesized enough. Small colored objects are missing, particularly in blue and red. Contours are blurred and not sharp enough, the fused images after GIHS method are shown in figure 5d. Figure 5e illustrates the fused image after HCT method this image is nice as a whole. Colors should be better synthesized. This would enhance the legibility of the image. Details are there, except for the most colored (blue, red). Errors in colors lead to interpretation errors. Contours should be sharper. Wavelet transform – pretty nice as a whole. Efforts to do with colors. Small size objects are missing, particularly those in red or blue. It can be seen from the Fig. 5f. Combined method (HSV-HCT-Wavelet) – have better contrast and clarity on the boundaries of the “object of interest – the background”. Contours are sharp enough and colors are well synthesized. It can be seen from the Fig. 5g.

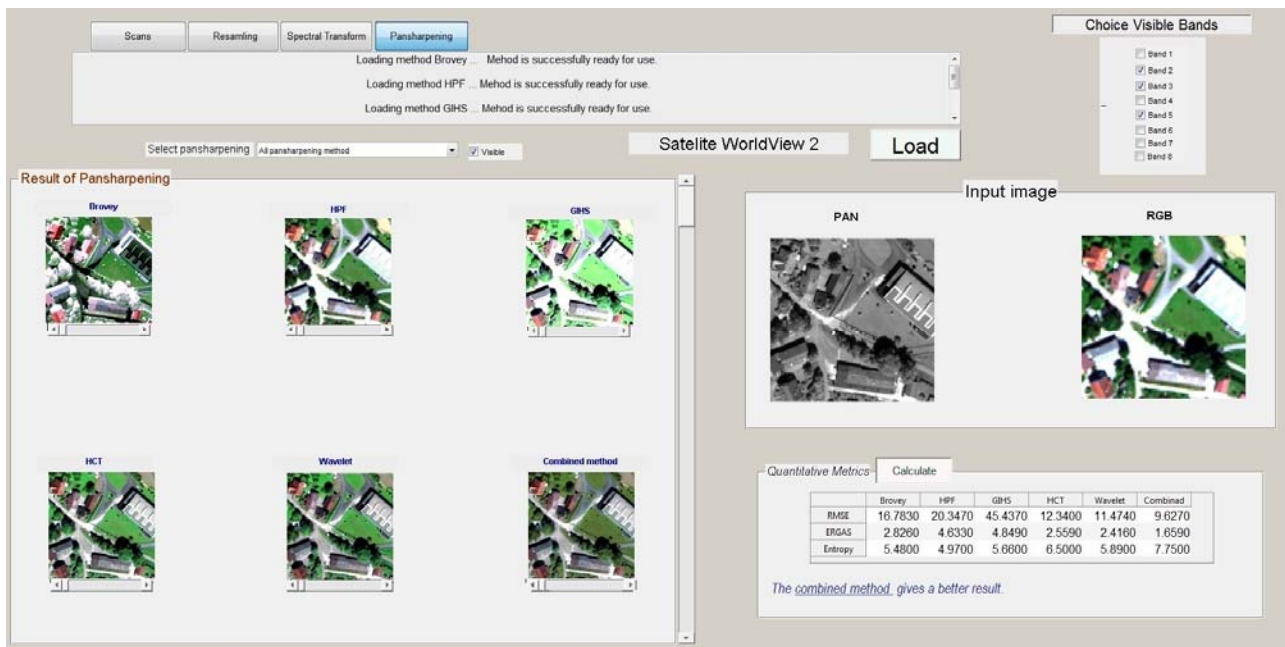


Figure 4 – Graphical user interface of proposed technology

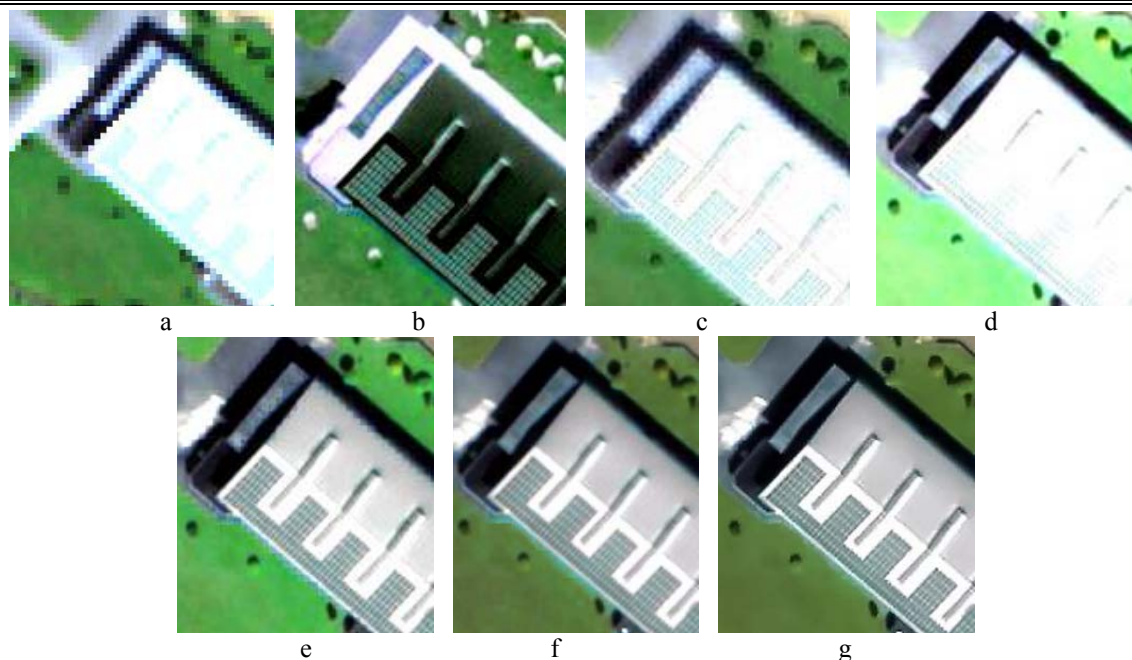


Figure 5 – The results of different classic fusion methods:
 a – Input RGB; b – Brovey; c – HPF, d – GIHS; e – HCT; f – Wavelet; g – combined method

At the same time, due to the subjectivity of human perception, the visual system perceives artifacts on the image differently, much depends on the experience of experts. Therefore, in order to determine the effect of each method on the quality of a multispectral image, a quantitative analysis of the information content of the primary and synthesized multispectral was carried out in the work, namely, found RMSE, relative dimensionless global error (ERGAS) and normalized vegetation index (NDVi).

Entropy is a measure of image information richness is an important indicator of the size of the entropy reflects the amount of information contained in the image number (entropy for PAN=7.51, RGB=7.299).

The obtained entropy, RMSE, ERGAS and NDVI values for synthesized multispectral images are shown in Table 1. We see that the combined method (HSV-HCT) compared to classical fusion methods gives a better result.

Table 1 – Comparison of fusion methods

Image/Method	RMSE	ERGAS	Entropy
PAN	–	–	7.51
RGB	–	–	7.299
GIHS	16.783	2.826	5.48
Brovey	20.347	4.639	4.97
HPF	15.437	4.849	5.66
HCT	12.311	2.659	6.05
Wavelet transform	11.474	2.416	5.89
Compiled method (HSV-HCT-Wavelet)	9.627	1.659	7.78

The efficiency of the combined method is also evidenced by the NDVI index values obtained in Table 2.

Table 2 – The results of the normalized vegetation index

Image/ Method	NDVI	NDVI for RMSE	NDV for ERGAS
Reference image of the satellite WorldView-2	1	0	0
GIHS	0.9614	0.0586	12.6028
Brovey	0.9725	0.2125	45.7155
HPF	0.9569	0.1526	32.8356
HCT	0.9578	0.0578	12.4439
Wavelet transform	0.9573	0.0603	17.568
Compiled method (HSV-HCT-Wavelet)	0.9924	0.0533	11.4107

6 DISCUSSION

The proposed technology of spatial resolution of satellites images is automatically improving the spectral quality of the sharpened image choosing the best result of the pansharpening method. As it evident from the Table 1 the quantitative evolutions of entropy for PAN is 7.51, for RGB – 7.299, after combined method – 7.78. Indicator of the size of the entropy reflects the amount of information contained in the image number, under normal circumstances, the greater the entropy of fusion images, the amount of information it contains, the more the better integration. The other quantitative evolutions of RMSE, ERGAS and NDVI have better result for the combined method with compared to classical fusion methods. This

is indicated by the value of the dimensionless global error and root mean square Error, which are the smallest (ERGAS = 1.659, RMSE = 1.659) compared to existing methods and indicates a minimum “number” of spectral distortions of satellite images and value entropy of the image pansharped by proposed technology is higher than others.

As it evident from the Table 2 the NDVI index values is efficiency of the combined method too. Obtaining a normalized vegetation index value for the Brovey and HPF methods indicates color distortion compared to the reference data. This is due to the fact that the Brovey and HPF methods are based on the fusion of three channel images and do not take into account information contained in the near infrared range, which affects the quality of spectral information of the synthesized image.

The results according to these criteria make it possible to establish a rating range of pansharpening methods from the worst result: Brovey, HPF, GIHS, soundet conversion, HCT and the combined HSV-HCT method. In addition, separate use of these methods leads to the appearance of spectral artifacts on the synthesized image compared to the reference one. The “ideal” pansharpening method is a method that increases spatial difference with minimization of color distortion.

The proposed automated technology increases the spatial resolution of multi-spectral aerospace images without color distortion.

CONCLUSIONS

The actual problem of automated increase of spatial resolution of primary multichannel images is solved in the work.

The scientific novelty of obtained results is that the proposed information technology can to increase the spatial resolution of multispectral satellite images with automation of a choice of the best pansharpening method on the basis of the analysis of quantitative and qualitative indicators is offered.

The practical significance of obtained results is that they can effectively carry out further recognition and construction of infrastructure facilities, etc. Testing showed that a synthesized image of high spatial resolution with maximum information content is achieved with the integrated use of fusion methods, which allows increasing the spatial fragmentation of a multi-channel image without significant color distortions.

Prospects for further research are to research the influence of factors that significantly affect the spatial and radiometric diversity of multichannel satellite image. Further research involves the use of a mechanism for selecting the most informative image channels. Note that the definition of criteria for assessing such informativeness and quality of the results is a separate unresolved issue.

ACKNOWLEDGEMENTS

The work is supported by the Research and Education Center of Geoinformation and Aerospace Technologies of Dnipro University of Technology.

REFERENCES

1. Kashtan V. Yu., Shedlovska Y. I., Hnatushenko V. V. Processing technology of multispectral remote sensing images [Electronic recourse], *International Young Scientists Forum on Applied Physics 2017, October 16–20, Lviv, Ukraine: Proceedings*. Lviv, 2017, pp. 355–358. DOI:10.1109/ysf.2017.8126647.
2. Zhang Y. Understanding image fusion, *Photogrammetric engineering and remote sensing*, 2004, Vol. 7, No. 6, pp. 657–661.
3. Palsson F., Sveinsson J. R., Ulfarsson M. O., Benediktsson J. A. MTF-based deblurring using a wiener filter for CS and MRA pansharpening methods, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, Vol. 9, No. 6, pp. 2255–2269.
4. Amro I., Mateos J., Vega M., Molina R., Katsaggelos A. K. A survey of classical methods and new trends in pansharpening of multispectral images, *EURASIP Journal on Advances in Signal Processing, Springer Open Journal*, 2011, No. 1, P. 79.
5. Kashtan V. Yu., Hnatushenko V. V. A Wavelet and HSV Pansharpening Technology of High Resolution Satellite Images, *Intelligent Information Technologies & Systems of Information Security IntelITSIS 2020, Khmelnytskyi, Ukraine, June 10–12, 2020*, pp. 67–76.
6. Hnatushenko V., Hnatushenko Vik., Kavats O., Shevchenko V. Pansharpening technology of high resolution multispectral and panchromatic satellite images, *Scientific Bulletin of National Mining University*, Issue 4, 2015, pp. 91–98.
7. Ghahremani M., Ghassemian H. Nonlinear IHS: A promising method for pan-sharpening, *IEEE Geoscience and Remote Sensing Letters*, Vol. 13, No. 11, 2016, pp. 1606–1610.
8. Ma J., Chen C., Li C., Huang J. Infrared and visible image fusion via gradient transfer and total variation minimization, *Information Fusion*, 2016, Vol. 31, pp. 100–109.
9. Rahmani S., Strait M., Merkurjev D., Moeller M., T. Wittman An adaptive IHS pansharpening method, *IEEE Geosci. Remote Sens. Lett.*, 2010, Vol. 7, No. 4, pp. 746–750.
10. Haitao Yin., Shutao Li Pansharpening with multiscale normalized nonlocal means filter: a two-step approach, *IEEE Transactions on Geoscience and Remote Sensing*, 2015, Vol. 53, No. 10, pp. 5734–5745.
11. Ghassemian H. A review of remote sensing image fusion methods, *Information Fusion*, 2016, 32, pp. 75–89. DOI:10.1016/j.inffus.2016.03.003
12. Wang Z., Ziou D., Armenakis C., Li D., Li Q. A comparative analysis of image fusion methods, *IEEE Transactions on Geoscience and Remote Sensing*, 2005, Vol. 43, No. 6, pp. 1391–1402.
13. Li X., Xu F., Lyu X., Tong Y., Chen Z., Li S., Liu D. A Remote-Sensing Image Pan-Sharpener Method Based on Multi-Scale Channel Attention Residual Network, *IEEE Access*, 2020, No. 8, pp. 27163–27177. DOI:10.1109/access.2020.2971502.
14. Aishwarya N., Abirami S., Amutha R. Multifocus image fusion using Discrete Wavelet Transform and Sparse Representation, *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSP-NET)*. Chennai, 2016, pp. 2377–2382. DOI: 10.1109/WiSPNET.2016.7566567.
15. Meinel G., Neubert M. A comparison of segmentation programs for high resolution remote sensing data, *International*

- Archives of Photogrammetry and Remote Sensing*, 2014, 35(Part B), pp. 1097–1105.
16. Hnatushenko V. V., Vasyliov V. V. Remote sensing image fusion using ICA and optimized wavelet transform, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B7, 2016, pp. 653–659, <https://doi.org/10.5194/isprs-archives-XLI-B7-653-2016>.
 17. Hnatushenko V. V., Shevchenko V. Yu., Kavats O. O. Improvement the spatial resolution of multichannel aerospace high spatial resolution imageson the base of hyperspherical transform, *Radio Electronics, Computer Science, Control*, 2015, No. 1, pp. 73–78 DOI 10.15588/1607-3274-2015-1-10.
 18. Tu T., Su S., Shyu H., Huang P. A new look at IHS-like image fusion methods, *Information Fusion*, 2001, Vol. 2, No. 3, pp. 177–186.
 19. Maglione P., Parente C., Vallario A. Pan-sharpening World-view-2: IHS, Brovey and Zhang methods in comparison, *Int. J. Eng. Technol.*, 2016, No. 8, pp. 673–679.
 20. Maglione P., Parente C., Vallario A. High resolution satellite images to reconstruct recent evolution of domitian coastline, *American Journal of Applied Sciences*, 2015, Vol. 12(7), pp. 506–515. DOI: 10.3844/ajassp.2015.506.515.
 21. Xu Li, Mingyi He, Lei Zhang Hyperspherical color transform based pansharpening method for WorldView-2 satellite images, *Proceedings of the 8th IEEE Conference on Industrial Electronics and Applications, Melbourne, Australia*, 2013, pp. 520–523. DOI: 10.1109/ICIEA.2013.6566424.
 22. Aishwarya N., Abirami S., Amutha R. Multifocus image fusion using Discrete Wavelet Transform and Sparse Representation, *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSP-NET)*. Chennai, 2016, pp. 2377–2382. DOI: 10.1109/WiSP-NET.2016.756656.
 23. Chen S., Zhang R., Su H., Tian J., Xia J. Scaling-up transformation of multisensor images with multiple resolutions, *Sensors*, 2009, pp. 1370–1381.
 24. Rouse J. W., Haas R. H., Schell J. A., Deering D.W. Monitoring Vegetation Systems in the Great Plains with ERTS, *3rd ERTS Symposium, NASA SP-351, Washington DC*, 10–14 December, 1973, pp. 309–317.
 25. Mustafa T., Hassoon I. K., Hassan M., Modher H. Using Water Indices (ndwi, mndwi, ndmi, wri and awei) to Detect Physical and Chemical Parameters by Apply Remote Sensing and GIS Techniques, *International Journal of Research, GRANTHAALAYAH*, 2017, Vol. 5, I. 10, pp. 117–128. DOI: 10.5281/zenodo.1040209.

Received 15.02.2021.

Accepted 29.04.2021.

УДК 004.93

АВТОМАТИЗОВАНА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ЗЛИТТЯ СУПУТНИКОВИХ ЗНІМКІВ

Гнатушенко В. В. – д-р техн. наук, професор, завідувач кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.

Каштан В. Ю. – канд. техн. наук, доцент кафедри інформаційних технологій та комп'ютерної інженерії Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.

АНОТАЦІЯ

Актуальність. На сьогоднішній день інформаційні технології широко використовуються при цифровій обробці зображень дистанційного зондування Землі. Актуальною є задача сумісної обробки супутникових зображень, отриманих різними космічними системами, які мають різне просторове розрізнення. Одними із найсучасніших супутників є WorldView-2 та WorldView-3, що дозволяють отримати восьмиканальне зображення високого просторового розрізнення. Незважаючи на розроблені методи, призначені для поліпшення якості результуючого зображення, з'являються нові наукові проблеми, які пов'язані із підвищенням вимог до результатів обробки знімків високого просторового розрізнення та розробкою автоматизованої технології обробки цих даних для подальшого їх тематичного аналізу. В роботі розглянуто концепції, принципи, обмеження та переваги традиційних методів злиття з програмною реалізацією.

Мета. Метою роботи є аналіз ефективності традиційних методів злиття, таких як Brovey, вейвлет-перетворення, GHS, HCT та комбінованого методу для супутникових зображень високого просторового розрізнення.

Метод. У роботі запропоновано інформаційну технологію злиття знімків високого просторового розрізнення з автоматизацією вибору найкращого методу злиття на основі аналізу кількісних та якісних показників. Запропонована технологія включає: масштабування мультиспектрального зображення до розміру панхроматичного зображення; використання еквалізації гістограми для коригування первинних зображень шляхом вирівнювання інтегральних областей з різною яскравістю; перетворення первинних зображень після спектральної корекції за традиційними методами «паншарпінг»; аналіз ефективності отриманих результатів. Технологія дозволяє визначити кращий метод злиття шляхом аналізу кількісних метрик: індексу NDVI, RMSE та ERGAS. Індекс NDVI для методів Brovey та HPF вказує на спектральні спотворення у порівнянні з еталонними даними. Це пов'язано з тим, що методи Brovey та HPF засновані на злитті трьохканальних зображень і не включають інформацію, що міститься в ближньому інфрачервоному діапазоні. Отримані значення RMSE та ERGAS демонструють перевагу комбінованого методу HSV-HCT-Wavelet над традиційними та найсучаснішими методами підвищення просторового розрізнення супутникових зображень. Це досягається, зокрема, попередньою обробкою первинних зображень, обробкою даних локалізованих спектральних компонент та використанням інформації, що міститься в інфрачервоному каналі зображення.

Результати. Розроблено програмне забезпечення, що реалізує запропонований підхід. Експериментальна оцінка проведена на восьмиканальних первинних супутникових знімках високого просторового розрізнення, отриманих супутником WorldView-2. Експериментальні результати показують, що синтезоване зображення з високою просторовою роздільною здатністю з високим інформаційним вмістом досягається за допомогою комплексного використання методів злиття, що дозволяє збільшити просторову роздільну здатність вихідного багатоканального зображення без спектральних спотворень.

Висновки. Проведені експерименти підтвердили ефективність запропонованої автоматизованої інформаційної технології, що дозволяє рекомендувати її при подальшому дослідженні таких задач як розпізнавання й оперативний моніторинг об'єктів інфраструктури.

© Hnatushenko V. V., Kashtan V. Yu., 2021

DOI 10.15588/1607-3274-2021-2-13

КЛЮЧОВІ СЛОВА: злиття зображень, супутникове зображення, паншарпенінг, автоматизація, WorldView, зображення високого просторового розрізнення, метааналіз.

УДК 004.93

АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ СЛИЯНИЯ СПУТНИКОВЫХ СНИМКОВ

Гнатушенко В. В. – д-р техн. наук, профессор, заведующий кафедрой информационных технологий и компьютерной инженерии Национального технического университета «Днепровская политехника», Днепр, Украина.

Каштан В. Ю. – канд. техн. наук, доцент кафедры информационных технологий и компьютерной инженерии Национального технического университета «Днепровская политехника», Днепр, Украина.

АННОТАЦИЯ

Актуальность. На сегодняшний день информационные технологии широко используются при цифровой обработке изображений дистанционного зондирования Земли. Актуальной является задача совместной обработки спутниковых изображений, полученных различными космическими системами, которые имеют разное пространственное разрешение. Одними из самых современных спутников являются WorldView-2 и WorldView-3, которые позволяют получить восьмиканальное изображение высокого пространственного разрешения. Несмотря на разработанные методы слияния, предназначенные для улучшения качества результирующего изображения, появляются новые научные проблемы, связанные с повышением требований к результатам обработки снимков высокого пространственного разрешения и разработкой автоматизированной технологии обработки этих данных для дальнейшего их тематического анализа. В работе рассмотрены концепции, принципы, ограничения и преимущества традиционных методов слияния.

Цель. Целью работы является анализ эффективности традиционных методов слияния, таких как Brovey, вейвлет-преобразования, GHS, НСТ и комбинированного метода для спутниковых изображений высокого пространственного разрешения.

Метод. В работе предложена информационная технология слияния снимков высокого пространственного разрешения с автоматизацией выбора наилучшего метода синтеза на основе анализа количественных и качественных показателей. Предложенная технология включает: масштабирование мультиспектрального изображения до размера панхроматические изображения; использование эквализации гистограммы для корректировки первичных изображений путем выравнивания интегральных областей с различной яркостью; преобразование первичных изображений после спектральной коррекции с традиционными методами паншарпинга; анализ полученных результатов. Технология позволяет определить лучший метод слияния путем анализа количественных метрик: индекса NDVI, RMSE и ERGAS. Индекс NDVI для методов Brovey и HPF указывает на искажение цвета по сравнению с эталонными данными. Это связано с тем, что методы Brovey и HPF основанные на слиянии трехканальных изображений и не включают информацию, содержащуюся в ближнем инфракрасном диапазоне. RMSE и ERGAS демонстрируют преимущество комбинированного метода HSV-НСТ-Wavelet над традиционными и современными методами повышения пространственного разрешения спутниковых изображений. Это достигается, в частности, предварительной обработкой первичных изображений, обработкой данных локализованных спектральных компонент и использованием информации, содержащейся в инфракрасном канале изображения.

Результаты. Разработано программное обеспечение, реализующее предложенный подход. Экспериментальная оценка, проведенная с использованием восьмиканальных первичных спутниковых снимков с высоким пространственным разрешением, полученных спутником WorldView-2. Экспериментальные результаты показывают, что синтезированное изображение высокого пространственного разрешения с высокой информативностью достигается с помощью комплексного использования методов синтеза, что позволяет увеличить пространственное разрешение исходного многоканального изображения без спектральных искажений.

Выводы. Проведенные эксперименты подтвердили эффективность предложенной автоматизированной информационной технологии, что позволяет рекомендовать его при дальнейшем исследовании таких задач как распознавание и оперативный мониторинг объектов инфраструктуры.

КЛЮЧЕВЫЕ СЛОВА: слияния изображений, спутниковое изображение, паншарпенінг, автоматизація, WorldView, изображение высокого пространственного разрешения, метааналіз.

ЛИТЕРАТУРА / LITERATURA

1. Kashtan V.Yu. Processing technology of multispectral remote sensing images [Electronic recourse] / V. Yu. Kashtan, Y. I. Shedlovskaya, V. V. Hnatushenko // International Young Scientists Forum on Applied Physics 2017, October 16–20, Lviv, Ukraine: Proceedings. – Lviv, 2017. – P. 355–358. DOI:10.1109/ysf.2017.8126647.
2. Zhang Y. Understanding image fusion / Y. Zhang // Photogrammetric engineering and remote sensing. – 2004. – Vol. 7, No. 6. – P. 657–661.
3. MTF-based deblurring using a Wiener filter for CS and MRA pansharpening methods / [F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, J. A. Benediktsson] // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2016. – Vol. 9, No. 6. – P. 2255–2269.
4. A survey of classical methods and new trends in pansharpening of multispectral images / [I. Amro, J. Mateos, M. Vega et al] // EURASIP Journal on Advances in Signal Processing, Springer Open Journal. – 2011. – No. 1. – P. 79.
5. Kashtan V. Yu. A Wavelet and HSV Pansharpening Technology of High Resolution Satellite Images / V. Yu. Kashtan, V. V. Hnatushenko // Intelligent Information Technologies & Systems of Information Security IntellITSIS 2020, Khmelnytskyi, Ukraine, June 10–12. – 2020. – P. 67–76.
6. Hnatushenko V. Pansharpening technology of high resolution multispectral and panchromatic satellite images / [V. Hnatushenko, Vik. Hnatushenko, O. Kavats, V. Shevchenko] // Scientific Bulletin of National Mining University. – Issue 4. – 2015. – P. 91–98.
7. Ghahremani M. Nonlinear IHS: A promising method for pan-sharpening / M. Ghahremani, H. Ghassemian // IEEE Geoscience and Remote Sensing Letters. – 2016. – Vol. 13, No. 11. – P. 1606–1610.

8. Infrared and visible image fusion via gradient transfer and total variation minimization / [J. Ma, C. Chen, C. Li, J. Huang] // *Information Fusion*. – 2016. – Vol. 31. – P. 100–109.
9. Rahmani S. An adaptive IHS pansharpening method / [S. Rahmani, M. Strait, D. Merkurjev et al] // *IEEE Geosci. Remote Sens. Lett.* – 2010. – Vol. 7, No. 4. – P. 746–750.
10. Haitao Yin. Pansharpening with multiscale normalized nonlocal means filter: a two-step approach / Yin Haitao, Li Shutao // *IEEE Transactions on Geoscience and Remote Sensing*. – 2015. – Vol. 53, No. 10. – P. 5734–5745.
11. Ghassemian H. A review of remote sensing image fusion methods / H. Ghassemian // *Information Fusion*. – 2016. – Vol. 32. – P. 75–89. DOI: 10.1016/j.inffus.2016.03.003.
12. Wang Z. A comparative analysis of image fusion methods / [Z. Wang, D. Ziou, C. Armenakis et al] // *IEEE Transactions on Geoscience and Remote Sensing*. – 2005. – Vol. 43, No. 6. – P. 1391–1402.
13. A Remote-Sensing Image Pan-Sharpener Method Based on Multi-Scale Channel Attention Residual Network / [X. Li, F. Xu, X. Lyu et al] // *IEEE Access*. – 2020. – No. 8. – P. 27163–27177. DOI:10.1109/access.2020.2971502.
14. Aishwarya N. Multifocus image fusion using Discrete Wavelet Transform and Sparse Representation / N. Aishwarya, S. Abirami, R. Amutha // 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, 2016. – P. 2377–2382. DOI: 10.1109/WiSPNET.2016.7566567.
15. Meinel G. A comparison of segmentation programs for high resolution remote sensing data / G. Meinel, M. Neubert // *International Archives of Photogrammetry and Remote Sensing*. – 2014. – 35(Part B). – P. 1097–1105.
16. Hnatushenko V.V. Remote sensing image fusion using ICA and optimized wavelet transform / V. V. Hnatushenko, V. V. Vasylyev // *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* – 2016. – XLI-B7. – P. 653–659. <https://doi.org/10.5194/isprs-archives-XLI-B7-653-2016>.
17. Hnatushenko V.V. Improvement the spatial resolution of multichannel aerospace high spatial resolution images on the base of hyperspherical transform / V. V. Hnatushenko, V. Yu. Shevchenko, O. O. Kavats // *Radio Electronics, Computer Science, Control*. – 2015. – No. 1. – P. 73–78. DOI 10.15588/1607-3274-2015-1-10.
18. A new look at IHS-like image fusion methods / [T. Tu, S. Su, H. Shyu, P. Huang] // *Information Fusion*. – 2001. – Vol. 2, No. 3. – P. 177–186.
19. Maglione P. Pan-sharpening Worldview-2: IHS, Brovey and Zhang methods in comparison / P. Maglione, C. Parente, A. Vallario // *Int. J. Eng. Technol.* – 2016. – No. 8. – P. 673–679.
20. Maglione P. High resolution satellite images to reconstruct recent evolution of domitian coastline / P. Maglione, C. Parente, A. Vallario // *American Journal of Applied Sciences*. – 2015. – Vol. 12(7). – P. 506–515. DOI: 10.3844/ajassp.2015.506.515.
21. Xu Li. Hyperspherical color transform based pansharpening method for WorldView-2 satellite images / Li Xu, He Mingyi, Zhang Lei // *Proceedings of the 8th IEEE Conference on Industrial Electronics and Applications, Melbourne, Australia*. – 2013. – P. 520–523. DOI: 10.1109/ICIEA.2013.6566424.
22. Aishwarya N. Multifocus image fusion using Discrete Wavelet Transform and Sparse Representation / N. Aishwarya, S. Abirami, R. Amutha // 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai. – 2016. – P. 2377–2382. DOI: 10.1109/WiSP-NET.2016.7566566.
23. Chen S. Scaling-up transformation of multisensor images with multiple resolutions / [S. Chen, R. Zhang, H. Su et al] // *Sensors*. – 2009. – P. 1370–1381.
24. Rouse J. W. Monitoring Vegetation Systems in the Great Plains with ERTS / [J. W. Rouse, R. H. Haas, J. A. Schell, D. W. Deering] // 3rd ERTS Symposium, NASA SP-351, Washington DC, 10–14 December. – 1973. – P. 309–317.
25. Using Water Indices (ndwi, mndwi, ndmi, wri and awei) to Detect Physical and Chemical Parameters by Apply Remote Sensing and GIS Techniques / [T. Mustafa, I. K. Hassoon, M. Hassan, H. Modher] // *International Journal of Research, GRANTHAALAYAH*. – 2017. – Vol. 5, I. 10. – P. 117–128. DOI: 10.5281/zenodo.1040209.

AN APPROACH WEB SERVICE SELECTION BY QUALITY CRITERIA BASED ON SENSITIVITY ANALYSIS OF MCDM METHODS

Polska O. V. – Senior lecturer of the Department of Computer Systems and Networks, Zaporizhzhia Polytechnic National University, Zaporizhzhia, Ukraine.

Kudermetov R. K. – PhD, Associate Professor, Head of the Department of Computer Systems and Networks, Zaporizhzhia Polytechnic National University, Zaporizhzhia, Ukraine.

Shkarupylo V. V. – PhD, Associate Professor, Associate Professor of the Department of Computer Systems and Networks, National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine.

ABSTRACT

Context. The problem of QoS based Web service from the list of Web services with equal or similar functionality was considered. This task is an essential part of the processes of finding, discover, matching and using Web services on the Internet due to the numerous offerings of Web services with equal or similar functionality. The reasonable selection of a suitable Web service takes into account a lot of user's quality requirements, such as response time, throughput, reliability, cost, etc. Such a task is usually formulated as an MCDM problem, in which the parameters are the Web service quality factors and the importance degree of these factors. The object of this research is a process of selection Web services using MCDM methods, taking into account the user's preferences and requirements to the Web service quality characteristics. The subject of the research is the LSP method, which, in addition to the degree of importance of the criteria used in all MCDM methods, simulates the user's reasoning about quality, taking into account, in particular, such characteristics of the criteria as mandatory, sufficiency, desirability, simultaneity and substitutability.

Objective. The objective of the work is to develop an approach for comparing the result of using the LSP method with the results of using other MCDM methods.

Method. A method for calculating the weights of input criteria that are not always explicitly specified in the LSP method was proposed. For this, the conjunctive coefficients of impact are used, which are calculated as a result of the sensitivity analysis of the Web service generalized quality criterion to changes the partial quality criteria. This method underlies the proposed approach to comparing the efficiency of the LSP method with other MCDM methods, which consists of using the obtained weights as the weights of the input criteria for the MCDM methods.

Results. The developed method and approach was verified experimentally. The Web service ranking produced by the LSP method was compared with the ones produced by SAW, AHP, TOPSIS and VIKOR methods. This comparison confirmed the efficiency of the proposed method and approach.

Conclusions. From the obtained results of comparing the LSP method and the MCDM methods considered in this study, it follows that the proposed method and approach provide the equivalent input conditions for these methods as for the LSP method, which is a necessary condition for the correct comparison of MCDM methods. The use of the proposed approach made it possible to study the sensitivities of the considered MCDM methods. In practical applications, this approach can be used to select a suitable MCDM method. The proposed method can be useful for creating professional evaluation systems in which it is necessary to assess the importance (weights) of tens and hundreds of quality criteria.

KEYWORDS: quality of Web service, Logical Scoring of Preference method, ranking measure, SAW, AHP, TOPSIS, VIKOR.

ABBREVIATIONS

AHP is an Analytic Hierarchy Process;
CPA/DPA is a Conjunction/Disjunction Partial Absorption;
DCG is a Discounted Cumulative Gain;
ELECTRE is an ELimination Et Choice Translating REality;
GCD is a Generalized Conjunction/Disjunction;
GL is a Graded Logic;
HPC/SPC is a Hard/Soft Partial Conjunction;
HPD/SPD is a Hard/Soft Partial Disjunction;
IDCG is an Ideal Discounted Cumulative Gain;
LSP is a Logical Scoring of Preference;
MCDM is a multi-criteria decision-making;
nDCG is a Normalized Discounted Cumulative Gain;
OWA is an Ordered Weighted Averaging;
QoS is a quality of service;
QUALIFLEX is a QUALItative FLEXible;
RCI is a Rank Consistency Index;

SAW is a Simple Additive Weighting;
TOPSIS is a Technique for Order of Preference by Similarity to Ideal Solution;
VIKOR is a VIse Kriterijumska Optimizacija Kompromisno Resenje;
WPM is a weighted power mean;
WsRF is a Web services Relevance Function.

NOMENCLATURE

E_0 is a global QoS criterion (preference) of the Web service;
 E_j is a j -th local QoS criterion of the Web service;
 RCI_k is a Rank Consistency Index of k -th MCDM method;
 r is a power in WPM function;
 S_i is an identifier of the i -th Web service alternative;
 W_1, W_2 are the weights of CPA function;

w_j is the weight of j -th local QoS criterion;

α is a degree of the *andness* parameter;

γ_j is a conjunctive coefficient of impact of the j -th QoS local criterion.

INTRODUCTION

Every day people are faced with the problem of choosing or selecting a subject and behavioural entities among their possible alternatives. This also applies to the choice of web services as they have become a part of people's daily life. For example, payment services, professional and personal E-shopping and E-commerce services, postal services, buying transport tickets services, booking a hotel services, etc. The problem of selecting Web services arises when creating and orchestrating composite services that involve several separate (atomic) services. This problem in many cases is due to the fact that there are a large number of services with the equal or similar functions, but the consumer may have different requirements for the quality characteristics of services, for instance, cost, response time, throughput, security, availability, etc. Such Web service characteristics are named non-functional characteristics or QoS. Therefore, the problem of choice is to select services suitable for the consumer QoS. This problem usually is formulated as MCDM [1].

The object of study is a process of QoS-based ranking Web services using MCDM methods.

There are many approaches to the selection of QoS-based Web services selecting with MCDM methods. However, these methods can produce different and even inconsistent ranking results. Therefore, it is recommended to compare and select the most appropriate methods depending on the context of the selection process. For the appropriate comparison of ranking methods, it is necessary to provide similar conditions and initial data and this process is also no simple task.

The subject of study is the LSP method, which allows more thoroughly than other MCDM methods to select the Web service on the base QoS properties and user's preferences.

The purpose of the work is to develop an approach that allows to correctly compare the results of Web service ranking by the LSP method with the ranking results of produced other MCDM methods.

Many popular MCDM methods are applied to the QoS-driven Web service selection, for example SAW, AHP, TOPSIS, VIKOR, ELECTRE, etc. A description of these and many other MCDM methods can be found, for example, in [2–4]. The LSP method [5] is different from these methods. The main differences are as follows: first, LSP takes into account user preferences more strictly through the use of such characteristics as simultaneity and substitutability of criteria; secondly, the use of composite aggregating functions makes it possible to more carefully aggregate small groups of criteria; third, the method allows adjusting the relationship between mandatory, sufficient and desirable criteria. The use of these criteria characteristics by the LSP method, together with such human

perception characteristics of importance and preference, make it possible to more accurately evaluate the quality of the selectable object and, respectively, to make a decision close to optimal.

1 PROBLEM STATEMENT

The LSP method is based on GL, which is an extension of traditional logic [5, 6]. Although GL can be implemented with different functions, the WPM is more often used:

$$E_0 = (w_1 E_1^r + \dots + w_n E_n^r)^{1/r}, \quad (1)$$
$$0 \leq w_j \leq 1, \quad \sum_{j=1}^m w_j = 1.$$

In this function, the parameter r is used to express logical relationships between criteria. The main logical relationships between the criteria are set by the *andness* or *orness* parameter. The andness degrees of gradations are presented in Table 1. The building of the criteria hierarchy is used in the LSP. This hierarchy is designed to form groups, subgroups and individual criteria, which are then, aggregated using GL functions with different andness values. The andness values are chosen depending on what requirements are imposed on quality characteristics: mandatory, sufficiency, desirability, as well as simultaneity and substitutability.

The MCDM methods can produce different results [7, 8]. It depends both on the aggregation function used in the method, and on the normalization techniques, which are needed in almost all MCDM methods. The LSP method uses mainly WPM functions to aggregate criteria. In addition, some criteria can be aggregated using the partial absorption functions CPA and DPA. For these functions, the weights of the input criteria are determined using penalties and rewards that the user allows for decreasing or increasing the values of some criteria relative to other criteria. Therefore, when building a composite LSP aggregator, a complete list of criteria weights may be missing.

To select the MCDM method for solving a specific problem, it is important (recommended) [9] to compare several methods with each other. Thus, the problem arises of how to compare the effectiveness of the LSP method with the effectiveness of other MCDM methods and how to compare the provided by these methods ranking results of the alternatives.

2 REVIEW OF THE LITERATURE

There are known many studies on the application of MCDM methods for QoS based Web service selection, e.g. AHP [1], SAW [10], WsRF [11], ELECTRE [12], TOPSIS [13, 14]. In these studies, QoS preference is mainly determined by assessing the importance of each quality characteristic. This importance is usually represented as a normalized numeric value called weight. A detailed survey of the Web service selection and MCDM methods can be found in [15]. The application of the LSP

method for QoS-based Web services ranking investigated in [16–18].

In papers [17, 18], the LSP method applied for automated QoS-based dynamic evaluation and selection of Web services. For this, the OWA method proposed by R. Yager (see, e.g., [19]) was used to determine the weights of the quality characteristics that are not mandatory. The disadvantage of this approach is that only one aggregator is used that implements the CPA function. This bounds the capabilities of the LSP method because OWA does not support the models of simultaneity and substitutability that are the significant feature of the LSP method.

For conclusively decision making with the MCDM method, it is recommended to compare the results obtained by this method with the results produced by other MCDM methods. The main described in the literature methods for comparison of the MCDM methods can be divided into methods comparison ranking sensitivities and methods comparison consistency rankings by ranking measures [20, 21]. The sensitivity analysis allows verifying the MCDM method in terms of the correct method parameters selection. The work [8] described an approach for comparing the SAW and TOPSIS methods based on the sensitivity of the ranking to the initial data and the weights of the criteria.

In paper [7], the results of comparison of decisions obtained by the TOPSIS, COPRAS, VIKOR and ELECTRE methods presented. For this, an analysis of the sensitivity of ratings to changes in criteria weight and methods consistency to varying measurement scales and criteria formulations used. In this work, one of the criteria for choosing the most suitable MCDM method is the stability of the produced by method rating to changes in the weight of the criteria. This approach is applicable if the ranges of the weights are known. In addition, taking into account that weights are normalized, the change of one weight of the input criteria requires the change of one or several other weights. This makes it difficult to analyze the true effect of each weight on the stability of the MCDM method. The authors of [22] proposed a method for finding new criterion weights when one of them is changed and used it to study the sensitivity of the QUALIFLEX and VIKOR methods on the weight variations.

In paper [8] the method and algorithm for statistically sensitivity analysis of the TOPSIS method were proposed. As the research result, the conclusion that the TOPSIS method is sensitive to input criteria distribution and the complying of the distributions of the TOPSIS criterion values of alternatives with the input criteria distributions was made. A significant computationally intensive may be required for the practical application of this method. Nevertheless, the applied approach can be useful for researching other MCDM methods, especially for methods in which it is difficult to clearly represent the causal relationship between the input values of the criteria and the result of alternatives ranking.

The authors of [23] used the sensitivity analysis of the VIKOR method to determine the stability solutions to the

criteria weights change. The sets of weights were determined by repeating the process of applying the AHP method and matrix of pairwise comparisons. The main drawback of this approach is that can impact to decision maker preference.

The integrated AHP-TOPSIS approach was proposed in [24]. In this approach, the AHP method was applied to determine the criteria weights and TOPSIS was adopted to select the optimal alternatives. Sensitivity analysis was performed by adjusting the weights obtained from the AHP procedures. The disadvantage of this sensitivity analysis is that it only aims to raise awareness among project participants of the potential impact on solutions by decision makers.

Summing up the works reviewed, it can be stated that the investigation of sensitivity analysis of MCDM methods is actual. The results of MCDM methods depend not only on the characteristics of the alternatives but also on the MCDM method used [7]. In addition, there is no consensus on how to determine the “quality” of the MCDM method and the reliability of its result [23].

3 MATERIALS AND METHODS

In this paper, a method to compare the results of the LSP method with other methods for QoS-based ranking Web services was proposed. The main idea of this method is to use the sensitivity of the LSP aggregation schema to variations of Web service quality characteristics to calculate the importance or weights of the input attributes for traditional MCDM methods. Attribute weights are calculated by normalizing sensitivity values. To test this idea, these weights were used as the input attribute weights in SAW, AHP, TOPSIS and VIKOR methods. The obtained ranking results were compared with the result of the LSP method ranking using the nDCG and RCI measures.

To build the aggregating schema of the LSP method, a tree of the Web services quality attributes was built. The Web services quality attributes tree is shown in Fig. 1. Formation of the criteria tree was carried out taking into account the existing Web service QoS factors standard [25], the dataset of the real Web services [26] and described in detail in [27].

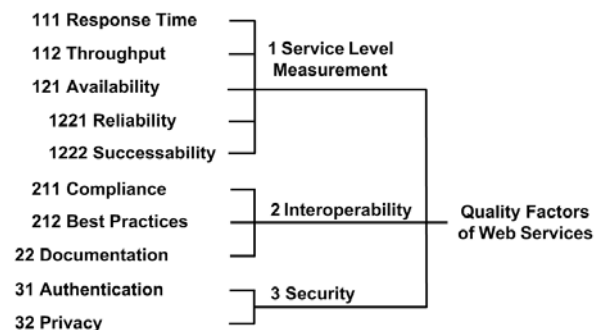


Figure 1 – The tree of the Web services quality attributes

The synthesis of the LSP aggregation schema was made taking into account the preferences of the “ordinary” user of the Web service. For this, the GCD func-

tions realized as WPM were used. The main parameters of GCD functions shown in Table 1. The WPM functions parameter r depends on the andness parameter α . The andness degree for each function was determined based on such requirements for quality attributes as mandatory, optional, sufficiency, simultaneity and substitutability. The mandatory type means that the requirement on quality characteristic is very important and must be satisfied. The optional type means that the satisfaction of the requirement on quality characteristic is desired, but not necessary. The importance of each quality attribute is assigned using weight.

Table 1 – GCD functions

GCD	Sym	α	r	$r, (n = 2)$	type
D	D	0	$r = +\infty$	$+\infty$	Sufficient
HPD	D++	1/16	$2 \leq r < +\infty$	20.36	
	D+	1/8		9.521	
	D+-	3/16		5.802	
	DA	1/4		3.929	
	D-+	5/16		2.792	
	D-	3/8		2.018	
SPD	D-	7/16	$1 < r < 2$	1.449	Optional
A	A	1/2	$r = 1$	1	
SPC	C-	9/16	$0 < r < 1$	0.619	Mandatory
	C-	5/8		0.261	
HPC	C+	11/16	$-\infty < r \leq 0$	-0.148	
	CA	3/4		-0.72	
	C+-	13/16		-1.655	
	C+	7/8		-3.51	
	C++	15/16		-9.06	
C	C	1	$r = -\infty$	$-\infty$	

In addition to the GCD functions, the LSP method uses the partial absorption aggregators CPA and DPA. These aggregators are designed to aggregate criteria that are of different types (mandatory, optional and sufficient). For instance, the CPA function aggregates mandatory criterion with optional criterion. This CPA function is shown in Fig. 2 [28].

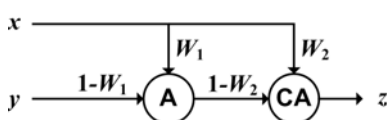


Figure 2 – An example of the CPA function

The input x of the CPA function in Fig. 2 is mandatory, and the input y is optional. The weights W_1 and W_2 are determined on the so-called penalty and reward, which are assigned by the user who selects the Web service.

The penalty is the degree to which the user agrees to reduce the aggregate satisfaction rate when a mandatory input is partially satisfied and an optional input is not satisfied. A reward is the degree to which a user wishes to increase the aggregate degree of satisfaction if a mandatory input is partially satisfied and an optional input is fully satisfied. The dependencies of the weights W_1 and W_2 on penalties and rewards for the partial absorption

functions are given in special tables developed by J. Djumović in [29]. The output value of this CPA aggregator is determined by the formula (2):

$$z = \left\{ (1 - W_2) [W_1 x + (1 - W_1) y]^r + W_2 x^r \right\}^{1/r}, \quad (2)$$

where $r = -0.72$.

As a result of the combination of WPM and CPA functions and taking into account the criteria tree, the LSP schema for the global aggregation of Web services quality criteria was synthesized. The detailed procedure of LSP aggregation schema synthesis explained in [30]. This aggregation schema is shown in Fig. 3.

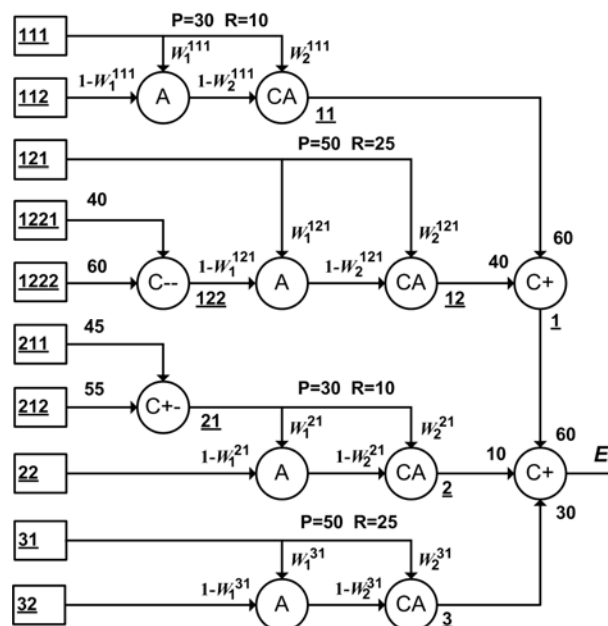


Figure 3 – Aggregate schema

The sensitivity was defined by keeping fixed all input attributes except one and then investigating the effects of the selected input on the aggregation schema output.

Obviously, the quality of service will improve as any of the quality attributes improve. For example, as the response time decreases, overall satisfaction with the Web service will increase. The response time attribute can be critical for the “ordinary” user of the Web service. This means that if the response time is longer than a certain value, then the Web service may be rejected by the user. In this case, the user’s satisfaction with the quality of Web service monotonously changes from complete dissatisfaction to complete satisfaction. On the other hand, an attribute such as throughput of the Web service may not be a critical factor (rather, this factor is critical for the Web service provider). In this case, increasing the throughput improves the overall quality of Web service, but the user can agree to some fixed minimum value for this attribute. These cases of changes of the output suitability to variation in critical (mandatory) and optional attributes (input suitability) are shown in Fig. 4. The sen-

sitivity curves in Fig. 4 are increasing concave functions for which the first derivative with respect to the variable x_j of the input attribute $\partial E_0/\partial x_j > 0$, and the second derivative $\partial^2 E_0/\partial x_j^2 < 0$.

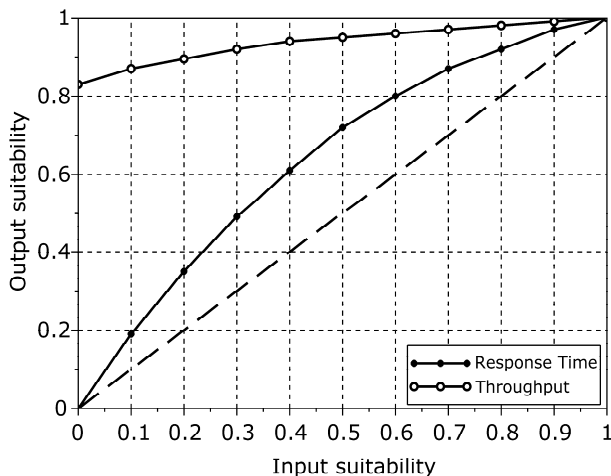


Figure 4 – The shapes of conjunctive sensitivity curves for mandatory and optional criteria

If all input attributes are fixed except one and are maximum, then the maximum impact of the unfixed attribute on the output is when $E_0 = x_j$. For this case in Fig. 4 corresponds to the dotted line. Sensitivity can be interpreted as the ratio of the area above the sensitivity curve to the area above the straight line of maximum impact. To assess the sensitivity of the output of the LSP aggregator to the input attribute variation, J. Dujmović introduced the conjunctive coefficient of impact [5], which takes into account the shape of the sensitivity curve:

$$\gamma_j = 200 \frac{y_j^{(\max)} - \int_0^1 f_j(x_j) dx_j}{y_j^{(\max)}}, [\%], \quad (3)$$

where $y_j^{(\max)}$ is the maximum value of the output suitability (maximum value of the output of the aggregator), $f_j(x_j)$ is the function of the sensitivity of the output suitability to a change in j -th input suitability (the value of the input attribute x_j).

The effect of each input attribute on the output of the aggregation schema in Fig. 3 is obtained experimentally, and then conjunctive coefficients of impact are calculated by formula (3). The input attribute weights are calculated using the formula (4):

$$w_j = \gamma_j / \sum_j \gamma_j. \quad (4)$$

These weights of input attributes can be used for SAW, AHP, TOPSIS, VIKOR methods in order to compare the ranking results of these methods with the results of ranking by the LSP method.

To evaluate the consistency of the ranking results by the LSP method and listed above methods, the nDCG and the RCI rank measures were used. The nDCG rank measure evaluate the degree of coincidence of the ranking results according to the positions of Web service alternatives in the ratings [31]. For computing the values of the nDCG of each MCDM methods, the following formulae were used [32]:

$$DCG = \sum_{i=1}^n \frac{rel(S_i)}{\log_2(i+1)}, \quad (5)$$

$$nDCG = \frac{DCG}{IDCG}, \quad (6)$$

where n is a number of Web service alternatives, $rel(S_i)$ is a relevance of Web service alternative at the i -th position and the $IDCG$ is the DCG of the rating produced by LSP method.

The RCI rank measure was used to evaluate the consistency of the ranking result of each MCDM methods with the ranking results of all considered MCDM methods. The RCI rank measure was introduced in [20], applied in [21] and modified in [33, 34]. The formula for calculating the RCI is as follows (7):

$$RCI_k = \frac{1}{(K-1)} \sum_{l=1}^{K-1} l \cdot t_{k,l}, \quad (7)$$

where K is a number of compared MCDM methods, $t_{k,l}$ is a total number of times when k -th MCDM method produced the same ranking as l other MCDM methods.

4 EXPERIMENTS

Based on the dataset [26] a list of 52 currently running Web services with similar functionality was compiled for a numerical experiment. Missing data, but needed according to the attribute tree presented in Fig. 1, was generated randomly. The method proposed in [1] was used to generate the values of the Authentication attribute. This method evaluates the quality characteristics presented as an ordered list. To calculate the values of the Authentication attribute, an ascending ordered list (Password, Certificate Based Authentication, One Time Password, Access Key Authentication, and Token Based Authentication) was used. Further, Web services with dominated quality characteristics vectors were excluded from the list of candidate Web services. As a result, a list of 7 Web service candidates (S_1 – S_7) for selection was formed. The values of these quality characteristics are presented in Table 2. The additional data of quality attributes, shown in Table 3, include the value ranges of the input attributes, their types and tendencies, which were used to build the LSP

aggregation schema. The elementary quality criteria of the Web services were calculated according to formulae (8) and (9) and taking to account the values, bands and tendencies of input quality attributes. The values of these criteria is shown in Table 4. Formula (8) is used to calculate criteria with a “benefit” tendency, which means that the higher value of the quality attribute is better than the lower one (for example throughput). Formula (9) is used to calculate criteria with a “cost” tendency, which means that the lower value of the quality attribute is better than the higher one (for example, response time).

Table 2 – Quality attributes for Web services

Criteria	Alternatives ID						
	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇
Response Time	114	113	108	132	120	100	140
Throughput	16.0	10.0	12.0	20.0	17.0	21.0	23.0
Availability	86	91	89	83	98	88	75
Reliability	73	70	74	68	83	59	88
Successability	86	85	90	77	96	90	98
Compliance	89	78	92	81	85	88	80
Best Practices	84	83	93	80	90	66	86
Documentations	60	62	90	89	72	52	95
Authentication	17	12	13	14	11	19	15
Privacy	1	0	1	1	0	1	0

Table 3 – Additional data for Quality attributes

QoS	min	max	type	tendency
Response Time, ms	50	150	mandatory	cost
Throughput, invokes/s	5	25	optional	benefit
Availability, %	60	100	mandatory	benefit
Reliability, %	40	100	optional	benefit
Successability, %	20	100	optional	benefit
Compliance, %	30	100	mandatory	benefit
Best Practices, %	40	100	mandatory	benefit
Documentations, %	20	100	optional	benefit
Authentication	2	20	mandatory	benefit
Privacy	0	1	optional	benefit

$$E_i(x_j) = \max \left[0, \min \left(1, \frac{x_{ij} - x_j^{(\min)}}{x_j^{(\max)} - x_j^{(\min)}} \right) \right], \quad (8)$$

$$E_i(x_j) = \max \left[1, \max \left(0, \frac{x_j^{(\max)} - x_{ij}}{x_j^{(\max)} - x_j^{(\min)}} \right) \right]. \quad (9)$$

Table 4 – QoS criteria

Criteria	Alternatives ID						
	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇
Response Time	0.36	0.37	0.42	0.18	0.30	0.50	0.10
Throughput	0.55	0.25	0.35	0.75	0.60	0.80	0.90
Availability	0.65	0.77	0.72	0.57	0.95	0.70	0.37
Reliability	0.55	0.50	0.57	0.47	0.72	0.32	0.80
Successability	0.82	0.81	0.87	0.71	0.95	0.87	0.97
Compliance	0.84	0.69	0.89	0.73	0.78	0.83	0.71
Best Practices	0.73	0.72	0.88	0.67	0.83	0.43	0.77
Documentations	0.50	0.52	0.87	0.86	0.65	0.40	0.94
Authentication	0.83	0.56	0.61	0.67	0.50	0.94	0.72
Privacy	1	0	1	1	0	1	0

The QoS criteria presented in Table 4 were used as inputs for the LSP aggregation schema (Fig. 3). The Web

services S₁–S₇ ranks and LSP aggregation schema sensitivities on changes of each criterion were obtained. Sensitivity curves are shown in Fig. 5 and Fig. 6. Then the sensitivity characteristics were used for calculating the weights criteria according to formula (4). These criteria weights are [0.236, 0.04, 0.162, 0.02, 0.034, 0.116, 0.128, 0.013, 0.174, 0.078].

In order to verify the proposed method, the values of considered list of input quality attributes (Table 2) was used to rank Web services by methods SAW, AHP, TOPSIS and VIKOR. For this the weights obtained as a result of calculating the sensitivity of the LSP aggregation schema were used.

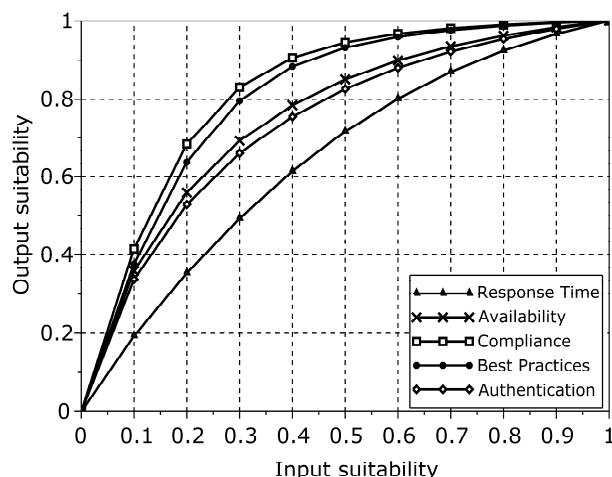


Figure 5 – Sensitivity curves for mandatory criteria

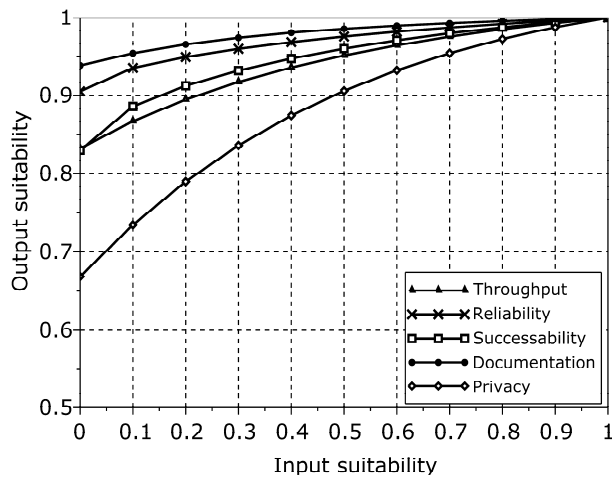


Figure 6 – Sensitivity curves for optional criteria

The sensitivity analysis of the SAW, AHP, TOPSIS and VIKOR methods was carried out similarly to the sensitivity analysis of the LSP method.

5 RESULTS

Table 5 shows the sensitivities of the considered MCDM methods to changes in each of the input attributes, as well as the average values of the sensitivity of each method. The sensitivity was calculated in accordance

with formula (3). As a function $f_i(x_i)$, the functions of the methods sensitivity curves was used. The maximum value of output suitability is obviously y_i equal to 1 (see Fig. 4).

Table 5 – Sensitivities of MCDM methods (γ_i , %)

	LSP	SAW	AHP	TOPSIS	VIKOR
Response Time	71.04	23.61	6.62	36.07	23.61
Throughput	12.010	3.99	2.04	14.92	3.99
Availability	48.61	16.16	3.32	20.74	16.16
Reliability	6.01	2.00	0.58	4.21	2.00
Successability	10.18	3.38	0.63	4.27	3.83
Compliance	34.88	11.59	1.53	9.97	11.59
Best Practices	38.38	12.76	3.26	20.33	12.76
Documentations	3.93	1.31	0.53	3.92	1.31
Authentication	52.33	17.39	7.40	49.37	17.39
Privacy	23.52	7.082	7.41	49.90	7.082
Average	30.09	10.00	3.33	21.37	10.00

Table 6 presents the results of Web service rankings obtained by MCDM methods and comparing these rankings using the rank measures nDCG and RCI. As mentioned above, the LSP ranking was chosen as the ideal ranking for the nDCG measure, since other methods use weights derived from the LSP method. The obtained high values of nDCG indicate that the SAW, AHP, TOPSIS and VIKOR methods determined the ranks of Web services in sufficient consistently with the LSP method. The RCI indicators received average values (the maximum indicator for these data is 7). This is since the considered methods differ in ranking and normalization of the input attributes techniques. These techniques can sufficiently affect alternatives with the low overall quality criterion. Nevertheless, the ranking results are quite consistent for the Web services which took place in the first three positions in the rankings.

Table 6 – Ranks of Web services and evaluations of nDCG and RCI measures

Service ID	LSP	SAW	AHP	TOPSIS	VIKOR
S ₁	3	2	2	2	3
S ₂	6	7	7	7	6
S ₃	2	3	3	3	2
S ₄	5	4	4	4	5
S ₅	4	5	5	6	4
S ₆	1	1	1	1	1
S ₇	7	6	6	5	7
nDCG, [%]	100	98	98	98.4	99.7
RCI	2.5	3.5	3.5	3.0	2.5

6 DISCUSSION

The comparison of the results of ranking by the LSP method and the SAW, AHP, TOPSIS and VIKOR methods confirmed the possibility of using the weights obtained from the LSP sensitivity analysis as the input attributes weights for the SAW, AHP, TOPSIS and VIKOR

methods. This indicates the efficiency of the proposed method. At the same time, according to the average values of the method sensitivities to the input criteria changes given in Table 5, it can be concluded that the LSP method has a higher overall sensitivity and especially to changes of mandatory criteria.

The develop a hierarchical tree of a set of input attributes for the LSP method facilitates to build the separate aggregators for small groups of input attributes (usually 2 to 5) with flexible consideration of such requirements for quality attributes as mandatory, sufficiency, optional, simultaneity and substitutability. These requirements are further taken into account when synthesizing the general aggregation schema. This approach makes it possible to more thoroughly take into account the user's preferences when QoS-based selecting the Web services. As a result, the mapping of the sensitivity of the general LSP aggregation schema to the input weights for other decision-making methods turns out to be more accurate and the decisions obtained by them are more relevant to the preferences of decision-makers.

As follows from the publications analyzed in this work, the problem of choosing the MCDM method is of current interest and it is recommended when choosing one to compare the results produced by several MCDM methods. Approaches to compare MCDM methods are constantly evolving. The proposed method can supplement already developed approaches for MCDM methods.

In many cases, for the considered MCDM methods, the main method for determining the input attributes weights is the AHP. Sometimes this method is used repeatedly to clarify the preferences of stakeholders or decision makers. The process of identifying user preferences is most advanced in the LSP method. Therefore, the proposed method, based on the determination of weights from the results of sensitivity analysis, can also be used to check or confirm the weights obtained by the AHP or other methods. This can be especially useful if the number of input attributes is a large.

CONCLUSIONS

In this work, the approach to comparing the effectiveness of the LSP method with other MCDM methods has been proposed.

The scientific novelty of this study is as follows:

- a method based on the LSP sensitivity analysis for calculating input criteria weights has been developed. These weights can be applied to other MCDM methods when compared to the LSP method;

- a comparison of the sensitivity of the LSP method with the sensitivity of SAW, AHP, TOPSIS and VIKOR methods to changes in input attributes have been made;

- to compare the ratings produced by the LSP, SAW, AHP, TOPSIS and VIKOR methods, the nDCG and RCI measures have been used.

The practical significance of the results obtained is that the developed method can be used to calculate the weights of input attributes, and the proposed approach can be used to verify the results of ranking alternatives, which

may be necessary when choosing the most suitable MCDM method.

Prospects for further research are the use of the proposed approach and method for comparing the LSP method with methods that do not have an explicit aggregation function, for example, ELECTRE and expanding the list of QoS criteria for selecting Web services.

REFERENCES

1. Tran V. X., Tsuji H., Masuda R. A new QoS ontology and its QoS-based ranking algorithm for Web services, *Simulation Modelling Practice and Theory*, 2009, Vol. 17, No. 8, pp. 1378–1398. DOI: 10.1016/j.simpat.2009.06.010
2. Tzeng G-H., Huang J-J. Multiple Attribute Decision Making : Methods and Applications. Boca Raton, Chapman and Hall (CRC Press), 2011, 336 p. DOI: 10.1201/b11032
3. Greco S., Ehrgott M., Figueira J. R. eds. Multiple Criteria Decision Analysis : State of the Art Surveys. New York, Springer, 2016, 1346 p. DOI: 10.1007/978-1-4939-3094-4
4. Alinezhad A., Khalili J. New Methods and Applications in Multiple Attribute Decision Making (MADM). Cham, Springer, 2019, 234 p. DOI: 10.1007/978-3-030-15009-9
5. Dujmović J. Soft Computing Evaluation Logic : The LSP Decision Method and Its Applications. Hoboken, John Wiley & Sons, Inc., 2018, 912 p. DOI: 10.1002/9781119256489
6. Dujmovic J. Graded logic for decision support systems, *International Journal of Intelligent Systems*, 2019, Vol. 34, № 11, pp. 2900–2919. DOI: 10.1002/int.22177
7. Pamucar D., Bozanic D., Randjelovic A. Multi-criteria decision making : An example of sensitivity analysis, *Serbian Journal of Management*, 2017, Vol. 12, No. 1, pp. 1–27. DOI: 10.5937/sjml2-9464
8. Simanavičienė R., Petraitytė V. Sensitivity Analysis of the TOPSIS Method in Respect of Initial Data Distributions, *Lithuanian Journal of Statistics*, 2016, Vol. 55, No. 1, pp. 45–51. DOI: 10.15388/LJS.2016.13866
9. Debnath N., Martellotto P., Daniele M. et al A method to evaluate QoS of web services required by a workflow, *ITS Telecommunications : 11th international conference, St. Petersburg, 23–25 August 2011 : proceeding. IEEE*, 2011, pp. 640–645. DOI: 10.1109/ITST.2011.6060134
10. Maheswari S., Karpagam G. R. Performance evaluation of semantic based service selection methods, *Computers and Electrical Engineering*, 2017, Vol. 71, pp. 966–977. DOI: 10.1016/j.compeleceng.2017.10.006
11. Al-Masri E., Mahmoud Q. H. QoS-based Discovery and Ranking of Web Services, *Computer Communications and Networks : 16th international conference, Honolulu, 13–16 August 2007 : proceeding. IEEE*, 2007, pp. 529–534. DOI: 10.1109/ICCCN.2007.4317873
12. Ma H., Zhu H., Hu Z. et al. Time-aware trustworthiness ranking prediction for cloud services using interval neutrosophic set and ELECTRE, *Knowledge-Based Systems – 2017*, Vol. 138, pp. 27–45. DOI: 10.1016/j.knsys.2017.09.027
13. Belouaar H., Kazar O., Rezeg K. Web service selection based on TOPSIS algorithm, *Mathematics and Information Technology : International conference, Adrar, 4–5 December 2017 : proceeding. IEEE*, 2018, pp. 177–182. DOI: 10.1109/MATHIT.2017.8259713
14. Sun R., Zhang B., Liu T. Ranking web service for high quality by applying improved Entropy-TOPSIS method, *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing : 17th IEEE/ACIS international conference, Shanghai, 30 May–1 June 2016: proceeding. IEEE*, 2016, pp. 249–254. DOI: 10.1109/SNPD.2016.7515909
15. Hosseinzadeh M., Hama H. K., Ghafour M. Y. et al. Service Selection Using Multi-criteria Decision Making : A Comprehensive Overview, *Journal of Network and Systems Management*, 2020, Vol. 28, № 4, pp. 1639–1693. DOI: 10.1007/s10922-020-09553-w
16. Yu H. Q. and Molina H. A modified Logic Scoring Preference method for dynamic Web services evaluation and selection, *Service Oriented Computing : 2nd European Seminar for Young Researchers, Leicester, 11–12 June 2007 : proceedings*. Leicester, 2007, pp. 87–93.
17. Yu H. Q., Reiff-Marganec S. A Method for Automated Web Service Selection, *Services – Part I : IEEE congress, Honolulu, 6–11 July 2008 : proceeding. IEEE*, 2008, pp. 513–520. DOI: 10.1109/SERVICES-1.2008.8
18. Yu H. Q., Reiff-Marganec S. Automated Context-Aware Service Selection for Collaborative Systems, *Advanced Information Systems Engineering : 21st International Conference, CAiSE 2009, Amsterdam, 8–12 June, 2009 : proceedings*. Berlin, Springer, 2009, pp. 261–274. (Lecture Notes in Computer Science, Vol. 5565). DOI: 10.1007/978-3-642-02144-2_23
19. Yager R. R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking, *IEEE Transactions on Systems, Man, Cybernetics*, 1988, Vol. 18, No. 1, pp. 183–190. DOI: 10.1109/21.87068
20. Chakraborty S., Yeh C.-H. A simulation comparison of normalization procedures for TOPSIS, *Computers and Industrial Engineering : International conference, Troyes, 6–9 July, 2009 : proceedings. IEEE*, 2009, pp. 1815–1820. DOI: 10.1109/iccic.2009.5223811
21. Vafaei N., Ribeiro R. A., Camarinha-Matos L. M. Data normalisation techniques in decision making: case study with TOPSIS method, *International Journal of Information and Decision Sciences*, 2018, Vol. 10, No. 1, pp. 19–38. DOI: 10.1504/ijids.2018.090667
22. Alinezhad A., Esfandiari N. Sensitivity Analysis in the QUALIFLEX and VIKOR Methods, *Journal of Optimization in Industrial Engineering*, 2012, Vol. 6, No. 1 (10), pp. 29–34.
23. Muñoz B., Romana M. G., Ordóñez J. Sensitivity Analysis of Multicriteria Decision Making Methodology Developed for Selection of Typologies of Earth-retaining Walls in an Urban Highway, *Transportation Research Procedia*, 2016, Vol. 18, pp. 135–139. DOI: 10.1016/j.trpro.2016.12.019
24. Tian G., Zhang H., Jia H. et al. Automotive style design assessment and sensitivity analysis using integrated analytic hierarchy process and technique for order preference by similarity to ideal solution, *Advances in Mechanical Engineering*, 2016, Vol. 8, No. 5, pp. 1–10. DOI: 10.1177/1687814016649885
25. Web Services Quality Factors Version 1.0, Candidate OASIS Standard 01, October 2012 [Electronic resource]. Access mode: <http://docs.oasis-open.org/wsrm/WS-Quality-Factors/v1.0/cos01/WS-Quality-Factors-v1.0-cos01.html>
26. QWS Dataset v2.0, November 2019 [Electronic resource]. Access mode: <https://github.com/qwsdata/qwsdata.github.io/releases>
27. Polska O. V. Kudermetov R. K., N. V. Shcherbak Model of web services quality criteria hierarchy, *Visnik Zaporiz'koho natsional'noho universitetu. Fiziko-matematichni nauki*,

- 2020, Vol. 2, pp. 43–51. DOI: 10.26661/2413-6549-2020-2-06
28. Dujmovic J. J. Continuous Preference Logic for System Evaluation, *IEEE Transactions on Fuzzy Systems*, 2007, Vol. 15, No. 6, pp. 1082–1099. DOI: 10.1109/TFUZZ.2007.902041
29. Dujmović, J. Partial absorption function, *Journal of the University of Belgrade, EE Dept., Series Mathematics and Physics*, 1979, No. 659, pp. 156–163.
30. Polska O. V., Kudermetov R. K., Shkarupylo V. V. The approach for QoS based web service selection with user's preferences, *Naukovi pratsi Donets'koho natsional'nogo tekhnichnogo universitetu. Problemi modelyuvannya ta avtomatizatsiyi proektuvannya*, 2020, No. 16, pp. 19–27. DOI: 10.31474/2074-7888-2020-2-19-27
31. Wang Y., Wang L., Li Y. et al. A Theoretical Analysis of NDCG Ranking Measures, *Learning Theory: 26th annual conference, COLT 2013, Princeton, 12–14 June, 2013: proceedings*. JMLR.org, 2013, pp. 25–54.
32. Fogli A., Sansonetti G. Exploiting semantics for context-aware itinerary recommendation, *Personal and Ubiquitous Computing*, 2019, Vol. 23, pp. 215–231. DOI: 10.1007/s00779-018-01189-7
33. Kudermetov R., Polska O., Shkarupylo V. et al. Normalization Techniques Comparison for QoS-based Web Services Selection by LSP Method, *Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS): 5th IEEE International Symposium Smart and Wireless Systems within the Conferences, Dortmund, 17–18 September 2020: proceedings*. IEEE, 2020, pp. 213–216. DOI: 10.1109/idaacs-sws50031.2020.9297098
34. Polska O., Kudermetov R., Alsayaydeh J. A. J. et al. QoS-aware Web-services Ranking: Normalization Techniques Comparative Analysis for LSP Method, *ARPN Journal of Engineering and Applied Sciences*, 2021, Vol. 16, No. 2, pp. 248–254.

Received 17.04.2021.

Accepted 03.06.2021.

УДК 004.42

ПІДХІД ДО ВИБОРУ ВЕБСЕРВІСІВ ЗА КРИТЕРІЯМИ ЯКОСТІ НА ОСНОВІ АНАЛІЗУ ЧУТЛИВОСТІ МЕТОДІВ MCDM

Польська О. В. – старший викладач кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка», Запоріжжя, Україна.

Кудерметов Р. К. – канд. техн. наук, доцент, завідувач кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка», Запоріжжя, Україна.

Шкарупило В. В. – канд. техн. наук, доцент, доцент кафедри комп'ютерних систем та мереж Національного університету біоресурсів і природокористування України, Київ, Україна.

АНОТАЦІЯ

Актуальність. Розглянуто задачу вибору вебсервісів за критеріями якості зі списку вебсервісів з однаковими або подібними функціональними характеристиками. Це завдання є невід'ємною частиною процесів пошуку, вибору, узгодження і використання таких вебсервісів в зв'язку з численними пропозиціями вебсервісів в Інтернеті. Розумний вибір необхідного вебсервісу може враховувати безліч вимог користувача до якісних характеристик, наприклад, часу відгуку, пропускну здатності, надійності, вартості тощо. Така задача зазвичай формулюється як багатокритеріальна задача прийняття рішень, в якій параметрами є чинники якості вебсервісів і ступеню значущості параметрів якості. Об'єктом дослідження є процеси ранжування вебсервісів за допомогою методів MCDM з урахуванням переваг та вимог користувача до характеристик якості вебсервісів. Предмет дослідження – метод LSP, який крім ступеню важливості критеріїв, що використовується у всіх методах MCDM, моделює міркування користувача про якість, враховуючи, зокрема, такі характеристики критеріїв, як обов'язковість, достатність, бажаність, одночасність та взаємозамінюваність. Мета дослідження – розробити підхід для порівняння результатів використання методу LSP з іншими методами MCDM.

Метод. Запропоновано метод обчислення ваг вхідних критеріїв, які не завжди задані явно в методі LSP. Для цього використовуються коефіцієнти кон'юнктивного впливу, які обчислюються в результаті процедури аналізу чутливості узагальненого критерію якості до змін вхідних критеріїв якості. Цей метод лежить в основі запропонованого підходу до порівняння ефективності методу LSP з іншими методами MCDM, який полягає в використанні отриманих ваг в якості ваг вхідних критеріїв для MCDM методів.

Результати. Розроблені метод і підхід було перевірено експериментально. В якості MCDM методів, з якими здійснювалось порівняння метод LSP, використовувалися методи SAW, АНР, TOPSIS і VIKOR. Порівняння результатів ранжування вебсервісів цими методами з результатом ранжирування методом LSP підтвердило працездатність запропонованих методу та підходу.

Висновки. З отриманих результатів порівняння методу LSP і розглянутих в даному дослідженні методів MCDM випливає, що запропоновані метод і підхід забезпечують для цих методів рівнозначні з методом LSP вхідні умови, що є необхідною умовою коректності порівняння методів MCDM. Використання запропонованого підходу дозволило дослідити чутливості розглянутих методів MCDM. У практичному застосуванні даний підхід можна використовувати для вибору підходящого MCDM методу. Запропонований метод може бути корисний для створення професійних систем оцінювання, коли необхідно оцінювати важливість (ваги) десятків і сотень критеріїв якості.

КЛЮЧОВІ СЛОВА: якість вебсервісу, метод логічного оцінювання вподобань, міри ранжування, SAW, АНР, TOPSIS, VIKOR.

ПОДХОД К ВЫБОРУ ВЕБ-СЕРВИСОВ ПО КРИТЕРИЯМ КАЧЕСТВА НА ОСНОВЕ АНАЛИЗА ЧУВСТВИТЕЛЬНОСТИ МЕТОДОВ MCDM

Польская О. В. – старший преподаватель кафедры компьютерных систем и сетей Национального университета «Запорожская политехника», Запорожье, Украина.

Кудерметов Р. К. – канд. техн. наук, доцент, заведующий кафедрой компьютерных систем и сетей Национального университета «Запорожская политехника», Запорожье, Украина.

Шкарупило В. В. – канд. техн. наук, доцент, доцент кафедры компьютерных систем и сетей Национального университета биоресурсов и природопользования Украины, Киев, Украина.

АННОТАЦИЯ

Актуальность. Рассмотрена задача выбора веб-сервисов по качественным характеристикам из списка веб-сервисов с одинаковыми или подобными функциональными характеристиками. Эта задача является неотъемлемой частью процессов поиска, выбора, согласования и использования таких веб-сервисов в связи с многочисленными предложениями веб-сервисов в Интернете. Разумный выбор необходимого веб-сервиса может учитывать множество требований пользователя к качественным характеристикам, например, времени отклика, пропускной способности, надежности, стоимости и т. д. Такая задача обычно формулируется как задача MCDM, в которой параметрами являются факторы качества веб-сервисов и степени значимости параметров качества. Объектом исследования являются процессы ранжирования веб-сервисов с помощью методов MCDM с учетом предпочтений и требований пользователя к характеристикам качества веб-сервисов. Предмет исследования – метод LSP, который помимо степени важности критериев, использующейся во всех методах MCDM, моделирует рассуждения пользователя о качестве, учитывая, в частности, такие характеристики критериев, как обязательность, достаточность, желательность, одновременность и взаимозаменяемость. Цель исследования – разработать подход для сравнения результатов использования метода LSP с другими методами MCDM.

Метод. Предложен метод вычисления весов входных критериев, которые не всегда заданы явно в методе LSP. Для этого используются коэффициенты конъюнктивного влияния, которые вычисляются в результате процедуры анализа чувствительности обобщенного критерия качества к изменениям входных критериев качества. Этот метод лежит в основе предложенного подхода к сравнению эффективности метода LSP с другими методами MCDM, который заключается в использовании полученных весов в качестве весов входных критериев для MCDM методов.

Результаты. Разработанный метод и подход были проверены экспериментально. В качестве MCDM методов, с которыми сравнивался метод LSP, использовались методы SAW, AHP, TOPSIS и VIKOR. Сравнение результатов ранжирования веб-сервисов этими методами с результатом ранжирования методом LSP подтвердило работоспособность предложенных метода и подхода.

Выводы. Из полученных результатов сравнения метода LSP и рассмотренных в данном исследовании методов MCDM следует, что предложенные метод и подход обеспечивают для этих методов равнозначные с LSP входные условия, что является необходимым условием корректности сравнения методов MCDM. Использование предложенного подхода позволило исследовать чувствительности рассмотренных методов MCDM. В практическом применении данный подход можно использовать для выбора подходящего MCDM метода. Предложенный метод может быть полезен для создания профессиональных систем оценивания, когда необходимо оценивать важность (веса) десятков и сотен критериев качества.

КЛЮЧЕВЫЕ СЛОВА: качество веб-сервиса, метод логического оценивания предпочтений, меры ранжирования, SAW, AHP, TOPSIS, VIKOR.

ЛИТЕРАТУРА / LITERATURA

1. Tran V. X. A new QoS ontology and its QoS-based ranking algorithm for Web services / V. X. Tran, H. Tsuji, R. Masuda // *Simulation Modelling Practice and Theory* – 2009. – Vol. 17, № 8. – P. 1378–1398. DOI: 10.1016/j.simpat.2009.06.010
2. Tzeng G-H. Multiple Attribute Decision Making: Methods and Applications / G-H. Tzeng, J-J. Huang. – Boca Raton: Chapman and Hall (CRC Press), 2011. – 336 p. DOI: 10.1201/b11032
3. Greco S. Multiple Criteria Decision Analysis: State of the Art Surveys / eds.: S. Greco, M. Ehrgott, J. R. Figueira. – New York: Springer, 2016. – 1346 p. DOI: 10.1007/978-1-4939-3094-4
4. Alinezhad A. New Methods and Applications in Multiple Attribute Decision Making (MADM) / A. Alinezhad, J. Khalili – Cham: Springer, 2019. – 234 p. DOI: 10.1007/978-3-030-15009-9
5. Dujmović J. Soft Computing Evaluation Logic: The LSP Decision Method and Its Applications / J. Dujmović. – Hoboken: John Wiley & Sons, Inc., 2018. – 912 p. DOI: 10.1002/9781119256489
6. Dujmovic J. Graded logic for decision support systems / J. Dujmovic // *International Journal of Intelligent Systems* – 2019. – Vol. 34, № 11. – P. 2900–2919. DOI: 10.1002/int.22177
7. Pamucar D. Multi-criteria decision making: An example of sensitivity analysis / D. Pamucar, D. Bozanic, A. Randjelovic // *Serbian Journal of Management*. – 2017. – Vol. 12, № 1. – P. 1–27. DOI: 10.5937/sjm12-9464
8. Simanavičienė R. Sensitivity Analysis of the TOPSIS Method in Respect of Initial Data Distributions / R. Simanavičienė, V. Petraitytė // *Lithuanian Journal of Statistics* – 2016. – Vol. 55, № 1. – P. 45–51. DOI: 10.15388/LJS.2016.13866
9. A method to evaluate QoS of web services required by a workflow / [N. Debnath, P. Martellotto, M. Daniele et al.] // *ITS Telecommunications: 11th international conference, St. Petersburg, 23–25 August 2011: proceeding.* – IEEE, 2011. – P. 640–645. DOI: 10.1109/ITST.2011.6060134
10. Maheswari S. Performance evaluation of semantic based service selection methods / S. Maheswari, G. R. Karpagam // *Computers and Electrical Engineering*. – 2017. –

- Vol. 71. – P. 966–977. DOI: 10.1016/j.compeleceng.2017.10.006
11. Al-Masri E. QoS-based Discovery and Ranking of Web Services / E. Al-Masri, Q. H. Mahmoud // *Computer Communications and Networks* : 16th international conference, Honolulu, 13–16 August 2007 : proceeding. – IEEE, 2007. – P. 529–534. DOI: 10.1109/ICCCN.2007.4317873
 12. Time-aware trustworthiness ranking prediction for cloud services using interval neutrosophic set and ELECTRE / [H. Ma, H. Zhu, Z. Hu et al.] // *Knowledge-Based Systems* – 2017. – Vol. 138. – P. 27–45. DOI: 10.1016/j.knosys.2017.09.027
 13. Belouaar H. Web service selection based on TOPSIS algorithm / H. Belouaar, O. Kazar, K. Rezeg // *Mathematics and Information Technology* : International conference, Adrar, 4–5 December 2017 : proceeding. – IEEE, 2018. – P. 177–182. DOI: 10.1109/MATHIT.2017.8259713
 14. Sun R. Ranking web service for high quality by applying improved Entropy-TOPSIS method / R. Sun, B. Zhang, T. Liu // *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing* : 17th IEEE/ACIS international conference, Shanghai, 30 May–1 June 2016: proceeding. – IEEE, 2016. – P. 249–254. DOI: 10.1109/SNPD.2016.7515909
 15. Service Selection Using Multi-criteria Decision Making : A Comprehensive Overview / [M. Hosseinzadeh, H. K. Hama, M. Y. Ghafour et al.] // *Journal of Network and Systems Management* – 2020. – Vol. 28, № 4. – P. 1639–1693. DOI: 10.1007/s10922-020-09553-w
 16. Yu H. Q. A modified Logic Scoring Preference method for dynamic Web services evaluation and selection / H. Q. Yu, and H. Molina // *Service Oriented Computing* : 2nd European Seminar for Young Researchers, Leicester, 11–12 June 2007 : proceedings. – Leicester, 2007. – P. 87–93.
 17. Yu H. Q. A Method for Automated Web Service Selection / H. Q. Yu, S. Reiff-Marganiec // *Services – Part I* : IEEE congress, Honolulu, 6–11 July 2008 : proceeding. – IEEE, 2008. – P. 513–520. DOI: 10.1109/SERVICES-1.2008.8
 18. Yu H. Q. Automated Context-Aware Service Selection for Collaborative Systems / H. Q. Yu, S. Reiff-Marganiec // *Advanced Information Systems Engineering* : 21st International Conference, CAiSE 2009, Amsterdam, 8–12 June, 2009 : proceedings. – Berlin : Springer, 2009. – P. 261–274. – (Lecture Notes in Computer Science, Vol. 5565). DOI: 10.1007/978-3-642-02144-2_23
 19. Yager R. R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking / R. R. Yager // *IEEE Transactions on Systems, Man, Cybernetics*. – 1988. – Vol. 18, № 1. – P. 183–190. DOI: 10.1109/21.87068
 20. Chakraborty S. A simulation comparison of normalization procedures for TOPSIS / S. Chakraborty, C.-H. Yeh // *Computers and Industrial Engineering* : International conference, Troyes, 6–9 July, 2009 : proceedings. – IEEE, 2009. – P. 1815–1820. DOI: 10.1109/iccie.2009.5223811
 21. Vafaei N. Data normalisation techniques in decision making: case study with TOPSIS method / N. Vafaei, R. A. Ribeiro, L. M. Camarinha-Matos // *International Journal of Information and Decision Sciences* – 2018. – Vol. 10, № 1. – P. 19–38. DOI: 10.1504/ijids.2018.090667
 22. Alinezhad A. Sensitivity Analysis in the QUALIFLEX and VIKOR Methods / A. Alinezhad, N. Esfandiari // *Journal of Optimization in Industrial Engineering* – 2012. – Vol. 6, № 1 (10). – P. 29–34.
 23. Muñoz B. Sensitivity Analysis of Multicriteria Decision Making Methodology Developed for Selection of Typologies of Earth-retaining Walls in an Urban Highway / B. Muñoz, M. G. Romana, J. Ordóñez // *Transportation Research Procedia* – 2016. – Vol. 18. – P. 135–139. DOI: 10.1016/j.trpro.2016.12.019
 24. Automotive style design assessment and sensitivity analysis using integrated analytic hierarchy process and technique for order preference by similarity to ideal solution / [G. Tian, H. Zhang, H. Jia et al.] // *Advances in Mechanical Engineering* – 2016. – Vol. 8, № 5. – P. 1–10. DOI: 10.1177/1687814016649885
 25. Web Services Quality Factors Version 1.0, Candidate OASIS Standard 01, October 2012 [Electronic resource]. – Access mode: <http://docs.oasis-open.org/wsrm/WS-Quality-Factors/v1.0/cos01/WS-Quality-Factors-v1.0-cos01.html>
 26. QWS Dataset v2.0, November 2019 [Electronic resource]. – Access mode: <https://github.com/qwsdata/qwsdata.github.io/releases>
 27. Polska O. V. Model of web services quality criteria hierarchy / O. V. Polska, R. K. Kudermetov, N. V. Shcherbak // *Вісник Запорізького національного університету. Фізико-математичні науки* – 2020. – № 2. – С. 43–51. DOI: 10.26661/2413-6549-2020-2-06
 28. Dujmovic J. J. Continuous Preference Logic for System Evaluation / J. J. Dujmovic // *IEEE Transactions on Fuzzy Systems* – 2007. – Vol. 15, № 6. – P. 1082–1099. DOI: 10.1109/TFUZZ.2007.902041
 29. Dujmović, J. Partial absorption function / J. Dujmović // *Journal of the University of Belgrade, EE Dept., Series Mathematics and Physics* – 1979. – № 659. – P. 156–163.
 30. Polska O. V. The approach for QoS based web service selection with user’s preferences / O. V. Polska, R. K. Kudermetov, V. V. Shkarupylo // *Наукові праці Донецького національного технічного університету, серія: «Проблеми моделювання та автоматизації проектування»* – 2020. – №16. – С. 19–27. DOI: 10.31474/2074-7888-2020-2-19-27
 31. A Theoretical Analysis of NDCG Ranking Measures / [Y. Wang, L. Wang, Y. Li et al.] // *Learning Theory* : 26th annual conference, COLT 2013, Princeton, 12–14 June, 2013 : proceedings. – JMLR.org, 2013. – P. 25–54.
 32. Fogli A. Exploiting semantics for context-aware itinerary recommendation / A. Fogli, G. Sansonetti // *Personal and Ubiquitous Computing* – 2019. – Vol. 23. – P. 215–231. DOI: 10.1007/s00779-018-01189-7
 33. Normalization Techniques Comparison for QoS-based Web Services Selection by LSP Method / [R. Kudermetov, O. Polska, V. Shkarupylo et al.] // *Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS)* : 5th IEEE International Symposium Smart and Wireless Systems within the Conferences, Dortmund, 17–18 September 2020: proceeding. – IEEE, 2020. – P. 213–216. DOI: 10.1109/idaacs-sws50031.2020.9297098
 34. QoS-aware Web-services Ranking: Normalization Techniques Comparative Analysis for LSP Method / [O. Polska, R. Kudermetov, J. A. J. Alsayaydeh et al.] // *ARNP Journal of Engineering and Applied Sciences* – 2021. – Vol. 16, № 2. – P. 248–254.

FORMALIZATION CODING METHODS OF INFORMATION UNDER TOROIDAL COORDINATE SYSTEMS

Riznyk V. V. – Dr. Sc., Professor, Professor of the Automatized Control System Department, Lviv Polytechnic National University, Lviv, Ukraine.

ABSTRACT

Contents. Coding and processing large information content actualizes the problem of formalization of interdependence between information parameters of vector data coding systems on a single mathematical platform.

Objective. The formalization of relationships between information parameters of vector data coding systems in the optimized basis of toroidal coordinate systems with the achievement of a favorable compromise between contradictory goals.

Method. The method involves the establishing harmonious mutual penetration of symmetry and asymmetry as the remarkable property of real space, which allows use decoded information for forming the mathematical principle relating to the optimal placement of structural elements in spatially or temporally distributed systems, using novel designs based on the concept of Ideal Ring Bundles (IRB). IRBs are cyclic sequences of positive integers which dividing a symmetric sphere about center of the symmetry. The sums of connected sub-sequences of an IRB enumerate the set of partitions of a sphere exactly R times. Two- and multidimensional IRBs, namely the “Glory to Ukraine Stars”, are sets of t -dimensional vectors, each of them as well as all modular sums of them enumerate the set node points grid of toroid coordinate system with the corresponding sizes and dimensionality exactly R times. Moreover, we require each indexed vector data “category-attribute” mutually uniquely corresponds to the point with the eponymous set of the coordinate system. Besides, a combination of binary code with vector weight discharges of the database is allowed, and the set of all values of indexed vector data sets are the same that a set of numerical values. The underlying mathematical principle relates to the optimal placement of structural elements in spatially and/or temporally distributed systems, using novel designs based on t -dimensional “star” combinatorial configurations, including the appropriate algebraic theory of cyclic groups, number theory, modular arithmetic, and IRB geometric transformations.

Results. The relationship of vector code information parameters (capacity, code size, dimensionality, number of encoding vectors) with geometric parameters of the coordinate system (dimension, dimensionality, and grid sizes), and vector data characteristic (number of attributes and number of categories, entity-attribute-value size list) have been formalized. The formula system is derived as a functional dependency between the above parameters, which allows achieving a favorable compromise between the contradictory goals (for example, the performance and reliability of the coding method). Theorem with corresponding corollaries about the maximum vector code size of conversion methods for t -dimensional indexed data sets “category-attribute” proved. Theoretically, the existence of an infinitely large number of minimized basis, which give rise to numerous varieties of multidimensional “star” coordinate systems, which can find practical application in modern and future multidimensional information technologies, substantiated.

Conclusions. The formalization provides, essentially, a new conceptual model of information systems for optimal coding and processing of big vector data, using novel design based on the remarkable properties and structural perfection of the “Glory to Ukraine Stars” combinatorial configurations. Moreover, the optimization has been embedded in the underlying combinatorial models. The favorable qualities of the combinatorial structures can be applied to vector data coded design of multidimensional signals, signal compression and reconstruction for communications and radar, and other areas to which the GUS-model can be useful. There are many opportunities to apply them to numerous branches of sciences and advanced systems engineering, including information technologies under the toroidal coordinate systems. A perfection, harmony and beauty exists not only in the abstract models but in the real world also.

KEYWORDS: elegant symmetry and asymmetry ensemble, relationship, information redundancy, combinatorial configuration, optimum vector data coding, code size, basis, trade-off.

ABBREVIATIONS

BVD is a big vector data;
ESAE is a Elegant Symmetry and Asymmetry Ensemble;
IRB is an Ideal Ring Bundle;
GUS is a “Glory to Ukraine Star” configuration;
OLAP is an online analytical processing;
OSR is an Optimum Structural Relationship;
GF is a Galois field;
PG is a projective geometry;

NOMENCLATURE

α_{\min} is a minimal angular distance;
 α_{\max} is a maximal angular distance;
 k_{it} is an i -th integer of t -tuple;
 m_i is a number of categories of i -th attribute, as well as a number of reference points on i -th ring axis in a toroidal coordinate system;

© Riznyk V. V., 2021
DOI 10.15588/1607-3274-2021-2-15

N is a general number of central angles multiple of α_{\min} ;
 n is a capacity of toroidal coordinate system basis;
 P_{\max} is a code size of t -dimensional vector data array;
 R is a number coding of the same vector data array;
 S is an order of rotational symmetry, a number of star-code combinations, as well as a number of nodal points in toroid coordinate system;
 t is a dimension of vector data array, number of attributes, as well as a number of significant digits of t -dimensional code.

INTRODUCTION

One of the key directions for development of modern information technologies is the design of multidimensional control systems of technical and technological objects, management of multidimensional systems and exchange of big vector data in information and communication networks, coding of information sources and other

practically important scientific and technical problems of today.

The object of study is the process of vector data coding and processing signals by t attributes and m_i categories of the data under toroid coordinate systems. Unlike customary coding designs, the vector data puts it in basis of the t -dimensional toroid coordinate system, where basis is the smallest of the general number coordinate set, which it creates taking m_i modulo addition. This scientific method reduces to the indexing t -dimensional data in the basis of t -dimensional toroidal reference system, where the basis is a set of less than the total number of coordinates of this coordinate set. Theoretically, is substantiated the existence of an infinitely large number of minimized basis, which give rise to numerous varieties of multidimensional "star" coordinate systems to find practical applications in modern and future multidimensional information technologies. Therefore, the proposed approach opens prospects for the development of the innovative information technologies grounded on the minimization of the basic structure of multidimensional coding systems and processing of big vector data in toroidal reference systems.

The subject of study is vector data coding methods of information under toroid coordinate systems. The known vector data coding methods [1–12] characterized by limited number of attributes and categories for signal coding and processing information encoded by two or more attributes and categories, as well as are high complexity.

The propose of the work is improving the efficiency of vector data coding methods and processing information encoded by two or more attributes and categories simultaneously.

1 PROBLEM STATEMENT

Mathematical problem in straight and inverse staging is reduced to establishing a mutually unambiguous display of vector binary code combinations sets according to vector data attribute-categories sets on the coordinate grid of the t -dimensional surface of the toroid.

The task is to increase the number of code combinations of t -dimensional binary code for the formation of information parameters of signals by the number of attributes and categories in the basis of the outlined t -dimensional coordinate system. We require the code combinations enumerate the node points set of the toroid coordinate grid with corresponding size and dimensionality.

The input variables are the number t attributes and the number of m_i indexed categories ($i = 1, 2, \dots, t$) of the attributes.

The output variable is the maximum number of binary code combinations taking account the number t attributes, and the number of m_i indexed categories ($i = 1, 2, \dots, t$) of the attributes and the number R of different ways of encoding the equally indexed sets "attribute-category" in the basis of the outlined t -dimensional toroid coordinate system.

We require the set of the underlying vector data one-to-one corresponds to set of nodal point coordinates of t -

dimensional toroidal coordinate system, using a sub-set of the set, and their combined sums for cover the set of the outlined system with the size of $m_1 \times m_2 \times \dots \times m_i$, the set of coordinates, which correspond to the t -sets of information parameters, and the number R of different ways of encoding the equally indexed sets "attribute-category" in the basis of the t -dimensional coordinate system of given sizes and dimensionality. The criterion to be, for example, increasing the level of data protection, crypto graphical and system security, while maintaining on code size looking for a favorable compromise between the contrary objectives.

There are other options for setting the task in theoretically defined limits of changing input and output variables. Thus formats of vector data and codes should be jointly designed to reach the best trade-off between performance and complexity for optimum processing information under toroid coordinate systems.

Therefore, it is advisable to formalize the relationships between the information parameters of vector data encoding systems and theoretical limitations for optimal solution of specific problems under toroid coordinate systems.

2 REVIEW OF THE LITERATURE

The main goal of modern information technology is development of effective BVD coding and processing, including large data amounts analysis. However, the design based on the traditional theory is not always applicable for multidimensional data processing. BVD often involve a number of factors, such as national defense, economic, and other indexes of infrastructure, which have led to difficulties in large-scale data processing. In recent times, a great number of new concepts, parallel algorithms, processing tools, platform, and applications are suggested and developed to improve the value of BVD [1–12]. Big spatial vector data management presented in global review [1]. A comprehensive survey of the existing techniques and technologies from both academia and industry given in [2]. The papers [3–5] present prospects and problems of BVD for distant sensing. A techniques for compose of a map procedure, which performs filtering, sorting, and a summary operations of BVD presented at the IEEE International Conferences on Data Engineering [6]. Developing a reversible rapid coordinate transformation BVD model for the cylindrical projection, we see in paper [7]. The paper [8] contains fast multidimensional ensemble empirical mode decomposition for the analysis of big spatial-temporal datasets. The geometric computing algorithms are always very complex and time-consuming, which makes big spatial data processing very slow, or even impossible [9]. A framework that couples cloud and high-performance computing for the parallel map projection of vector-based big spatial data regarded in [10, 11]. In the United States, large payer data amassed to explore large data to advance knowledge discovery in nursing methodologies, clinical trials and lab research [12]. The idea of topological coordinates for toroidal chemical structures is in agreement with describing the

physics of toroidal confined plasmas [13]. The projection provides large-scale spatial modeling of BVD under a common coordinate system. However, algorithmic complexity of the map projections represents a pressing computational challenge.

Modern theory of combinatorial configurations are such spatial structures as perfect difference sets [14, 17], algebraic constructions based on cyclic groups in extensions of Galois fields [15], manifolds [16], structures connecting algebra through geometry [17]. Many scientists also have suggested that the entire universe is a torus. This notion used to help visualize aspects of higher dimensional toroidal spaces. It was the torus mathematical model useful for describe geometric objects in spatial coordinates. The torus topology is superior to geometry for describing such objects because relate with philosophical spatial relationships. In one dimension, a usual single-holed torus is the 1 – torus as a ring shape object. In two dimensions, we see a usual torus, also called the 2 – torus. In analogy with this concept, in three and more dimensions, the t -dimensional torus, or t – manifold is an object that exists in dimension $t + 1$. Regarding the remarkable properties and structural perfection of two-and three-dimensional numerical rings [18], models of optimal multidimensional coding systems [19], and GUS combinatorial configurations [20], easy to see their toroidal topological structure.

Research into the underlying coding methods under toroid coordinate systems provide an ability optimal coding of multidimensional signals and objects, decreasing information redundancy in increasing code size of vector data coding under limitation on word length and geometric dimensions of toroid coordinate grid for trade-off processing vector data arrays of given number attributes and categories. Therefore, the formalization of these relationships is of very need, because make it possibility finding optimal solution by required criterion.

3 MATERIALS AND METHODS

“Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty and perfection.” H. Weyl.

Symmetry and asymmetry relation in geometric structure is the most familiar type of them. The more general meaning of symmetry-asymmetry is in combinatorial configurations as a whole. In this context, symmetries and asymmetries underlie some of the most profound results found in modern physics, including aspects of space and time [21]. Finally, discusses interpenetrating symmetry and asymmetry in the humanities, covering its rich and varied use in architecture, philosophy, and art. Space-time symmetries are features of space-time that described as exhibiting some form of symmetry [22].

Let us define the quant interval α_{\min} for finding basic harmonious relationship between S – fold rotational symmetry and embedded into the symmetry asymmetry (Fig. 1).

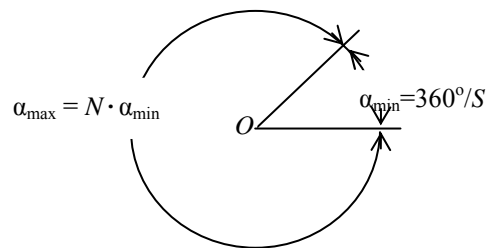


Figure 1 – A chart for define the quant interval α_{\min} for finding basic harmonious relationship between S – fold rotational symmetry and embedded into the symmetry asymmetry

Definition. The S – old rotational symmetry, which provides ability to split it into pairs of mutually complementary asymmetries, where we require a set of all $N = (S-1)R = (n-1)n$ angular distances $[\alpha_{\min}, \alpha_{\max}]$. A set of angular spaces between everything possible lines that diverge from common center O allows an enumeration of integers $[1, S-1]$ exactly R -times, is the Elegant Symmetry and Asymmetry Ensemble (ESAE) [19].

For example, the 7-fold ($S = 7$) rotational symmetry is the ESAE because splits into two mutually complementary asymmetries cyclically in ORS spatial proportions as both $\{1 : 2 : 4\}$ for the first, and the $\{1 : 1 : 3 : 2\}$ – the second ones clockwise, where $\alpha_{\min} = 360^\circ/7$, $\alpha_{\max} = 6\alpha_{\min}$. In turn, the first of them is the IRB $\{1, 2, 4\}$ with $n = 3$, $R = 1$, while the second is the IRB $\{1, 1, 3, 2\}$ with $n = 4$, and $R = 2$.

From the Fig. 1 and underlying definition follows formula (1):

$$S - 1 = \frac{n \cdot (n - 1)}{R} \quad (1)$$

Note, the role of symmetry in physics is important in simplifying solutions to many problems, e.g. exact solutions of Einstein’s field equations of general relativity [23], and study of isometrics in two or three-dimensional Euclidian space [24]. Only one angular interval in one-fold ($S = 1$) rotational symmetry enumerates the set $\{1\}$ exactly once ($n = 1$) is singleton, known as a unit set [25].

The idea of “perfect” systems comes from remarkable properties of symmetry in combinatorics [14], and nature laws [21, 22]. Let us regard the n -stage ring sequence of two-dimensional ($t = 2$) vectors with non-negative integers $\{(k_{11}, k_{12}), (k_{21}, k_{22}), \dots, (k_{i1}, k_{i2}), \dots, (k_{n1}, k_{n2})\}$, which form two-dimensional coordinate grid of sizes $m_1 \times m_2$ over of a torus, taking modulo m_1 , and modulo m_2 sums accordingly. The coordinate grid basis is “Glory to Ukraine Star” (GUS) combinatorial configurations [20].

Because “perfect” toroidal codes” created by GUS configurations, we call their, shortly, star-codes.

Planar projection of spatially disjointed the ring axes $m_1 = 2$ and $m_2 = 3$ coordinate points of two-dimensional ($t = 2$) toroid grid $m_1 \times m_2 = 2 \times 3$ with common reference point $(0,0)$ given in Fig. 2.

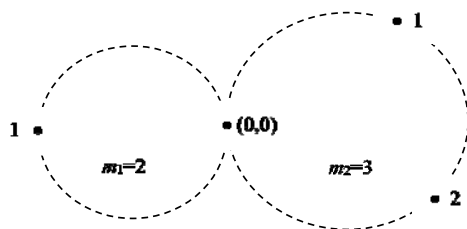


Figure 2 – Planar projection of spatially disjointed the ring axes $m_1=2$ and $m_2=3$ coordinate points of two-dimensional ($t=2$) toroid reference grid $m_1 \times m_2=2 \times 3$ with common reference point $(0,0)$

Hence, we have designed toroid coordinate system for vector data coding two ($t=2$) attributes with two ($m_1=2$) categories of the first, and three ($m_2=3$) for the second attribute accordingly under star-code basis.

Next, we consider a more general model of the t -dimensional toroid coordinate system for vector data coding more of two attributes with arbitrary number of categories under star-code basis.

Clearly, a t -dimensional toroid coordinate system designed for vector data coding t attributes and m_i categories of each of them ($i=1,2,\dots,t$) requires t concurrent disjointed axes $m_1, m_2, \dots, m_i, \dots, m_t$ with common reference point for forming t -dimensional coordinate grid of the system with sizes $m_1 \times m_2 \times \dots \times m_t$. So, the underlying multidimensional information toroid coordinate system can be described by parameters S, n, R, t, m_i ($i=1, 2, \dots, t$). Here t is dimension of vector data array, number of attributes, and number of significant digits of t -dimensional code, m_i is a number of categories of i -th attribute, and number of reference points on i -th ring axis in a toroidal coordinate system. Besides, information about vector data array depends of geometric sizes $m_1 \times m_2 \times \dots \times m_t$ of the toroid coordinate system.

The harmonious relationship between S – fold ESAE behavior is in agreement with classic combinatorial theory [14] based on algebraic group theory [17], and expands on the other combinatorial configurations [18–20] well.

Theorem. The maximum code size P_{\max} of n -digit binary t – dimensional star-code interconnected of numeric values $R, m_1, m_2, \dots, m_1, \dots, m_t$ by the formula (2):

$$P_{\max} = R \cdot \prod_1^t m_i \leq 2^n - 1, \quad (2)$$

$$(m_1, m_2, \dots, m_i, \dots, m_t) = 1.$$

Proof. The code size of n -digit binary code cannot be greater than the number of nonzero binary code combinations formed by it.

Corollary 1. The maximum code size P_{\max} of n -digit binary t -dimensional star-code is equal to the order S of the rotating symmetry, which spatial field creates the toroid coordinate grid with sizes $m_1 \times m_2 \times \dots \times m_i \times \dots \times m_t$.

Corollary 2. If the difference between the numerical values of the left and right parts of inequality (2) reduce,

the star-code size increases, but security of the code is lost, and vice versa.

Corollary 3. Increasing the sizes of the t -dimensional toroidal coordinate grid allows intensification the converting form of vector data information with star-code, expanding the possibility of spatial reconstructing the coordinate system by combining the ratio between numerical values of interconnected formula (2).

Corollary 4. The star-code information capacity is the bigger than code size of it.

Application of the underlying relationships between the information parameters of vector data encoding systems and theoretical limitations for optimal solution of specific problems under toroid coordinate systems will make it possible to reach the best trade-off between performance and complexity for optimum processing information under toroid coordinate systems.

4 EXPERIMENTS

Experimental verification of the effectiveness of this information technology relates to optimization of vector data coding and processing under toroidal coordinate system, using novel designs based on the remarkable properties of ESAE, and the generalization of these methods and results to the improvement of a larger class of information systems.

The basic ideas of BVD processing under toroidal coordinate system are as following:

- determine sizes of toroidal coordinate system and its dimensionality accordingly to entity-attribute-value of a BVD list;
- make indexing entity-attribute-value list under the toroidal coordinate system;
- fetch from an information base applicable vector star-code with respect to computer power and processing program;
- make Big Vector Data processing under the star toroidal coordinate system.

The efficiency is that the less number of encoded signals use for given vector data processing than at present, while maintaining on the other significant operating characteristics of the coding system underlying methods provide opportunities to apply them to configure suitable relation big vector data models, e.g., using t attributes, and m_i categories for i -th attribute, $i=1, 2, \dots, t$ of the model.

Let us form toroid coordinate system for vector data coding two ($t=2$) attributes with two ($m_1=2$) categories of the first, and three ($m_2=3$) for the second attribute accordingly under star-code basis. For example, 2-tuple set of three ($n=3$) integer of the IRB $\{1, 2, 4\}$ forms modulo-2 and modulo-3 accordingly the basis $\{(1,1), (0,2), (0,1)\}$ of two-dimensional ($t=2$) toroid coordinate system as follows: $(0,0) \rightarrow (0,2) + (0,1)$; $(1,0) \rightarrow (1,1) + (0,2)$; $(1,2) \rightarrow (0,1) + (1,1)$. Take the $\text{mod}_1(n-1) = \text{mod}_1 2$ sum of the ring axis $m_1=2$, and the $\text{mod}_2 n = \text{mod}_2 3$ sum of the ring axis $m_2=3$. After arranged all vectors, we derive the coordinate grid $m_1 \times m_2=2 \times 3$ over a torus based

on the GUS $\{(1,1), (0,2), (0,1)\}$ exactly once ($R=1$) (Table 1).

Table 1 – Forming the coordinate grid $m_1 \times m_2 = 2 \times 3$ based on the GUS $\{(1,1), (0,2), (0,1)\}$

№	Forming the reference grid $m_1 \times m_2 = 2 \times 3$			
	References	Summation the mod ₁ 2, and mod ₂ 3		
		(1,1)	(0,2)	(0,1)
1	(0,0)		+	+
2	(0,1)			+
3	(0,2)		+	
4	(1,0)	+	+	
5	(1,1)	+		
6	(1,2)	+		+

From Table 1, easy to see the GUS $\{(1,1), (0,2), (0,1)\}$ forms the coordinate grid 2×3 over torus surface. The set of 2D coordinate grid one-to-one corresponds to a set of 2D vector data attributes, the first numerical index indicates vector data by two ($m_1=2$) categories, while the second – three ($m_2=3$) of ones by the same vector. The underlying property allows on 2D vector data processing over the 2×3 range concurrently.

Reasoning along similarly, we form toroid coordinate system for 2-tuple set of four ($n=4$) integer for the IRB $\{1, 1, 3, 2\}$ with $n=4, R=2$. We use the basis $\{(1,1), (1,1), (1,0), (0,2)\}$ of 2D ($t=2$) toroid coordinate system: $(0,0) \rightarrow (1,1) + (1,0) + (0,2)$, and $(0,0) \rightarrow (1,0) + (0,2) + (1,1)$; $(0,1) \rightarrow (1,1) + (1,0)$, and $(0,1) \rightarrow (0,2) + (1,1) + (1,1)$; $(0,2) = (0,2)$, and $(0,2) \rightarrow (1,1) + (1,1)$; $(1,0) = (1,0)$, and $(1,0) \rightarrow (0,2) + (1,1)$; $(1,2) \rightarrow (1,0) + (0,2)$, and $(1,2) \rightarrow (1,1) + (1,1) + (1,0)$. After arranged all vectors, we derive the coordinate grid $m_1 \times m_2 = 2 \times 3$ over torus surface based on the GUS $\{(1,1), (1,1), (1,0), (0,2)\}$ exactly twice ($R=2$) as shows in the Table 2.

Table 2 – Forming the coordinate grid $m_1 \times m_2 = 2 \times 3$ over torus based on the GUS $\{(1,1), (1,1), (1,0), (0,2)\}$

№	Forming the reference grid $m_1 \times m_2 = 2 \times 3$				
	References	Summation the mod ₁ 2, and mod ₂ 3			
		(1,1)	(1,1)	(1,0)	(0,2)
1	(0,0)		+	+	+
2	(0,0)	+		+	+
3	(0,1)	+			+
4	(0,1)	+	+	+	
5	(0,2)			+	+
6	(0,2)	+	+		+
7	(1,0)			+	
8	(1,0)		+		+
9	(1,1)	+			
10	(1,1)		+		
11	(1,2)			+	+
12	(1,2)	+	+		+

Generally, each i -th t -dimensional vector $K_i = (k_{i1}, k_{i2}, \dots, k_{it})$ of t -dimensional sequence $\{K_1, K_2, \dots, K_b, \dots, K_n\}$

is a t -stage of non-negative integers as sub-sequence of the sequence, a set of ordered vector data attributes are indicated accordingly to t categories of the BVD processing under t -dimensional toroidal coordinate system over coordinate grid $m_1 \times m_2 \times \dots \times m_t$.

5 RESULTS

Fragments of the results of conducted experiments presented in the Tables 3–6.

Let us consider some specific examples using the methods for coding vector data sets of a particularly defined value in the basis of toroidal coordinate systems with corresponding geometric dimensions in the formalization (2).

The Table 3 contains six ($n^2 - n = 6$) 3-digit ($n = 3$) binary 2D ($t = 2$) star-code combinations, which interrelated with numerical values of indexed vector data in the a priori of acceptable limits of their changes (2). In the Table 3 shows that the code size P_{\max} of 3-digit ($n = 3$) binary 2D ($t = 2$) star-code comes near theoretically maximum, as basis of the toroid coordinate grid with sizes $m_1 \times m_2 = (n-1)n = 2 \times 3$ created under the first embedded into of the ESAE asymmetry, namely the GUS $\{(0,1), (0,2), (1,0)\}$.

Table 3 – The forming 3-digit 2D star-code $\{(0,1), (0,2), (1,0)\}$ under torus coordinate system 2×3

№	The 3-digit 2D star-code under grid $m_1 \times m_2 = 2 \times 3$			
	Vector data	Digit weights of the binary star-code		
		(0,1)	(0,2)	(1,0)
1	(0,0)	1	1	0
2	(0,1)	1	0	0
3	(0,2)	0	1	0
4	(1,0)	0	0	1
5	(1,1)	1	0	1
6	(1,2)	0	1	1

The Table 4 contains 12 binary 2D ($t = 2$) 4-digit ($n=4$) combinations ($n^2 - n = 12$) for coding two attributes ($t = 2$) both with two ($m_1 = 2$) category of the first, and three ($m_2 = 3$) – the second attribute concurrently.

Table 4 – The fragment of forming 4-digit 2D star-code $\{(1,1), (1,1), (0,2), (1,0)\}$ under torus coordinate system 2×3

№	The 4-digit 2D star-code under grid $m_1 \times m_2 = 2 \times 3$				
	Vector data	Digit weights of the binary star-code			
		(1,1)	(1,1)	(0,2)	(1,0)
1	(0,0)	0	1	1	1
2	(0,0)	1	0	1	1
3	(0,1)	1	0	0	1
4	(0,1)	1	1	1	0
.....					
14	(1,2)	0	0	1	1
15	(1,2)	1	1	0	1

In the Table 4 shows that the code size $P(n) = 2^n - 1 = 15$ of 4-digit ($n = 4$) binary 2D ($t = 2$) star-code with $n = 4$ and $R = 2$. Note, the basis of the toroid coordinate grid created under the second embedded into of the ESAE asymmetry, namely the GUS $\{(1,1), (1,1), (1,0), (0,2)\}$.

Table 5 illustrates fragment of forming 4-digit 2D star-code $\{(1,2), (2,4), (1,3), (2,1)\}$ under torus coordinate system 3×5 . In the Table 5 shows that the code size $P(n) = 2^n - 1 = 15$ of 4-digit ($n = 4$) binary 2D ($t = 2$) star-code with $n = 4$ and $R = 1$. Note, the basis of the torus coordinate grid created under the GUS $\{(1,2), (2,4), (1,3), (2,1)\}$.

Table 5 – The fragment of forming 4-digit 2D star-code $\{(1,2), (2,4), (1,3), (2,1)\}$ under torus coordinate system 3×5

№	The 4-digit 2D star-code under grid $m_1 \times m_2 = 3 \times 5$				
	Vector data	Digit weights of the binary star-code			
		(1,2)	(2,4)	(1,3)	(2,1)
1	(0,0)	1	1	1	1
2	(0,1)	1	1	0	0
3	(0,2)	0	1	1	0
4	(0,3)	1	0	0	1
.....
14	(2,3)	0	1	1	1
15	(2,4)	0	1	0	0

In the Table 6 shows forming 5-digit 2D star-code $\{(1,1), (1,4), (2,2), (2,2), (2,3)\}$ under torus coordinate system 3×5 . To see this we observe that the code size $P(n) = 2^n - 1 = 31$ of 5-digit ($n = 5$) binary 2D ($t = 2$) star-code with $n = 5$ and $R = 2$. Note, the basis of the torus coordinate grid created under the GUS $\{(1,1), (1,4), (2,2), (2,2), (2,3)\}$.

Table 6 – The fragment of forming 5-digit 2D star-code $\{(1,1), (1,4), (2,2), (2,2), (2,3)\}$ under coordinate system 3×5

№	The 5-digit 2D star-code under grid $m_1 \times m_2 = 3 \times 5$					
	Vector data	Digit weights of the binary star-code				
		(1,1)	(1,4)	(2,2)	(2,2)	(2,3)
1	(0,0)	1	1	0	1	1
2	(0,1)	1	1	1	0	1
3	(0,2)	0	1	1	0	0
4	(0,3)	0	1	0	1	0
.....
30	(2,3)	0	1	0	1	1
31	(2,4)	0	1	1	0	1

From the Tables 3–6 follows that the underlying formula (2) make it possible to reach the best trade-off between performance and complexity for optimum processing of information under toroid coordinate system, while maintaining or improving on information capacity of the system.

The equations (1) and (2) formalize a large class of t -dimensional codes from non-redundant vector codes of high performance [19] to coding self-correcting vector

data signals with faster than classic codes transmission of multidimensional information by noise communication channels [20]. One of them is optimum binary monolithic star-code for processing of two-or multidimensional vector signals under toroidal coordinate system. The monolithic star-code forms code combinations as a sequence of solid connected symbols “1” or “0” in the code combination [19, 20]. This property makes it possible to use self-correcting vector data signals for faster than classic codes transmission in a noise channel, using limited number of bits.

From the underlying formalization (1) and (2) have been defined the next types of the optimum weighed n -digital binary ring monolithic codes based on IRBs [20]:

a) numerical code is a set of ring n -sequences of positive integers which sums of connected sub-sequences of the set enumerate the set of integers $[1, N]$ exactly R -times, $N = n(n-1)/R$;

b) two-dimensional code is a set of ring n -sequences of 2-tuple nonnegative integers which modular sums of connected 2-tuple taken modulo m_1 and m_2 , respectively, allows an enumeration of nodal points of coordinate grid $m_1 \times m_2$ over 3-torus exactly R -times with frame of axes $m_1 \cdot m_2 = n(n-1)/R$;

c) multidimensional code is a set of ring n -sequences of t -tuple nonnegative integers which modular sums of connected t -tuple taken modulo m_1, m_2, \dots, m_t , respectively, allows an enumeration of nodal points of coordinate grid $m_1 \times \dots \times m_t$ over $(t+1)$ -torus exactly R -times with frame of axes $m_1 \cdot m_2 \dots \cdot m_t = n(n-1)/R$.

6 DISCUSSION

As it evident, the Table 3 demonstrates the advantages of two-dimensional ($t = 2$) binary vector data coding in the minimized 3-digit ($n = 3$) 2D database of the toroid coordinate system a reference grid with sizes $m_1 \times m_2 = 2 \times 3$. The mutually unambiguous compliance with a set of indexed data “attribute-categories” of a set of binary vector code combinations formed by this database have been achieved in the system. In turn, it was possible due to reducing the natural redundancy in the system.

Instead, Table 4 gives example of the introduction of artificial redundancy by increasing to two ($R = 2$) the number of ways to cover the grid for increasing 2 times security of the vector data, due to increment a register capacity by one: from $n = 3$ to 4. According to the formula (1) the sizes of reference grid are the same that for the parameters $S = 7, n = 3, R = 1$ because $S = 7, n = 4, R = 2$, and consequently $m_1 \times m_2 = 2 \times 3$ also.

Reasoning along similarly, in the Tables 5 and 6 show methods of vector data coding design in the minimized database of the toroid coordinate system a reference grid with sizes $m_1 \times m_2 = 3 \times 5$.

In the formula (2), the underlying rule displayed by the number R of methods of covering all node points of t -dimensional outlined toroidal coordinate system for encoding of indexed vector data “attribute-categories” sets with t attributes and appropriated number of categories in the set. Besides, using n -digit base reconstruction of t -

dimensional binary vector provide code we the rule: repetition is a loss of information in order to increase so-many times security of vector data coding under the coordinate system. In addition, there are elegant ensembles of star-codes, which allow you to form different sets of allowed code combinations on a set of modular sums of the base vectors of the toroid coordinate system according to several ring schemes with cross routes of their bypass ($R = 1$) or more ($R > 1$) in ways [18–20]. From the formula (1) and (2) follows increasing the number of allowed code combinations of the coding method and improving the characteristics of vector data encoding in toroid coordinate systems by performance and interference, as well as the star-code information capacity increases faster than code size of the code with growing dimensionality and number of its binary digits.

The study aims at patterns of transformation of flat harmoniously conjugated asymmetric structures of rotating symmetry into spatial structural forms of higher dimensions, where they take the form of two-and multidimensional toroid coordinate systems formed by a minimized binary code basis with vector weight digits. Such combinatorial configurations differ favorably from the classical ones [14–18] of spatial forms with central, mirror or axial symmetries, forming families on a set of the ESAEs with varieties of symmetrical shape [19].

The growth of vector code combination by one bit doubles the code size of the encoding method in the outlined toroid coordinate system with the corresponding dimensions and dimensions. The number of indexed attributes and categories in the form of any long a priori of integer t -tuple allows one code word to encode, forward and process in the basis of the system simultaneously as many signs of vector data as symbols contained in the t -tuples. It respectively provides increasing the performance of the system at t times, where t is the number of ring axes of the toroid coordinate grid.

Formalization allows you to reach a balanced compromise on contradictory goals related to the power and reliability of the information method. It outlines theoretically a large-scale information model of harmoniously built of multidimensional geometric space as a hypothetical system of a perfectly structured source of information as t -dimensional locked spheres. This system has a priori infinitely large number of sets of coordinate sub-systems of quantum “density”, generated by a minimized basis of t -dimensional binary n -bit code with compression coefficient value approaching to $2^n/n$.

Formalization of methods of optimal encoding of multidimensional data arrays in spatial coordinate systems is connected not only with the establishment of mathematical connection of information parameters with harmoniously organized symmetrical and asymmetric spatial structures, but also with the processing of a set of philosophical, methodological, scientific and applied, educational problems of analysis and synthesis of multidimensional information systems.

Among other results, it is interesting to observe the behavior of mathematical models of geometric objects

with non-uniform structure [18–20]. These objects can be in progress for originate of such geometric structures already in the higher dimensions, with the advent of new elegant forms of spatial configurations enriched with numerous varieties of thin structure multiplied on growing of their combinatorial varieties. It follows an intuitive conclusion about the hypothetical model of toroidal geometric structure in the behavior of natural and man-made objects in the real world [21, 22].

CONCLUSIONS

The urgent problem of coding and processing large information content solved, using formalization of interdependence between information parameters of vector data coding systems on a single mathematical platform. The study of the theoretical relationship between the detected pattern of harmonious conjuring of symmetrical and asymmetrical structures, the basic principle of optimal structural relations (OSR) was established. It reflects the essence of the proposed concept of converting multidimensional form of information with binary code in a structured field of toroid coordinate systems of corresponding dimensions and dimensionalities for provide optimum coding and processing vector data arrays with numbering attributes and categories under minimized basis of outlined t -dimensional toroidal coordinate system, using the underlying formalization (1), (2).

The scientific novelty of obtained results is the formalization of interdependence between information parameters of vector data coding systems on a single mathematical platform. This approach allows forming the outlined t – dimensional toroidal coordinate system, using smaller as all number of coordinates set basis of t -tuples. This, in turn, provides optimum coding and processing vector data arrays with numbering attributes and categories under minimized basis of outlined t -dimensional toroidal coordinate system by holding criterion and limitations to achieve a favorable compromise between contradictory goals in changing of computational values within of theoretically defined by formulas (1) and (2) boundaries. The upper limit of the information capacity encoding method for given code size and numbering attributes, and categories of vector data sets, as well as R of various ways of coding the same “attribute-category” sets defined.

Prospect for further research are the development of new direction in multidimensional systems engineering, for improving such quality indices as information capacity, reliability, transmission speed, positioning precision, and ability to reproduce the maximum number of combinatorial varieties in the system with a limited number of elements and bonds, using remarkable properties and structural perfection of GUS combinatorial configurations. The experimental results allow recommending the proposed methodology for direct applications to information and computational technologies, telecommunications, radio-and electronic engineering, radio-physics, and other engineering areas, as well as in education. These design techniques allows configure optimum two-and multidimensional vector data coding system, using innovative

methods based on the underlying combinatorial models offering ample scope for progress in systems engineering, cybernetics, and industrial informatics.

Physical results – a better understanding the role of geometric structure in the behavior of natural and man-made objects.

The existence of an a priori of an infinitely large number of minimized basis, which give rise to numerous varieties of multidimensional “star” coordinate systems, opens up new possibilities for solving a wide range of mathematical and applied problems of computer science, cybernetics, and management on the platform of system mathematics. We take into account the developments of modern theory of systems as a set of philosophical, methodological and scientific, and applied problems of analysis and synthesis of multidimensional systems.

A perfection and harmony exists not only in the abstract models but in the real world also.

ACKNOWLEDGEMENTS

This work involves the results performed in 2012–2017 years in Automated Control Systems Department of Lviv Polytechnic National University. I grateful to our colleagues for their active participation in support and understanding important significance in development of fundamental and applied research in vector information technologies, using design based on multidimensional combinatorial optimization of the technology. The basic results of the research presented in completed work on the R&S project “Designing Software for Vector Data Processing and Information Protection Based on Combinatorial Optimization” (State registration 0113U001360). State account number 0218U000988. All authors declare that no have financial support.

REFERENCES

1. Xiaochuang Y., Guoqing Li Big spatial vector data management: a review, *Big Earth Data*, 2018, Vol. 2, No. 1, pp. 108–129. DOI: 10.1080/20964471.2018.1432115
2. Bose N. R. Applied Multidimensional Systems Theory. Pennsylvania State University State College, PA, USA: Springer Int. Publishing AG, 2017, 192 p.
3. Chi M., Plaza A., Benediktsson J. A., Chi M., Plaza A., Benediktsson J. A., Sun Z., Shen J., and Zhu Y. Big data for remote sensing: Challenges and opportunities, *IEEE*, 104 (11), pp. 2207–2219, DOI:10.1109/Jproc.2016.598228
4. Ma Y., Wu H., Wang L., Huang B., Ranjan R., Zomaya A., Jie W. Remote sensing big data computing: Challenges and opportunities, *Future Generation Computer Systems*, 2015, No. 51, pp. 47–60. DOI:10.1016/j.future.2014.10.029
5. Pekturk M. K., Unal M. A review on real-time big data analysis in remote sensing applications, *25th Signal Processing and Communications Applications Conference (SIU)*. Antalya, Turkey, 15–18, May 2017.
6. Eldawy A., Mokbel M. F., Alharthi S., Alzaidy A., Tarek K., Ghani S. SHAHED A MapReduce-based system for querying and visualizing spatio-temporal satellite data, *IEEE International Conference on Data Engineering*. Seoul, South Korea, 13–17 April 2015.
7. Ye S., Yan T., Yue Y., Lin W., Li L., Yao X., Zhu D. Developing a reversible rapid coordinate transformation model for the cylindrical projection, *Computers & Geosciences*, 89, pp. 44–56. DOI:10.1016/j.cageo.2016.01.007
8. Wu Z., Feng J., Qiao F., Tan M. Z. Fast multidimensional ensemble empirical mode decomposition for the analysis of big spatio-temporal datasets, *Philos Trans A Math Phys Eng Sci.*, 2016, 374(2065), 2015.01.97
9. Ray S., Simion B., Brown A. D., Johnson R. A parallel spatial data analysis infrastructure for the cloud, *ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 5–8 November 2013. Orlando, FL, USA.
10. Tang W., Feng W. Parallel map projection of vector-based big spatial data: Coupling cloud computing with graphics processing units, *Computers, Environment and Urban Systems*, 2014, Vol. 61, pp. 187–197. DOI:10.1016/j.compenvurbusys.2014.01.001
11. Eldawy A., Mokbel M. F., Jonathan C. Hadoop Viz: A MapReduce framework for extensible visualization of big spatial data, *The 32nd IEEE International Conference on Data Engineering*, 16–20 May 2016. Helsinki, Finland, IEEE, 2016.
12. Connie DeLaney W., Charlotte Weaver A., Judith Warren J., Thomas Clancy R., Roy L. Simpson Big Data-Enabled Nursing: Education, Research and Practice. Cham, Switzerland, Springer, 2017, 488 p. <https://www.bookdepository.com/Big-Data-Enabled-Nursing-Connie-White-Delaney/9783319532998>
13. Laszlo I., Rassat A., Fowler P. W., Graovas A. Topological coordinates for toroidal structures, *Chemical Physics Letters: Elsevier Science B. V*, 2001, Vol. 342, pp. 369–374.
14. Hall Jr. Combinatorial Theory, 2 edition, Wiley-Interscience, 1998, 464 p. DOI: 10.1002 / 9781118032862
15. Rotman Joseph. Galois Theory, Second ed.: Springer, 1998, 106 p. DOI:10.1007/978-1-4612-0617-0
16. Lee John M. Introduction to topological manifolds. Springer-Verlag, New York, Inc., 2000, 395 p. DOI: 10.1007 / 978-1-4419-7940-7
17. Moore E. H., Pollatsek H. S. Difference Sets, *Student mathematical library Volume 67: Connecting Algebra, Combinatorics, and Geometry*, AMS, 2013, 298 p. ISBN 978-0-8218-9176-6
18. Riznyk W. Application of the Golden Numerical Rings for Configure Acoustic Systems of Fine Resolution, *Acta Physica Polonica A.*, 2011, Vol. 119, pp. 1046–1049. DOI: 10.12693/APhysPolA.119.1046
19. Riznyk V. Multi-modular Optimum Coding Systems Based on Remarkable Geometric Properties of Space, *Advances in Intelligent Systems and Computing*. Springer, 2017, Vol. 512, pp. 129–148. DOI 10.1007/978-3-319-45991-2_9
20. Riznyk V. V. Multidimensional Systems Optimization Developed from Perfect Torus Groups, *International Journal of Applied Mathematics and Informatics*, 2015, Vol. 9, pp. 50–54.
21. Weyl Hermann. Symmetry. New Jersey, Princeton University Press, Princeton, 1952, 175 p.
22. Wigner Eugen P. Symmetries and reflections. Bloomington, Indiana University Press, 1967, 280 p.
23. Spacetime symmetries [Electronic resource]. Access mode: https://en.wikipedia.org/wiki/Spacetime_symmetries.
24. Euclidian group [Electronic resource]. Access mode: https://en.wikipedia.org/wiki/Euclidian_group.
25. Singleton (mathematics) [Electronic resource]. – Access mode: [https://en.wikipedia.org/wiki/Singleton_\(mathematics\)](https://en.wikipedia.org/wiki/Singleton_(mathematics))

Received 07.12.2020.
Accepted 18.04.2021.

ФОРМАЛІЗАЦІЯ МЕТОДІВ КОДУВАННЯ ІНФОРМАЦІЇ В ТОРОЇДНИХ СИСТЕМАХ КООРДИНАТ

Різник В. В. – д-р техн. наук, професор, професор кафедри автоматизованих систем управління Національного університету «Львівська політехніка», Львів, Україна.

АНОТАЦІЯ

Актуальність. Кодування та опрацювання великого інформаційного контенту актуалізує проблему формалізації взаємозалежності між інформаційними параметрами систем кодування векторних даних на єдиній математичній платформі. Об'єктом дослідження є модель кодування масивів векторних даних в базисі тороїдних систем координат. Мета роботи – формалізація методів кодування векторних даних в оптимізованому базисі тороїдних систем координат.

Метод. В основу дослідження покладено встановлений факт гармонійного спряження симетричних та асиметричних структур у вигляді оптимальних співвідношень між симетрією і вбудованими в неї асиметричними просторовими структурами з груповими властивостями, притаманними для векторних комбінаторних конфігурацій «зіркового» класу – наборів t -вимірних векторів двійкового коду як базису t -вимірної системи координат тороїда. Для опису математичної моделі «зіркових» систем координат використано математичний апарат теорії множин, теоретико-числові і теоретико-групові методи комбінаторного аналізу та елементи алгебричної теорії ідеальних кільцевих в'язанок. Метод дозволяє формалізувати взаємозв'язок інформаційних параметрів векторного коду (число кодових комбінацій, розрядність, кількість способів кодування однакових наборів) з геометричними параметрами системи координат (розмірність і розміри сітки по числу атрибутів і категорій векторних даних). Виведено функціональну залежність між вищезгаданими параметрами, що дозволяє досягати вигідного компромісу між суперечними цілями (наприклад, продуктивністю і надійністю методу кодування).

Результати. Розроблені методи реалізовані при оптимальному кодуванні векторних даних в базисі тороїдних систем координат.

Висновки. Проведені обчислювальні експерименти підтвердили працездатність запропонованого математичного забезпечення і дозволяють рекомендувати його для використання на практиці при вирішенні задач формалізації методів оптимального кодування масивів векторних даних та опрацювання інформації різної природи в тороїдних системах координат визначених розмірів і розмірності.

КЛЮЧОВІ СЛОВА: елегантний ансамбль симетрії – асиметрії, закономірність, інформаційна надмірність, комбінаторна конфігурація, оптимальне кодування векторних даних, потужність коду, базис, компроміс.

ФОРМАЛИЗАЦИЯ МЕТОДОВ КОДИРОВАНИЯ ИНФОРМАЦИИ В ТОРОИДНЫХ СИСТЕМАХ КООРДИНАТ

Ризник В. В. – д-р техн. наук, профессор, профессор кафедры автоматизированных систем управления Национального университета «Львовская политехника», Львов, Украина.

АННОТАЦИЯ

Актуальность. Кодирование и обработка большого информационного контента актуализирует проблему формализации взаимосвязности между информационными параметрами систем кодирования векторных данных на единой математической платформе. Объектом исследования является модель кодирования массивов векторных данных в базисе тороидных систем координат. Цель работы – формализация методов кодирования векторных данных в оптимизированном базисе тороидных систем координат.

Метод. В основу исследования положено установленный факт гармонического сопряжения симметричных и асимметричных структур в виде оптимальных соотношений между симметрией и встроенной в нее асимметричными пространственными структурами с групповыми свойствами, присущими для векторных комбинаторных конфигураций «звездного» класса – наборов t -мерных векторов двоичного кода как базиса t -мерной системы координат тороида. Для описания математической модели «звездных» систем координат использован математический аппарат теории множеств, теоретико-числовые и теоретико-групповые методы комбинаторного анализа и элементы алгебраической теории идеальных кольцевых вязанок. Метод позволяет формализовать взаимосвязь информационных параметров векторного кода (число кодовых комбинаций, розрядность, количество способов кодирования одинаковых наборов) с геометрическими параметрами системы координат (размерность и размеры сетки по числу атрибутов и категорий векторных данных). Выведено функциональную зависимость между вышеупомянутыми параметрами, что позволяет достигать выгодного компромисса между противоречивыми целями (например, производительность и надежность метода кодирования).

Результаты. Разработаны методы и реализованы при оптимальном кодировании векторных данных в базисе тороидных систем координат.

Выводы. Проведенные вычислительные эксперименты подтвердили работоспособность предложенного математического обеспечения и позволяют рекомендовать его для использования на практике при решении задач формализации методов оптимального кодирования массивов векторных данных и обработки информации различной природы в тороидных системах координат определенных размеров и размерности.

КЛЮЧЕВЫЕ СЛОВА: элегантный ансамбль симметрия-асимметрия, закономерность, информационная избыточность, комбинаторная конфигурация, оптимальное кодирование векторных данных, мощность кода, базис, компромисс.

ЛІТЕРАТУРА / LITERATURA

1. Xiaochuang Y. Big spatial vector data management: a review / Y. Xiaochuang, Li Guoqing // *Big Earth Data*. – 2018. – Vol. 2, № 1. – P. 108–129. DOI: 10.1080/20964471.2018.1432115
2. Bose N. R. *Applied Multidimensional Systems Theory* / N. R. Bose. – Pennsylvania State University State College, PA, USA: Springer Int. Publishing AG, 2017. – 192 p.
3. Big data for remote sensing: Challenges and opportunities / [M. Chi, A. Plaza, J. A. Benediksson et al.] // *IEEE*. – 104 (11). – P. 2207–2219. DOI:10.1109/Jproc.2016.5982228
4. Remote sensing big data computing: Challenges and opportunities / [Y. Ma, H. Wu, L. Wang, et al.] // *Future Generation Computer Systems*. – 2015. – № 51. – P. 47–60. DOI:10.1016/j.future.2014.10.029
5. Pekturk M. K. A review on real-time big data analysis in remote sensing applications / M. K. Pekturk, M. Unal // *25th Signal Processing and Communications Applications Conference (SIU)*, Antalya, Turkey, 15–18, May 2017
6. A MapReduce-based system for querying and visualizing spatio-temporal satellite data / [A. Eldawy, M. F. Mokbel, S. Alharthi et al.] // *IEEE International Conference on Data Engineering*, Seoul, South Korea, 13–17 April 2015.
7. Developing a reversible rapid coordinate transformation model for the cylindrical projection / [S. Ye, T. Yan, Y. Yue et al.] // *Computers & Geosciences*. – 89. – P. 44–56. DOI:10.1016/j.cageo.2016.01.007
8. Fast multidimensional ensemble empirical mode decomposition for the analysis of big spatio-temporal datasets / [Z. Wu, J. Feng, F. Qiao, Tan M. Z.] // *Philos Trans A Math Phys Eng Sci*. – 2016. – 374(2065), 2015.01.97
9. A parallel spatial data analysis infrastructure for the cloud / [S. Ray, B. Simion, A. D. Brown, R. Johnson] // *ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 5–8 November 2013. – Orlando, FL, USA.
10. Tang W. Parallel map projection of vector-based big spatial data: Coupling cloud computing with graphics processing units / [W. Tang, W. Feng] // *Computers, Environment and Urban Systems*. – 2014. – Vol. 61. – P. 187–197. DOI: 10.1016/j.compenvurbsys.2014.01.001
11. Eldawy A. Hadoop Viz: A MapReduce framework for extensible visualization of big spatial data / [A. Eldawy, M. F. Mokbel, C. Jonathan] // *The 32nd IEEE International Conference on Data Engineering*, 16–20 May 2016. – Helsinki, Finland: IEEE, 2016.
12. *Big Data-Enabled Nursing: Education, Research and Practice* / [W. Connie DeLaney, A. Charlotte Weaver, J. Judith Warren et al.]. – Cham, Switzerland : Springer, 2017. – 488 p. <https://www.bookdepository.com/Big-Data-Enabled-Nursing-Connie-White-Delaney/9783319532998>
13. Topological coordinates for toroidal structures / [I. Laszlo, A. Rassat, P. W. Fowler, A. Graoavas] // *Chemical Physics Letters: Elsevier Science B. V.* – 2001. – Vol. 342. – P. 369–374.
14. Hall Jr. *Combinatorial Theory* / Jr. Hall. – 2 edition : Wiley-Interscience, 1998. – 464 p. DOI: 10.1002 / 9781118032862
15. Rotman Joseph. *Galois Theory* / Joseph Rotman. – Second ed.: Springer, 1998. – 106 p. DOI:10.1007/978-1-4612-0617-0
16. Lee John. *Introduction to topological manifolds* / John M. Lee. – Springer-Verlag, New York, Inc., 2000. – 395 p. DOI:10.1007 / 978-1-4419-7940-7
17. Emily H. Moore *Difference Sets* / Emily H. Moore, Harriet S. Pollatsek // *Student mathematical library. – Connecting Algebra, Combinatorics, and Geometry*. AMS. – 2013. – Vol. 67. – 298 p. ISBN 978-0-8218-9176-6
18. Riznyk W. Application of the Golden Numerical Rings for Configure Acoustic Systems of Fine Resolution / W. Riznyk // *Acta Physica Polonica A*. – 2011. – Vol. 119. – P. 1046–1049. DOI: 10.12693/APhysPolA.119.1046
19. Riznyk V. Multi-modular Optimum Coding Systems Based on Remarkable Geometric Properties of Space / V. Riznyk // *Advances in Intelligent Systems and Computing*. – Springer, 2017. – Vol. 512. – P. 129–148. DOI 10.1007/978-3-319-45991-2_9
20. Riznyk V. V. Multidimensional Systems Optimization Developed from Perfect Torus Groups / V. V. Riznyk // *International Journal of Applied Mathematics and Informatics*. – 2015. – Vol. 9. – P. 50–54. ISSN: 2074-1278.
21. Weyl Hermann. *Symmetry* / Hermann Weyl. – New Jersey: Princeton University Press, Princeton, 1952. – 175 p.
22. Wigner Eugen P. *Symmetries and reflections* / Eugen P. Wigner. – Bloomington : Indiana University Press, 1967. – 280 p.
23. Spacetime_symmetries [Electronic resource]. – Access mode: https://en.wikipedia.org/wiki/Spacetime_symmetries
24. Euclidian group [Electronic resource]. – Access mode: https://en.wikipedia.org/wiki/Euclidian_group
25. Singleton (mathematics) [Electronic resource]. – Access mode: [https://en.wikipedia.org/wiki/Singleton_\(mathematics\)](https://en.wikipedia.org/wiki/Singleton_(mathematics))

INFORMATION TECHNOLOGY OF CLIMATE MONITORING

Talakh M. V. – PhD, Assistant Professor of Department of Computer Science, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine.

Holub S. V. – Dr. Sc., Professor of Department of Automated Systems Software, Cherkasy State Technological University, Cherkasy, Ukraine.

Turkin I. B. – Dr. Sc., Professor, Head of Department of Software Engineering, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine.

ABSTRACT

Context. Information monitoring technology is used to reduce information uncertainty about the regularity of air temperature changes during managing work in hard-to-reach places [1]. The task was to create a method for modelling one of the climatic indicators, air temperature, in the given territories in the information monitoring technology structure. Climate models are the main tools for studying the response of the ecological system to external and internal influences. The problem of reducing information uncertainty in making managerial decisions is eliminated by predicting the consequences of using planned control actions using climate modelling methods in information monitoring technology. The information technology of climate monitoring combines satellite observation methods and observations on climate stations, taking into account the spatial and temporal characteristics, to form an array of input data. It was made with the methods for synthesizing models of monitoring information systems [1] and methods of forming multi-level model structures of the monitoring information systems [1] for converting observation results into knowledge, and with the rules for interpreting obtained results for calculating the temperature value in the uncontrolled territories.

Objective of the work is to solve the problem of identifying the functional dependence of the air temperature in a given uncontrolled territory on the results of observations of the climate characteristics by meteorological stations in the information technology of climate monitoring structure.

Method. The methodology for creating information technologies for monitoring has been improved to expand its capabilities to perform new tasks of forecasting temperature using data from thermal imaging satellites and weather stations by using a new method of climate modelling. A systematic approach to the process of climate modelling and the group method of data handling were used for solving problems of functional dependence identification, methods of mathematical statistics for evaluating models.

Results. The deviation of the calculated temperature values with the synthesized monitoring information systems models from the actual values obtained from the results of observations by artificial earth satellites does not, on average, exceed 2.5°C. Temperature traces obtained from satellite images and weather stations at similar points show similar dynamics.

Conclusions. The problem of the functional dependence identification of air temperature in uncontrolled territories on the results of observations at meteorological stations is solved. The obtained results were used in the process of creating a new method of climate modelling within information technology of climate monitoring. Experimental confirmation of the hypothesis about the possibility of using satellite images in regional models of temperature prediction has been obtained. The effectiveness of the application of the methodology for the creation of monitoring information technologies during the implementation of the tasks of reducing uncertainty for management decisions during works in non-controlled territories has been proven.

KEYWORDS: information technology, monitoring, climate models, observations, air temperature, thermal imagery, Landsat satellites, inductive modelling, machine learning.

ABBREVIATIONS

GMDH is a group method of data handling;
MIS is a monitoring information systems;
RT is a real temperature;
PT is a predicted temperature;
AP is an absolute error;
MinV is a min value;
MaxV is a max value;
DFC is a direction of function's change;
SV is a variable name, showing reliable changing of function's value;
IMT is an information monitoring technology;
ITCM is an information technology of climate monitoring;
MIS is a monitoring information systems;
AID is an array of input data;
OLAP is an online analytical processing.

NOMENCLATURE

Y is a set of the dependent variables;
 X is an independent variable;

m is a number of temperature observation points;
 n is a number of features obtained from the results of observations at meteorological stations;
 L_{λ} is a spectral density of the radiant luminosity at the sensor aperture, that is, the power of the energy flux (Watt) per m^2 of the earth's surface per steradian;
 Q_{cal} is a pixel value;
 $Q_{cal_{min}}$ is a minimum pixel value, which corresponds to $L_{min_{\lambda}}$;
 $Q_{cal_{max}}$ is a maximum pixel value, which corresponds to $L_{max_{\lambda}}$;
 $L_{min_{\lambda}}$ is a spectral luminosity on the sensor, which is scaled to $Q_{cal_{min}} \left[\frac{W}{m^2 sr \mu m} \right]$;

L_{\max} , is a spectral luminosity on the sensor, which is scaled to $Q_{cal_{\max}} \left[\frac{W}{m^2 sr \mu m} \right]$;

M_L is a rescaling rate;

A_L is an additive rescaling rate;

T_C is a brightness temperature on the satellite [C];

K_1 is a channel-dependent thermal conversion constant (Landsat 4 TM – 671.62, Landsat 5 TM – 607.76, Landsat 7 ETM+ – 666.09);

K_2 is a channel-dependent thermal conversion constant [K] (Landsat 4 TM – 1284.30, Landsat 5 TM – 1260.56, Landsat 7 ETM+ – 1282.71);

y_1 is an air temperature measured using a satellite image in the vicinity of the Chernivtsi weather station;

y_2 is an air temperature measured using a satellite image in the vicinity of the Kolomyia weather station;

y_3 is an air temperature measured using a satellite image in the vicinity of the Vizhnitsa weather station;

x_1 is a day of observation;

x_2 is a months of observation;

x_3 is a year of observation;

x_4 is a Y coordinate of the location of the meteorological station “Chernivtsi”;

x_5 is a X coordinate of the location of the meteorological station “Chernivtsi”;

x_6 is a height above sea level of observation points in the vicinity of the meteorological station “Chernivtsi”;

x_7 is an air temperature measured by the meteorological station “Chernivtsi”;

x_8 is a cloudiness index measured by a satellite in the vicinity of the Chernivtsi weather station;

x_9 is a Y coordinate of the “Kolomyia” weather station location;

x_{10} is a X coordinate of the “Kolomyia” weather station

x_{11} is a height above sea level of observation points in the vicinity of the weather station “Kolomyia”;

x_{12} is an air temperature measured by the weather station “Kolomyia”;

x_{13} is a cloudiness index measured by a satellite in the vicinity of the weather station “Kolomyia”;

x_{14} is a Y coordinate of the location of the weather station “Vizhnitsa”;

x_{15} is a X coordinate of the location of the weather station “Vizhnitsa”;

x_{16} is a height above sea level of observation points in the vicinity of the weather station “Vizhnitsa”;

x_{17} is an air temperature measured by the weather station “Vizhnitsa”;

x_{18} is a cloudiness index measured by a satellite in the vicinity of the weather station “Vizhnitsa”.

INTRODUCTION

Information technology monitoring [1] It is built to provide information for decision-making processes in the relevant subject area [2, 3]. In the process of managing work in hard-to-reach places, monitoring is ordered to

© Talakh M. V., Holub S. V., Turkin I. B., 2021
DOI 10.15588/1607-3274-2021-2-16

identify external disturbances in the form of changes in weather indicators, in particular air temperature, and their influence on the consequences of decisions made.

Temperature as one of the fundamental climatic markers of the region is utilized to dissect environmental change and establish climate models of various levels [4]. Presently, normal temperature data of the specific region are received, from meteorological stations. Simultaneously, the grid of meteorological stations isn't adequately extended all through the territory of Ukraine. Additionally, as a source of temperature information, thermal imaging of satellites can be utilized to get temperature information for each surface point (accuracy is resolved simply by the goal of room pictures). However, the frequency of thermal imaging isn't regular, and in general, is much lower than the frequency of obtaining data from meteorological stations. In this way, the utilization of space images permits us to consider the spatial component of monitoring, and data from meteorological stations – temporal component, and their combination will permit to create of a complex system of the temperature monitoring of certain territory [5, 6].

Simulation by observed data is an essential condition for monitoring since it permits to assess rapidly the current circumstance and foresee their turn of events. Specifically, the effectiveness of the utilization of inductive modeling methods for modeling climatic parameters, one of which is temperature, is demonstrated. Thus, the development of the method for determining the value of air temperature, based on establishing functional dependencies between the investigated parameters is relevant.

The object of study is the processes of environmental monitoring.

The subject of study is the processes of forming an array of input data and the synthesis of climate models in the structure of information technology for environmental monitoring.

The purpose of the study is to reduce information uncertainty in a process of management decisions during the work in uncontrolled territories by expanding the capabilities of information technology for environmental monitoring by developing a new method for modelling air temperature in hard-to-reach places.

1 PROBLEM STATEMENT

To reduce information uncertainty in management decision-making, it is proposed to calculate the air temperature at the worksite. A hypothesis was formulated about the existence of a dependence of the air temperature value in the uncontrolled area on the air temperature values measured in the territories, where the air temperature can be measured by meteorological stations. The information content of the indicators and the power of the synthesizer of the monitoring information system will be enough to identify this dependence. The hypothesis is based on knowledge about the emergence of the ecosystem.

It is necessary to solve the problem of identifying the functional dependence of the air temperature in a given area $Y = \{y_1, y_2, \dots, y_m\}$ obtained from the results of processing thermal satellite images with different cloud cover with a given discreteness from the air temperature measured at stationary meteorological stations, form a set $X = \{x_1, x_2, \dots, x_n\}$ values, consistent in time:

$$y_i = f(x_1, x_2, \dots, x_n), i = \overline{1, m}. \quad (1)$$

Problem 1 will be considered solved if the built model has the value of the quality criterion no worse than the specified one. To evaluate the models, the absolute error of modeling the temperature values obtained from the results of processing thermal satellite images on the test sequence of observation points of the AID [1]:

$$AE = RT - PT. \quad (2)$$

2 LITERATURE REVIEW

Presently some studies are demonstrating the effectiveness of using thermal imaging data for estimating and analyzing the air temperature. Specifically, papers [7, 8] show a bunch of ways to deal with the determination of air temperature dependent on data received from remote sensors. Frequently, this determination of temperature is done to take care of practical issues [9]. Regularly, time series analysis is utilized to examine the temporal dynamics of temperature. This chance exists since data from most satellites are accessible on normal from the eighties of the last century [10]. Numerous investigations show the utilization of thermal channels for the determination of thermal anomalies in urban areas, plants, or water bodies. A few works show strategies for identifying and foreseeing the distribution of forest fires by thermal imaging of satellites [11, 12].

There are various ways to deal with how information can be obtained and handled. Specifically, the examination [13] utilized a model to approximate the annual temperature cycle and acquire periodic parameters to describe annual changes in surface temperature, including average and maximum values, both during daytime and at night.

The effectiveness of utilizing the GMDH for forecasting the value of climatic factors has appeared in numerous studies. Remarkably, the paper [14] shows ways to modeling and predicting average wind speeds per hour. A few investigations offer integrated algorithms to build climatic models. In this manner, the article [15] proposed a technique that was obtained by integrating the kernel principal component analysis method with the locally weighted gathering GMDH, which can be gotten by combining GMDH with the spatial regression and weighted least squares regression. The paper [16] presents another case of the successful utilization of a few algorithms. Specifically, for this situation, GMDH was utilized for establishing a climate model, neural network – for predicting wind speed and power in the short term and wavelet noise decrease – for separating high-

frequency emissions in wind speed information, giving a smooth neural network training.

The work [17] proposes the utilization of inductive simulation techniques for the recovery of incomplete or short-term data recorded in the large territory and the making of forecasting climatic models on their basis.

This paper improves the procedure for evaluating the air temperature in hard-to-reach places, utilizing the model of the temperature dependence of the determined by thermal imagery from the temperature got at existing weather stations and creating a monitoring information system.

3 MATERIALS AND METHODS

As is known, monitoring is a technology for providing information to decision-making processes [2, 3]. The necessities for this information technology are determined by the sort of data that the super-system needs for the formation of control actions in managing systems. On account of environmental monitoring, it will be a different level of biological systems, usually – ecosystem. Mainly, the identification of patterns in the temperature regime of the region is significant to foresee conceivable undesirable occasions and to introduce, if necessary, adjustments to economic and other activities. The arrangement of numerous issues related to the investigation of the laws of the temperature system requires the securing, handling, and transformation of large amounts of information – the results of observations. Today, the rules of obtaining and processing various climate characteristics are a remarkably complex process.

The highlights of the climate system present significant limitations on the selection of techniques and tools reasonable for climate prediction. In this manner, the assessment of future environmental change is on a fundamental level possible just as a particular likelihood distribution; along these lines, the undertaking of foreseeing atmosphere is diminished to deciding the probability distribution functions of different attributes of the climate system.

Climate models are attempting to clarify climatic processes with the assistance of fundamental parts of the system – they are simplified versions of reality. The result of these rearrangements is little on a planetary scale; in any case, they are noticeable when foreseeing climate “on the ground”. Simultaneously, the worldwide climate can just anticipate general patterns; however, they are not entirely reasonable for practical use. Due to the “gross” spatial resolution, global climate models poorly reproduce the regional climate. Regional models are one of the approaches to explain information as indicated by local highlights of the orography, atmospheric circulation, and so. The use of local models makes it conceivable to distinguish the impacts related to local characteristics of the underlying surface all the more precisely [5–7].

Temperature was chosen as the principal climatic parameter for further work among climatic indicators. The temperature regime list of indicators of the region and the frequency of observations allows showing the state of the

studied system in the obtained data array. Processing of the results of the observations was done by statistical methods (for pointers of meteorological stations) and classical techniques for geospatial investigation for information acquired by satellite images. Transformation of preparing results into information about the parameters of the temperature system of the area is a mix of the two referenced above strategies. The monitoring results are acquired as two-, three-dimensional conditions. However, the indicators of the temperature system of the territory are controlled by countless variables (principally the dynamics in space and time). The potential outcomes of statistical modeling techniques for showing multifactor impacts are limited [6].

In this case, monitoring information is needed to determine the dependence of temperature values at a certain point in space on a temperature at other points in space, which are characterized by the presence of weather stations. The problem of climate change is currently extremely relevant as in the field of basic research, so far in practice. The development of reliable estimates and objective forecasts of environmental impacts related to climate dynamics determined the researcher's interest in this topic, in particular, the intensity and nature of climate transformations. The task of accurate prediction of temperature values cannot be realized. Therefore, a justified decision of the most significant and informative components accessible for adequate research is required.

The issue explored in this article arises when creating MIS. The proposed arrangement removes the current logical inconsistency between the need to reflect the properties of a complex climate system in the structure of the model and the restricted capacities of the modeling techniques utilized for this. As a result, information about the features of observational objects are not reflected in the model structure and are not used in the decisions making by monitoring results. Consequently, there is a need to develop and implement new techniques for distinguishing and considering these factors, handling and analyzing the received information, which ought to be founded on a combination of exploration results utilizing climatology and meteorology methods joined with current techniques for preparing big data and simulating complex systems. There is a need to reduce the time for receiving conclusions by automating routine work by the widespread use of ITM.

The list of features and their numeric characteristics (contain the results of satellite observations and observations of the air temperature of meteorological stations) form an array of input data (AID) for synthesizing the model of the monitoring object. The input data array is in the form of a two-dimensional table:

Table 1 – The array of observation results

Dependent variable	Independent variables			
	x_1	x_2	...	x_n
y_1	x_{11}	x_{12}	...	x_{1n}
y_2	x_{21}	x_{22}	...	x_{2n}
...
y_m	x_{m1}	x_{m2}	...	x_{mn}

It was proposed to determine the form of the operator f by decoupling the typical problem of identifying functional dependence with an atypical procedure for synthesizing the MIS model [1]. As a more complex tool for solving the problem of air temperature modeling based on the analyzed data, it was proposed to use inductive modeling methods.

According to the methodology for creating information monitoring technologies [1] in the first stage of the formation of the IMT, the super-system formulates requirements for the content of information that must be obtained from the monitoring results.

That is the reason, that the prerequisite for MIS for this situation is to depict the dependence of temperature values (with specific characteristics of the location, altitude, and the nature of the underlying surface), controlled by space photographs, from the estimations of this indicator at other points determined by meteorological stations and cloudiness.

In the second stage of the IMT formation, the number of levels of information conversion, and the local tasks for each of the levels is determined. For this situation, it is important to form two levels of information transformation. At the first level, the indications of the condition of the climate system are changed over into digital characteristics by the results of observations. Means are estimating instruments for determining the temperature (both directly and with the assistance of thermal photos) and regular maintenance of these processes. At the second level, the information is transformed from the form of the array of numerous characteristics into the form of analytical models, the structure of which reflects the change in temperature values at a certain point of space based on satellite images, from its values at other locations where there is a meteorological station.

First level of monitoring implementation results for the analyzed area during 1985–2016. They are used as components of the initial description of monitoring objects and as variables for the synthesis of second-level models.

Y is a set of the dependent variables (modeled indicators). In this case it is air temperature determined on meteorological stations (48 observation points).

Variables 2–3 are independent variables and form the set X .

Here introduced the consequences of studies whose purpose was to decide the idea of the impact of air temperature values determined at meteorological stations and cloudiness index on air temperature, defined by thermal imaging from satellite and meteorological stations. Simultaneously as the examined indicators utilized the estimation of changing it for a particular geographic location.

The study was conducted for three meteorological stations in the Chernivtsi and Ivano-Frankivsk regions (Chernivtsi, Vyzhnytsia, and Kolomyia meteorological stations). Each of the weather stations had 16 observation points. For each weather stations, air temperature was additionally determined from satellite images. As control

points were used the regions which are remote from the meteorological station, for which it is possible to determine only the temperature by thermal imaging. Since sweeping advancement was to establish the possibility of predicting the territory temperature with no meteorological stations, the temperature was chosen as the dependent variable based on the result of the thermal imaging.

At the third stage of building the ITM, the structure of its information transformation subsystem is formed. The method of ascending model synthesis and tools of MIS was used. Under the given dependent and independent variables, models are synthesized at each level of information transformation; models of a separate level are combined into executions. The output signals of the models of the lower penalty form arrays of initial data for the synthesis of the models of the upper penalty [18, 19]. For the integration of models of the second level of information conversion of the MIS temperature based on the satellite of the material stations, a multi-row GMDH algorithm was used [3]. In the process of synthesizing the model, data obtained during 1985–2017 were used for the territory of the weather station. For testing the model, we used the results of observations obtained during the same period for the lands where there are no meteorological stations.

The informativeness of the indicators, which form the array of input data, determines the ITM's ability to perform the monitoring task. The value of independent variables (Table 1) was obtained as observations from meteorological stations in a suitable form for the forming of the AID. The results of satellite observations, which were used to form the dependent variables of modeling, are not suitable for use in the AID without conversion. Their transformation is required.

To download satellite images and save them in a temporary folder, the API service of the American Geological Research Center glovis.usgs.gov was used [20]. API of two weather services were used simultaneously – worldweatheronline.com and rp5.ua. The first service allows to get the coordinates of weather stations, and the second is used to get temperature data. It's possible, because, each pixel besides attribute values, also has geographic coordinates, so analyzing each pixel sequentially, the system sends requests to weather servers for the same period for satellites and weather stations [21, 22]. Data were automatically downloaded in the vicinity of weather stations. These data characterize each of the observation points at different observation times during the study period. The satellites of the Landsat series of series 4, 5, 7, 8 [20, 23], namely, their thermal channels (the specific channel number depends on the series). The choice of satellites is due to the presence of a thermal infrared sensor image (10.4–12.5 μm). Data from the Landsat series satellites have been presented since 1982 and are freely available.

Each scene is read by pixel (an image concerning a particular date, which has a set of specific characteristics that are contained in a file with metadata). A special method has been developed to implement this process.

© Talakh M. V., Holub S. V., Turkin I. B., 2021
DOI 10.15588/1607-3274-2021-2-16

As a result, this method returns an object containing already processed (converted to surface temperature) pixel values.

The conversion of the attributive values of thermal channels pixels of space images into the temperature value takes place in several stages. The first is the definition of radiant luminosity. For Landsat satellites of series 4, 5, and 7, the radiant luminosity on the satellite is calculated by the expression:

$$L_{\lambda} = \left(\frac{L_{\max_{\lambda}} - L_{\min_{\lambda}}}{Q_{cal_{\max}} - Q_{cal_{\min}}} \right) (Q_{cal} - Q_{cal_{\min}}) + L_{\min_{\lambda}} \quad (3)$$

For Landsat 8, the satellite radiant luminosity is calculated using the formula (4):

$$L_{\lambda} = M_L Q_{cal} + A_L \quad (4)$$

At the second stage, the obtained results are recalculated into absolute dimensionless values of reflectivity, which characterizes the ratio of light reflected by the surface and the light entering it. Unlike luminosity, which significantly depends on the lighting conditions, the vividness is its own relatively unchanged characteristic of the object of observation [24]. This exact value was used as the temperature determined from space images.

To calculate the temperature on the satellite for the Landsat series 4, 5, and 7, the following expression was used:

$$T_c = \frac{K_2}{\ln\left(\frac{K_1}{L_{\lambda}} + 1\right)} - 273.15 \quad (5)$$

For the Landsat 8, and K_1 values are given in the image metafile. The values obtained from the metadata are converted from Kelvin to Celsius values.

Temperature data from meteorological stations were received from meteorological servers [21, 22]. Terrestrial meteorological stations are used to evaluate the reliability of temperature measurements from remote sensing data for Landsat satellites.

Since spatial data were obtained from different sources, they were in different cartographic projections. For unification, all data were transformed into a ITM projection, in which all the coordinates given in the the work are presented.

The main disadvantage of this approach is the effect of changes in the shadows and illumination of the earth's surface, vegetation, and objects when analyzing satellite images limits the use of the space-time approach. However, the development of a method for modeling, detecting, and removing shadows on images allows, after additional processing, to carry out the comparison of multi-time images.

To compare the values of temperature measurements, according to the data of meteorological stations and the results of the thermal band from satellite images, we se-

lected images from the Landsat -4, -5, -7, and -8 satellites.

As a result of converting every image to matrix (with attribute values of pixels) we obtain digital data that do not have physical dimensions. It is necessary to convert the space data to the real temperature value. All the required information about the input data is contained in the metadata file supplied with each image (scene).

All input images have resolution 30 m per pixel and are sufficiently accurate to predict climate indicators. When comparing meteorological and satellite data, the concept of the effective forecast radius of a meteorological station (the experimentally established territory for which the proper temperature can be measured at the meteorological station) was used. It is known that on average, such a radius is about 30 km. In this case, 16 points from satellite images closest to the weather station were used to restore weather data.

Also, a significant aspect that was considered during arranging an experiment was the frequency of data acquisition. The fixation of temperature indicators at meteorological stations occurs several times a day (from 4 to 12 times, depending on the type of meteorological station). However, fully satisfying the time component of the weather station research is almost devoid of spatial (in case of working with many stations model would not still be regional). Simultaneously, another source of input data, in particular, satellite images, is completely good in the spatial component of the analysis; however, the Landsat satellite captures 16–18 days, that is, the time during which the satellite turns around the Earth and how often it records. Also, the cloudiness value of more than 50% makes it almost impossible to use a specific scene to an image, which further reduces the frequency of data acquisition. So, one of the tasks of creating IMT within this research was to develop the methods for the common use of all these types of input data.

Primary analysis of the initial data array shows that the graphs of the course of temperatures obtained by both methods are almost identical, which suggests the expediency of using this approach in the future. However, absolute values at specific points are different, which may depend on diversity of factors. Especially, cloudiness, that was considered in the model. Moreover, also from the type of underlying surface that affects the temperature determined by the image, while in the meteorological stations the temperature indicators are measured at a level of 2.6 m above the surface of the earth.

4 EXPERIMENTS

To solve problem 1, the temperature values were determined from satellite images in the vicinity of meteorological stations in Chernivtsi, Kolomyia, and Vyzhnyt-sya. After that, the values of air temperatures in these territories were compared with the values of these indicators measured by meteorological stations, and an array of input data was formed in the form of Table 1. The models were built in the form of the dependence of the air temperature determined from satellite images in the area

of one of the three meteorological stations on the air temperature measured by the other two weather stations.

The list of indicators that are used as elements of the set of independent X and dependent variables Y are presented in nomenclature.

Since the air temperature in the uncontrolled territories had to be calculated without data received from the satellite, the cloudiness indicators, denoted by the variables x_8, x_{13}, x_{18} , were removed from the array of input data. Therefore, the problem of interpreting the functional dependencies of individual elements of the set Y on certain elements of the set X was solved according to the expressions:

$$y_1 = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}, x_{14}, x_{15}, x_{16}, x_{17}); \quad (6)$$

$$y_2 = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{14}, x_{15}, x_{16}, x_{17}); \quad (7)$$

$$y_3 = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}, x_{14}, x_{15}, x_{16}). \quad (8)$$

Thus, expression (6) reflects the dependence of the value y_1 of the air temperature measured by the satellite in the area of the Chernivtsi meteorological station on the characteristics of observation points and on the results of measurements by the Kolomyia and Vizhnyt-sa meteorological stations x_9-x_{17} . Expression (7) reflects the dependence of the value y_2 of the air temperature measured by the satellite in the area of the Kolomyia meteorological station on the characteristics of observation points and the measurement results by the Chernivtsi and Vizhnyt-sa meteorological stations $x_4-x_7, x_{14}-x_{17}$. Expression (8) reflects the dependence of the value y_2 of the air temperature measured by the satellite in the area of the Vizhnyt-sa meteorological station on the characteristics of observation points and the results of measurements by the Chernivtsi and Kolomyia meteorological stations x_4-x_7, x_9-x_{12} .

To solve problems (6), (7), (8), the multi-row GMDH algorithm was used [18]. The array of input data was formed based on the results of simultaneous observations of 16 points around each of the 3 meteorological stations and the transformation of satellite images obtained during 1985 – 2016. Data for 14 indicators of the set X were obtained from the results of 928 measurements of air temperature at 3 meteorological stations, to obtain indicators of the set Y were processed 58 satellite images. That is, according to Table 1, the values $m = 928, n = 14$.

To build the model, observation points were used, built using 57 satellite images. 16 observation points, formed from the results of measurements of 16 observation points around the meteorological stations using the last 58th satellite image, used as a test data set. They were not used in the process of building the model.

5 RESULTS

Figure 1 shows the values of the absolute errors of modeling the air temperature on the test data set.

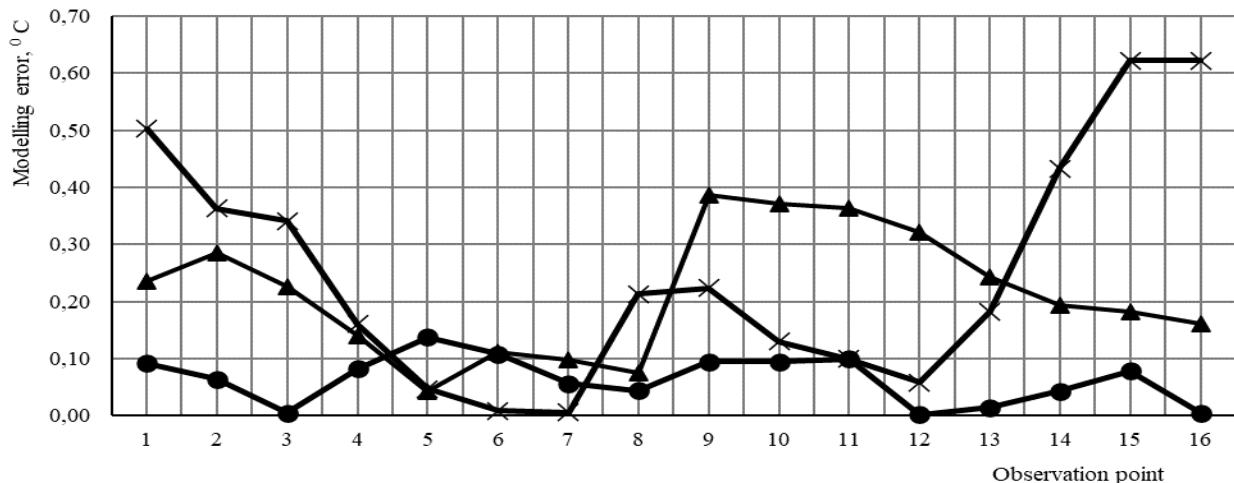


Figure 1 – Modeling errors of air temperature on uncontrolled areas:
 ● Chernivtsi, ▲ – Kolomyia, × – Vizhnitsa

6 DISCUSSION

It should be noted that the models were tested on data, that did not take part in their creation, so obtained the results reflect the persistence of the model. Such a choice of parameters for modeling and generating an array of input data is determined by the need to avoid multicollinearity to the consistency of two or more variables in the models.

Modeling error, calculated by expression (2), varied in the range from 0.02% to 7.82%, which did not exceed 0.7°C. at separate observation points. And in this case:

- modelling error, reflecting dependence (6), varies in the range (0.02–1.25)%, which does not exceed 0.2°C, the standard deviation is 0,05°C;

- modelling error, reflecting dependence (7), varies in the range (0.54–5.13)%, which does not exceed 0.4°C, the standard deviation is 0.17°C;

- modelling error, reflecting dependence (8), varies in the range (0.06–7.82)%, which does not exceed 0.7°C, the standard deviation of the simulation results is 0.21°C.

The general characteristics of the models, considered (Fig. 1), confirms that the characteristics of the simulation results depend on the distance between the plotted points and the comparison of their absolute heights. Comparing the models, it was found that the difference in indicators (longitude, latitude, and altitude) between the modeled point and the variable, which showed a change in the function direction at about the same level. This suggests that for each point for which the temperature simulation is carried out, the point(s) can be set that has the most significant effect on the obtained value.

The study of existing climatic models of various levels allows us to assert that the temperature depends more on its dynamics in space and time. So, during the processing of space images, the date, as well as the specific shooting season, is essential. This is taken into account when carrying out an atmospheric correction and is one of the necessary procedures for preparing the analyzed data [9].

The use of thermal satellite images to predict temperature is also important because they contain information about cloudiness, as well as the location of the studied points and even heights, which is also one of the decisive factors influ-

encing the obtained value. Also, the type of underlying surface is taken into account on satellite images and has a noticeable effect on the temperature regime [9].

An analysis of the results obtained shows that the synthesized MIS models are accurate and adequate since the difference between the air temperature values calculated by the models and their actual values measured at meteorological stations, and the value of the standard deviation of the simulation results from the actual values are acceptable. Even though the accuracy and adequacy of the simulation results were evaluated on test data sets, the obtained models are stable.

Thus, the hypothesis about the possibility of calculating the air temperature value in the uncontrolled territory by monitoring information systems using the results of measurements of operating meteorological stations was obtained experimentally. The possibility of identifying functional dependencies of air temperature in remote areas on the available indicators has been proved.

Obtained results make it possible to substantiate the possibility of obtaining climatic characteristics of remote areas by identifying their dependences on the results of measurements at active meteorological stations, which are difficult to access or uncontrollable.

CONCLUSIONS

The problem of identification of the functional dependence of air temperature in uncontrolled places on the results of observations at meteorological stations is solved. The results obtained were used in the process of creating a new method of climate modeling within ITCM. Experimental confirmation of the hypothesis about the possibility of using satellite images in regional models of temperature prediction has been obtained. The effectiveness of the application of the methodology for the creation of information technologies for monitoring in the implementation of the tasks of reducing uncertainty in making decisions on the management of works in uncontrolled territories has been proved.

The analysis and synthesis of existing approaches, methods, and means of temperature prediction, including by satellite imagery, in particular, by thermal imaging, has been carried out. It was revealed that the most effective means of

solving the problem is the use of methods of inductive modeling.

The analysis and synthesis of existing approaches, methods, and means of temperature prediction, including by satellite imagery, in particular, by thermal imaging, has been carried out. It was revealed that the most effective means of solving the problem is the use methods of inductive modeling.

During the investigation of the variables' selection of the input data array, it was shown that satellite imagery data could be used in regional models of temperature prediction, and temperature traces obtained from satellite imagery and weather stations at similar points show similar dynamics.

The effectiveness of GMDH using of the multi-row algorithm for forecasting temperature for areas with no meteorological stations is shown.

The scientific novelty of obtained results is the improvement of the methodology for creating information technologies for multilevel monitoring by using the existing method of adaptive synthesis of models to solve the new problem of creating a new method of climate modeling. In contrast to the existing ones, the method combines the processes of satellite observations and observations of the climate on the soil surface, taking into account the spatial and temporal characteristics, to form an array of input data (AID), with the processes of synthesis of models of monitoring information systems (MIS) [1] and the processes of forming multi-level model structures MIS for converting observation results, and with the rules for interpreting the results obtained for calculating the temperature value in uncontrolled territories. This makes it possible to expand the capabilities of monitoring information systems to perform new tasks of calculating air temperature values in uncontrolled areas during the time between space observations.

The practical significance of the obtained results is in the possibility of constructing an information monitoring system that ensures a continuous determination of air temperature in areas where there are no meteorological stations. The results can be widely used to reduce the uncertainty of decision-making processes while eliminating both global and local problems. In particular, for planning work in hard-to-reach and uncontrolled territories, observing climate changes and creating climate models, or identifying the state of vegetation and the presence of traces of human activity in the studied territories.

Prospects for further research are to improve developed monitoring information system due to the use of a complex mechanism is needed that combines OLAP technology and IoT search mechanisms, which will affect the final result, for example, taking into account the features of the underlying surface or the presence of "heat islands" over cities.

REFERENCES

1. Golub S. V. Multilevel modelling in environmental monitoring technologies: monograph. Cherkasy, Ed. from. ChNU named after Bohdan Khmelnytsky, 2007, 220 p. ISBN 978-966-353-062-8.
2. Kunytska S., Holub S. Eds: Palagin A., Anisimov A., Morozov A., Shkarlet S. Multi-agent Monitoring Information Systems, *Mathematical Modeling and Simulation of Systems. MODS. Advances in Intelligent Systems and Computing*. Cham, Springer, 2020, Vol. 1019, pp. 164–171. DOI: 10.1007/978-3-030-25741-5_17
3. Holub S., Burliai I. The accuracy improving modelling of fire-fighting process in the Information system of fire safety monitoring, *Journal of the Technical University of Gabrovo*, 2014, Vol. 47, pp. 13–16. <http://izvestia.tugab.bg/index.php?m=20&tom=5>.
4. Bathiany S., Dakos V., Scheffer M., Lenton T. M. Climate models predict increasing temperature variability in poor countries, *Science Advances*, 2018, Vol. 4, Issue 5, pp. 1–10. DOI: 10.1126/sciadv.aar5809
5. Rosselló J. C., Poyatos R., Ninyerola M., Llorens P. Combining remote sensing and GIS climate modelling to estimate daily forest evapotranspiration in a Mediterranean mountain area, *Hydrology and Earth System Sciences*, 2011, Vol. 15, Issue. 5, pp. 1563–1575. DOI:10.5194/hess-15-1563-2011.
6. Yang J. Fu R., Gong P., Zhang M. The role of satellite remote sensing in climate change studies, *Nature Climate Change*, 2013, Vol. 3, Issue 1, pp. 875–883. DOI:10.1038/nclimate1908.
7. Niclos R., Valiente J. A., Barbera M. J., Caselles V. Land Surface Air Temperature Retrieval From EOS-MODIS Images, *IEEE Geoscience and Remote Sensing Letters*, 2014, Vol. 11, No. 8, pp. 1380–1384. DOI: 10.1109/LGRS.2013.2293540.
8. Converting Landsat TM and ETM+ thermal bands to temperature, The Yale Center for Earth Observation, 2010, [Online]. Available: http://geography.middlebury.edu/data/gg1002/Handouts/Landsat_D_N_Temp.pdf.
9. Meaden G. J., Aguilar-Manjarrez J. Advances in geographic information systems and remote sensing for fisheries and aquaculture, *FAO Fisheries and Aquaculture Technical Paper*, 2013, No. 552, 425 p. ISBN 978-92-5-107391-9
10. Merchant C. J., Embury O., Bulgin C. E., et al. Satellite-based time-series of sea-surface temperature since 1981 for climate applications, *Sci Data*, 2019, Vol. 6, Issue 223, pp. 1–18. [Online]. Available: <https://www.nature.com/articles/s41597-019-0236-x>. DOI: 10.1038/s41597-019-0236-x
11. Heyerdahl E. K., McKenzie D., Daniels L. D., Hessl A. E., Littell J. S., Mantua N. J. Climate drivers of regionally synchronous fires in the inland Northwest, *International Journal of Wildland Fire*, 2008, Vol. 17, pp. 40–49. DOI:10.1071/WF07024
12. Sundara Kumar K., Udayabhaskar P., Padmakumari K. Estimation of Land surface temperature to study Urban Heat Island Effect Using Landsat ETM+ image, *International journal of engineering, science and technology*, 2018, Vol. 4, Issue 2, pp. 807–814. ISSN : 0975-5462
13. Zhao Wei An, Juelin He, Yanhong Wu, Donghong Xiong, Fengping Wen, Ainong Li An Analysis of Land Surface Temperature Trends in the Central Himalayan Region Based on MODIS Products, *Remote Sensing*, 2019, Vol. 11(8): 900, pp. 1–24. [Online]. Available: <https://www.mdpi.com/2072-4292/11/8/900>. DOI:10.3390/rs11080900.
14. Abdel-Aal R. E., Elhadidy M.A., Shaahid S. M. Modeling and forecasting the mean hourly wind speed time series using GMDH-based abductive networks, *Renewable Energy*, 2009, Vol. 34, Issue 7, pp. 1686–1699. DOI: 10.1016/j.renene.2009.01.001.
15. Elattar E., Taha I. B M An Advanced Intelligent Method for Wind Power Prediction, *International Journal of Scientific and Engineering Research*, 2013, Vol. 5, pp. 1–10. Available: https://www.researchgate.net/profile/Ibrahim-Taha-10/publication/292995732_An_Advanced_Intelligent_Method_for_Wind_Power_Prediction/links/56b4fc6708ae44bb3305798b/An-Advanced-Intelligent-Method-for-Wind-Power-Prediction.pdf
16. Makhlofi S. Wind Speed and Wind Power Forecasting Using Wavelet Denoising-GMDH Neural Network, *The 5th International Conference on Electrical Engineering, November 6–7. Algiers, Algeria*, 2018, pp. 29–31. DOI: 10.1109/GTDAAsia.2019.8715985.
17. Saghafiyan B., Ghasemi S., Nasserli M. Backcasting long-term climate data: evaluation of hypothesis, *Theoretical and Applied Climatology*, 2017, Volume 132, Issue 3–4, pp. 717–726, DOI:10.1029/2010WR009104.
18. Madala H. R., Ivakhnenko A. G. Inductive Learning Algorithms for Complex System Modeling. Boca Raton; Ann Arbor; London; Tokyo: CRC Press, 1994, 368 p. ISBN 9781315894393.
19. Müller, J.-A., Lemke, F. Self-Organising Data Mining. Hamburg, Libri, 2000, 250 p. U.S. Geological Survey [Electronic resource]. Access mode: <https://www.usgs.gov/>. DOI: 10.1080/0232929031000136135.

20. USGS [Electronic resource]. Access mode: <https://www.usgs.gov/>.
21. RP5 [Electronic resource]. Access mode: <http://rp5.ua>.
22. Worldweather [Electronic resource]. Access mode: <https://www.worldweatheronline.com>.
23. NASA. Landsat Science [Electronic resource]. Access mode: <http://landsat.gsfc.nasa.gov>.
24. Zabelin S. A., Tulegulov A. D. Metodika atmosfernoj korrektsii snimkov Landsat [Tekst], *Bulletin ENU named by. L. N. Gumileva*, 2011, No. 6, pp. 147–154. ISSN 2616-6771.

Received 23.12.2020.

Accepted 25.04.2021.

УДК 504.064.37: 004.942

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ КЛІМАТИЧНОГО МОНІТОРИНГУ

Талах М. В. – канд. біол. наук, асистент кафедри комп'ютерних наук Чернівецького національного університету ім. Юрія Федьковича, Чернівці, Україна.

Голуб С. В. – д-р техн. наук, професор кафедри програмного забезпечення автоматизованої системи Черкаського державного технологічного університету, Черкаський державний технологічний університет, Черкаси, Україна.

Туркін І. Б. – д-р техн. наук, професор, завідувач кафедри інженерії програмного забезпечення Національного аерокосмічного університету «Харківський авіаційний інститут», Харків, Україна.

АНОТАЦІЯ

Актуальність. Для зниження інформаційної невизначеності про закономірності зміни температури повітря при управлінні роботами у важкодоступних місцях використовується інформаційна технологія моніторингу [1]. Виконувалось завдання створення методу моделювання одного із кліматичних показників, температури повітря, на заданих територіях в структурі інформаційної технології моніторингу. Кліматичні моделі є основними інструментами для дослідження реакції екологічної системи на зовнішні та внутрішні впливи. Проблема зниження інформаційної невизначеності при прийнятті управлінських рішень усувається шляхом прогнозування наслідків застосування запланованих керуючих впливів за допомогою використання в інформаційній технології моніторингу методів кліматичного моделювання. В інформаційну технологію кліматичного моніторингу поєднані методи супутникових спостережень і спостережень за кліматом на поверхні ґрунту з урахуванням просторово-часових характеристик, для формування масиву вхідних даних, із методами синтезу моделей моніторингових інформаційних систем [1] і методами формування багаторівневих модельних структур моніторингових інформаційних систем [1] для перетворення результатів спостережень у знання, та із правилами інтерпретації отриманих результатів для розрахунку значення температури на невідконтрольованих територіях.

Мета роботи полягає в розв'язанні задачі ідентифікації функціональної залежності температури повітря на заданій невідконтрольованій території від результатів спостережень за характеристиками клімату метеорологічними станціями на поверхні ґрунту в структурі інформаційної технології кліматичного моніторингу.

Метод. Удосконалено методологію створення інформаційних технологій моніторингу з метою розширення її можливостей для виконання нових завдань прогнозування температури за допомогою даних тепловізійних супутників та метеостанцій шляхом використання нового методу кліматичного моделювання. Застосовувався системний підхід до процесу кліматичного моделювання, метод групового урахування аргументів для розв'язання задачі ідентифікації функціональної залежності, методи математичної статистики для оцінки моделей.

Результати. Відхилення розрахованих значень температури за синтезованими моделями МІС від дійсних значень, отриманих за результатами спостережень штучними супутниками Землі у середньому не перевищує 2,5°C. Сліди температури, отримані від супутникових зображень та метеорологічних станцій у подібних точках, мають схожу динаміку.

Висновки. Розв'язана задача ідентифікації функціональної залежності температури повітря у невідконтрольованих місцях від результатів спостережень на метеорологічних станціях. Отримані результати використані в процесі створення нового методу кліматичного моделювання в межах інформаційної технології кліматичного моніторингу. Отримано експериментальне підтвердження гіпотези про можливість використання супутникових зображень в регіональних моделях прогнозування температури. Доведена ефективність застосування методології створення інформаційних технологій моніторингу при виконанні завдань зниження невизначеності при прийнятті рішень із управління роботами на невідконтрольованих територіях.

КЛЮЧОВІ СЛОВА: інформаційна технологія, моніторинг, кліматичні моделі, спостереження, температура повітря, тепловізори, супутники Landsat, індуктивне моделювання, машинне навчання.

УДК 504.064.37: 004.942

ІНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ КЛИМАТИЧЕСКОГО МОНИТОРИНГА

Талах М. В. – канд. біол. наук, асистент кафедри комп'ютерних наук Черновицького національного університету ім. Юрія Федьковича, Черновці, Україна.

Голуб С. В. – д-р техн. наук, професор кафедри програмного забезпечення автоматизованих систем Черкаського державного технологічного університету, Черкаси, Україна.

Туркін І. Б. – д-р техн. наук, професор, завідувач кафедри інженерії програмного забезпечення Національного аерокосмічного університету «Харківський авіаційний інститут», Харків, Україна.

АННОТАЦИЯ

Актуальность. Для снижения информационной неопределенности о закономерности изменения температуры воздуха при управлении работами в труднодоступных местах используется информационная технология мониторинга [1]. Выполнялось задание создания метода моделирования одного из климатических показателей, температуры воздуха, на заданных территориях в структуре информационной технологии мониторинга. Климатические модели являются основными инструментами для исследования реакции экологической системы на внешние и внутренние воздействия. Проблема снижения информационной неопределенности при принятии управленческих решений устраняется путем прогнозирования последствий применения запланированных управляющих воздействий с помощью использования в информационной технологии мониторинга методов климатического моделирования. В информационную технологию климатического мониторинга объединены методы спутниковых наблюдений и наблюдений за климатом на поверхности почвы с учетом пространственно-временных характеристик, для формирования массива входных данных массива входных данных, с методами синтеза моделей мониторинговых информационных систем [1] и методами формирования многоуровневых модель-

ных структур мониторинговых информационных систем [1] для преобразования результатов наблюдений в знания, и с правилами интерпретации полученных результатов для расчета значения температуры на неподконтрольных территориях.

Цель работы заключается в решении задачи идентификации функциональной зависимости температуры воздуха на заданном неподконтрольной территории от результатов наблюдений за характеристиками климата метеорологическими станциями на поверхности почвы в структуре информационная технологии мониторинга.

Метод. Усовершенствована методология создания информационных технологий мониторинга с целью расширения ее возможностей для выполнения новых задач прогнозирования температуры с помощью данных тепловизионных спутников и метеостанций путем использования нового метода климатического моделирования. Применялся системный подход к процессу климатического моделирования, метод группового учета аргументов для решения задач идентификации функциональной зависимости, методы математической статистики для оценки моделей.

Результаты. Отклонение рассчитанных значений температуры с синтезированными моделями МИС от действительных значений, полученных по результатам наблюдений искусственными спутниками Земли в среднем, не превышает 2,5°C. Следы температуры, полученные от спутниковых изображений и метеорологических станций в подобных точках, имеют похожую динамику.

Выводы. Решена задача идентификации функциональной зависимости температуры воздуха в неподконтрольных местах от результатов наблюдений на метеорологических станциях. Полученные результаты использованы в процессе создания нового метода климатического моделирования в пределах информационной технологии мониторинга. Получено экспериментальное подтверждение гипотезы о возможности использования спутниковых изображений в региональных моделях прогнозирования температуры. Доказана эффективность применения методологии создания информационных технологий мониторинга при выполнении задач сنجия неопределенности при принятии решений по управлению работами на неподконтрольных территориях.

КЛЮЧЕВЫЕ СЛОВА: информационная технология, мониторинг, климатические модели, наблюдения, температура воздуха, тепловизоры, спутники Landsat, индуктивное моделирование, машинное обучение.

ЛІТЕРАТУРА / LITERATURE

1. Голуб С. В. Багаторівневе моделювання в технологіях моніторингу оточуючого середовища. Черкаси : Вид. від. ЧНУ імені Богдана Хмельницького, 2007. – 220 с. ISBN 978-966-353-062-8
2. Kunytska S. Multi-agent Monitoring Information Systems / S. Kunytska, S. Holub // Mathematical Modeling and Simulation of Systems. MODS. Advances in Intelligent Systems and Computing / Eds: A. Palagin, A. Anisimov, A. Morozov, Shkarlet S. – Cham : Springer, 2020. – Vol 1019. – P. 164–171. DOI: 10.1007/978-3-030-25741-5_17
3. Holub S. The accuracy improving modelling of firefighting process in the Information system of fire safety monitoring / S. Holub, I. Burliai // Journal of the Technical University of Gabrovo. – 2014. – Vol. 47. – P. 13–16. https://www.academia.edu/10082252/The_accuracy_improving_of_modeling_of_firefighting_process_in_the_information_system_of_fire_safety_monitoring
4. Climate models predict increasing temperature variability in poor countries / [S. Bathiany, V. Dakos, M. Scheffer, T. M. Lenton] // Science Advances. – 2018. – Vol. 4, Issue 5. – P. 1–10.
5. Combining remote sensing and GIS climate modelling to estimate daily forest evapotranspiration in a Mediterranean mountain area / [J. C. Rosselló, R. Poyatos, M. Ninyerola, P. Llorens] // Hydrology and Earth System Sciences. – 2011. – Vol. 15, Issue 5. – P. 1563–1575. DOI: 10.5194/hess-15-1563-2011.
6. The role of satellite remote sensing in climate change studies / [J. Yang, R. Fu, P. Gong, M. Zhang] // Nature Climate Change. – 2013. – Vol. 3, Issue 1. – P. 875–883. DOI:10.1038/nclimate1908.
7. Land Surface Air Temperature Retrieval From EOS-MODIS Images / [R. Niclos, J. A. Valiente, M. J. Barbera, V. Caselles] // IEEE Geoscience and Remote Sensing Letters. – 2014. – Vol. 11, No. 8. – P. 1380–1384. DOI: 10.1109/LGRS.2013.2293540.
8. Converting Landsat TM and ETM+ thermal bands to temperature / The Yale Center for Earth Observation. – 2010, [Online]. Available: http://geography.middlebury.edu/data/gg1002/Handouts/Landsat_DN_Temp.pdf.
9. Meaden G. J. Advances in geographic information systems and remote sensing for fisheries and aquaculture / G. J. Meaden, J. Aguilar-Manjarrez // FAO Fisheries and Aquaculture Technical Paper. – 2013. – No. 552. – 425 p. ISBN 978-92-5-107391-9
10. Merchant C. J. Satellite-based time-series of sea-surface temperature since 1981 for climate applications / [C. J. Merchant, O. Embury, C. E. Bulgim et al.] // Sci Data. – 2019. – Vol. 6, Issue 223. – P. 1–18. [Online]. Available: <https://www.nature.com/articles/s41597-019-0236-x>. DOI: 10.1038/s41597-019-0236-x.
11. Climate drivers of regionally synchronous fires in the inland Northwest / [E. K. Heyerdahl, D. McKenzie, L. D. Daniels et al.] // International Journal of Wildland Fire. – 2008. – Vol. 17. – P. 40–49. DOI:10.1071/WF07024.
12. Sundara Kumar K. Estimation of Land surface temperature to study Urban Heat Island effect Using Landsat ETM+ image / K. Sundara Kumar, P. Udayabhaskar, K. Padmakumari // International journal of engineering, science and technology. – 2018. – Vol. 4, Issue 2. – P. 807–814. ISSN : 0975-5462.
13. Zhao Wei An Analysis of Land Surface Temperature Trends in the Central Himalayan Region Based on MODIS Products / Wei Zhao, Juelin He, Yanhong Wu et al] // Remote Sensing. – 2019. – Vol. 11(8): 900. – P. 1–24. [Online]. Available: <https://www.mdpi.com/2072-4292/11/8/900>. DOI:10.3390/rs11080900
14. Abdel-Aal R. E. Modeling and forecasting the mean hourly wind speed time series using GMDH-based abductive networks / R. E. Abdel-Aal, M. A. Elhadidy, S. M. Shaahid // Renewable Energy. – 2009. – Vol. 34, Issue 7. – P. 1686–1699. DOI: 10.1016/j.renene.2009.01.001.
15. Elattar E. An Advanced Intelligent Method for Wind Power Prediction / E. Elattar, I. B. M. Taha // International Journal of Scientific and Engineering Research. – 2013. – Vol. 5. – P. 1–10. Available: https://www.researchgate.net/profile/Ibrahim-Taha-10/publication/292995732_An_Advanced_Intelligent_Method_for_Wind_Power_Prediction/links/56b4fc6708ae44bb3305798b/An-Advanced-Intelligent-Method-for-Wind-Power-Prediction.pdf.
16. Makhloufi S. Wind Speed and Wind Power Forecasting Using Wavelet Denoising-GMDH Neural Network / S. Makhloufi // The 5th International Conference on Electrical Engineering, November 6–7. – Algiers, Algeria : 2018. – P. 29–31. DOI: 10.1109/GTDAAsia.2019.8715985.
17. Saghafian B. Backcasting long-term climate data: evaluation of hypothesis / B. Saghafian, S. Ghasemi, M. Nasseri // Theoretical and Applied Climatology. – 2017. – Volume 132, Issue 3–4. – P. 717–726, 2017. Doi:10.1029/2010WR009104.
18. Madala H. R. Inductive Learning Algorithms for Complex System Modeling / H. R. Madala, A. G. Ivakhnenko. – Boca Raton; Ann Arbor; London; Tokyo: CRC Press, 1994. – 368 p. ISBN 9781315894393.
19. Müller J.-A. Self-Organising Data Mining / J.-A. Müller, F. Lemke. – Hamburg : Libri, 2000. – 250 p. DOI:10.1080/0232929031000136135.
20. Геологічний сервіс. USGS [Електронний ресурс]. – Режим доступу: <https://www.usgs.gov/>.
21. RP5 [Електронний ресурс]. – Режим доступу: <http://rp5.ua>.
22. Worldweather [Електронний ресурс]. – Режим доступу: <https://www.worldweatheronline.com>.
23. NASA. Landsat Science [Електронний ресурс]. – Режим доступу: <http://landsat.gsfc.nasa.gov>
24. Забелин С. А. Методика атмосферной коррекции снимков Landsat [Текст] / С. А. Забелин, А. Д. Тулегулов // Вестник ЕНУ им. Л. Н. Гумилева. – 2011. – № 6. – С. 147–154. ISSN 2616-6771.

РЕАЛИЗАЦИЯ КОНЕЧНО-ЭЛЕМЕНТНОЙ БИБЛИОТЕКИ КЛАССОВ С ИСПОЛЬЗОВАНИЕМ ОБОБЩЕННОГО ПРОГРАММИРОВАНИЯ

Чопоров С. В. – д-р техн. наук, профессор, профессор кафедры программной инженерии Запорожского национального университета, Запорожье, Украина.

Игнатченко М. С. – аспирант кафедры программной инженерии Запорожского национального университета, Запорожье, Украина.

Кудин А. В. – канд. физ.-мат. наук, доцент кафедры программной инженерии Запорожского национального университета, Запорожье, Украина.

Кривохата А. Г. – канд. физ.-мат. наук, старший преподаватель кафедры программной инженерии Запорожского национального университета, Запорожье, Украина.

Гоменюк С. И. – д-р техн. наук, профессор, декан математического факультета, профессор кафедры программной инженерии Запорожского национального университета, Запорожье, Украина.

АННОТАЦИЯ

Актуальность. Для компьютерного моделирования сложных объектов и явлений различной природы на практике часто используют численный метод конечных элементов. Его программная реализация (особенно для исследования новых классов задач) является достаточно трудоемким процессом.

Высокая стоимость разработки программного обеспечения обуславливает актуальность разработки новых подходов к повышению эффективности программирования и сопровождения (в т.ч. добавление новых функций).

Цель. Цель работы – создание новой эффективной архитектуры программ конечно-элементного анализа проблем математической физики, позволяющей легко расширять их функциональность для решения новых классов задач.

Метод. Предложен метод разработки программ для конечно-элементного анализа с использованием обобщенного программирования, что дает возможность существенно упростить архитектуру программного обеспечения и сделать его более удобным для сопровождения и модификации за счет разделения алгоритмов и структур данных.

Предложена новая архитектура классов, реализующих конечно-элементный расчет, позволяющая легко расширять функциональность программ за счет добавления новых типов конечных элементов, методов решения систем линейных алгебраических уравнения, параллельных вычислений и т.д.

Результаты. Предложенный подход был программно реализован в виде библиотеки классов на языке C++. Проведен ряд вычислительных экспериментов, подтвердивших его работоспособность при решении практических задач.

Выводы. Разработанный подход можно использовать как для создания систем конечно-элементного анализа общего назначения с открытой архитектурой, так и для реализации специализированных программных пакетов, ориентированных на решение конкретных классов задач (механики разрушения, эластомеров, контактного взаимодействия и т.п.).

КЛЮЧЕВЫЕ СЛОВА: метод конечных элементов, конечный элемент, объектно-ориентированное программирование, обобщенное программирование, архитектура программного обеспечения.

АББРЕВИАТУРЫ

ПО – программное обеспечение;
МКЭ – метод конечных элементов;
КЭ – конечный элемент;
СЛАУ – система линейных алгебраических уравнений;

НОМЕНКЛАТУРА

u_i – компоненты вектора перемещений;
 u_i^l – искомое значение функции $u_i(\mathbf{x})$ в l -ой вершине (узле) конечного элемента;
 ε_{ij} – компоненты тензора деформации;
 σ_{ij} – компоненты тензора напряжения;
 X_i – компоненты вектора объемных нагрузок;
 \bar{X}_i – компоненты вектора поверхностных нагрузок;
 Ω – область расчета;
 Γ – поверхность области расчета;
 $N_i(\mathbf{x})$ – функция формы КЭ;
 M – количество узлов КЭ;

D – матрица упругости, содержащая характеристики материала;
 N – вектор компонент функций формы;
 V – вектор производных от функций формы;
 p – вектор компонент объемной нагрузки;
 q – вектор компонент поверхностной нагрузки;
 K – матрица жесткости;
 F – вектор компонент нагрузки.

ВВЕДЕНИЕ

Использование вычислительной техники для замены физических испытаний виртуальным компьютерным экспериментом позволяет существенно уменьшить издержки при проектировании и производстве сложных современных инженерно-технических систем.

Применение МКЭ является одним из наиболее эффективных при реализации компьютерного анализа. Таким образом, возникает необходимость в постоянном совершенствовании существующего и разработке нового ПО для компьютерного моделирования различных объектов и процессов.

Объектом исследования являлся процесс разработки ПО для математического моделирования и анализа сложных инженерно-технических систем.

Наиболее важным этапом проектирования является оценка прочности и долговечности создаваемой техники. Это приводит к необходимости анализа ее напряженно-деформированного состояния в процессе эксплуатации. В большинстве случаев аналитическое решение соответствующих задач не представляется возможным, поэтому на практике применяют различные численные методы, наиболее популярным среди которых в силу своей универсальности является МКЭ [1].

Для применения МКЭ в настоящее время разработано большое количество различных коммерческих программных систем, самыми известными среди которых являются Ansys [2], COMSOL [3], MSC Nastran [4] и др. [5].

Альтернативой им являются ПО с открытым исходным кодом, среди которого можно выделить FreeFEM [6], GetFEM [7], OpenCAD [8] и др. [5, 9, 10].

Сложность стоящих перед промышленностью задач требует постоянной разработки новых методов и алгоритмов решения различных классов задач (механики разрушения, гидро- и аэродинамики, контактных задач и т. д.), а особенно новых композитных материалов. Все это обуславливает необходимость создания такого ПО для конечно-элементного анализа, архитектура которого бы позволяла легко встраивать новые типы КЭ для учета особенностей того или иного конструкционного материала (например, эластомера); новые методы решения СЛАУ; поддержку параллельных вычислений и т. п.

Предметом исследования являлись методы разработки программ конечно-элементного анализа для создания удобной к модификации архитектуры соответствующего ПО.

Одним из возможных путей решения этой задачи является использование обобщенной парадигмы программирования, позволяющей отделить наиболее общие алгоритмы от соответствующих структур данных.

Целью работы являлось создание такой архитектуры ПО для конечно-элементного анализа, которая бы позволила сократить время на его разработку и сопровождение.

1 ПОСТАНОВКА ЗАДАЧИ

Анализ проектируемой конструкции на прочность и долговечность обычно сводится к моделированию ее напряженно-деформированного состояния в процессе эксплуатации, параметры которого в статике можно найти, минимизируя следующий энергетический функционал (вариационный принцип Лагранжа) [1]:

$$\int_{\Omega} \sigma_{ij} \varepsilon_{ij} d\Omega - \int_{\Omega} X_i u_i d\Omega - \int_{\Gamma} \bar{X}_i u_i d\Gamma = 0. \quad (1)$$

Одним из наиболее эффективных способов минимизации функционала (1) является использование © Чопоров С. В., Игнатченко М. С., Кудин А. В., Кривохага А. Г., Гоменюк С. И., 2021
DOI 10.15588/1607-3274-2021-2-17

МКЭ. Для его применения исходную область Ω разбивают на некоторое множество КЭ заданной формы $\Omega_k, k = \overline{1, N}$. Аппроксимация искомой функции (например, компонент вектора перемещений) для заданного КЭ описывается следующим соотношением:

$$u_i(\mathbf{x}) = \sum_{l=1}^M u_i^l N_l(\mathbf{x}). \quad (2)$$

Как функции формы чаще всего используют полиномы, коэффициенты которых определяются из следующего соотношения:

$$N_l(x_m) = \begin{cases} 1, & l = m, \\ 0, & l \neq m. \end{cases}$$

Подставляя (2) в (1) с учетом соотношений Коши и обобщенного закона Гука, получим соотношение, которое в матричном виде можно записать так:

$$\int_{\Omega_k} \mathbf{B}^T \mathbf{D} \mathbf{B} d\Omega_k = \int_{\Omega_k} \mathbf{N}^T \mathbf{p} d\Omega_k + \int_{\Gamma_k} \mathbf{N}^T \mathbf{q} d\Gamma_k. \quad (3)$$

Минимизация соотношения (3) при помощи, например, метода вариации произвольных постоянных позволяет окончательно получить СЛАУ, которую в матричной форме можно записать следующим образом:

$$\mathbf{K} \mathbf{u} = \mathbf{F}. \quad (4)$$

Таким образом, в общем виде возникает задача создания такого ПО, которое бы автоматизировало все стадии решения задачи с помощью МКЭ: от разбиения исходной области на КЭ заданного типа, до решения СЛАУ вида (4). При этом существенный интерес представляет обеспечение возможности создания лаконичного кода и простоты изменения типов контейнеров и алгоритмов.

2 ОБЗОР ЛИТЕРАТУРЫ

Обобщенное программирование – это парадигма разработки ПО, базирующаяся на идее написания алгоритмов таким образом, чтобы они не зависели от типов обрабатываемых данных. Например, на языке программирования C++ обобщенную (шаблонную) процедуру поиска максимального значения среди двух объектов можно описать при помощи кодов:

```
template <typename T> T max(const T& lhs, const T& rhs)
{
    return lhs < rhs ? rhs : lhs;
}
```

Эта функция может быть использована для любых типов данных T, для которых определена операция сравнения «меньше».

Общую теорию использования обобщенного программирования описано в работе [11]. Методологию написания программ с использованием шаблонных классов на языке C++ рассмотрено в [12–14].

Одними из первых публикаций, посвященных использованию парадигмы обобщенного программирования при разработке систем конечно-элементного анализа, являются работы [15, 16]. В них исследованы самые общие принципы применения шаблонов при реализации базовых контейнерных классов и итераторов.

В [17] рассмотрена обобщенная реализация низкоуровневых матричных операций, используемых при создании ПО для МКЭ.

Принципы реализации шаблонных классов для параллельной реализации векторов большой длины предложены в [18].

В работе [19] приведены примеры использования элементов функционального и обобщенного программирования для реализации разных типов КЭ и базовых операций над ними: вычисление функций формы, их производных, а также различных вспомогательных операций над ними. Основной упор сделан на реализацию векторных вычислений с использованием соответствующих возможностей современных графических процессоров.

Повышение эффективности конечно-элементных вычислений при использовании обобщенного программирования обосновывается в [20].

Различные варианты обобщенной реализации структур данных и алгоритмов для описания КЭ, дискретных моделей, параллельных вычислений и т.п. приведены в [21, 22].

В большинстве известных работ обобщенное программирование применяется для создания низкоуровневых процедур, реализующих вспомогательные операции или структуры данных, необходимые для разработки конечно-элементного ПО.

Таким образом, разработка высокоуровневых подходов для реализации новых типов КЭ или алгоритмов расчета является актуальной.

3 МАТЕРИАЛЫ И МЕТОДЫ

Алгоритм применения МКЭ можно условно разделить на два основных этапа: 1) формирование локальных матриц жесткости, массы и сопротивления (демпфирования) для каждого КЭ и добавление их к соответствующим глобальным матрицам (ансамблирования); 2) учета краевых условий, формирования СЛАУ и ее решения.

Первый этап – формирование глобальных матриц может быть достаточно нетривиальной и длительной процедурой в случае использования специальных типов КЭ, например, оболочечных. Поэтому его программная реализация является важнейшей частью разработки любого ПО для конечно-элементного анализа. На практике требуется создание программ с такой архитектурой, которая бы позволяла сравнительно легкое добавление к уже существующей библиотеке новые типы КЭ, предназначенные для решения

специальных классов задач (многослойных оболочек, композитов, контактного взаимодействия и т.д.) или оптимизированные по скорости вычислений.

Наиболее распространенными на практике формами КЭ являются треугольные и четырехугольные элементы в плоском случае, а также тетраэдральные и кубические элементы в трехмерном пространстве. Большинство автоматических генераторов сеток ориентированы на дискретизацию исходных геометрических областей на элементы именно таких типов.

Для описания формы наиболее распространенных на практике изопараметрических КЭ предлагается следующая иерархия классов, UML-диаграмма которой приведена на рис. 1.

Здесь базовый абстрактный класс Shape описывает наиболее общие свойства формы изопараметрического элемента, такие как количество его узлов, коэффициенты функции формы, параметры квадратур для численного интегрирования и т.п. Кроме того, данный класс содержит набор методов, предназначенных для вычисления коэффициентов функций формы, Якобиана, значений функций формы и ее производных в заданных точках и т.п. Большинство этих методов являются абстрактными и реализованы в производных от Shape классах.

Классы Shape1D2, Shape2D3 и Shape3D4 инкапсулируют наиболее распространенные линейные формы КЭ: стержень, треугольник и тетраэдр. Производные от них классы реализуют либо элементы более высоких степеней аппроксимации (Shape1D3 – квадратичный трехузловой стержень; Shape2D6 – квадратичный шестиузловой треугольник; Shape3D10 – квадратичный десятиузловой тетраэдр), либо элементы другой формы (Shape2D4 – четырехугольник; Shape3D8 – кубический элемент).

Соответствующую иерархию классов шаблонов, описывающих различные типы КЭ, можно представить следующим образом (рис. 2).

Здесь базовый абстрактный класс FE описывает наиболее общие свойства и методы КЭ: его форму, количество степеней свободы, упругие и физические характеристики, толщину, локальные матрицы жесткости, массы и сопротивления и т.п. В данном классе также задекларирован абстрактный метод генерации локальных матриц.

Производные от FE классы шаблонов содержат реализацию процедуры построения локальных матриц для конкретного типа элемента. В качестве параметра они принимают форму элемента. Например, класс FE2DP<Shape2D3> описывает конечный элемент пластины треугольной формы с линейной аппроксимацией, а FE2DP<Shape2D4> – элемент четырехугольной формы с билинейной аппроксимацией.

Такой способ описания классов, реализующих изопараметрические КЭ, позволяет, с одной стороны, отделить реализацию формы элемента от алгоритмов вычисления локальных матриц (численного интегрирования), а с другой – существенно упростить создание элементов новой формы и/или типа.

Второй этап применения МКЭ – решение СЛАУ также является достаточно длительной и ресурсоемкой процедурой, особенно при решении сверхбольших систем уравнений, возникающих, например, при создании новой аэрокосмической техники.

Совместно с МКЭ обычно используются два типа методов решения СЛАУ: прямые и итерационные. Прямые методы позволяют получить точное решение (без учета машинной погрешности), но обычно они требуют значительных объемов оперативной памяти

компьютера и могут быть неэффективными при решении больших СЛАУ. Приближенные итерационные методы обычно требуют существенно меньше оперативной памяти, поэтому часто используются для решения систем уравнений высоких порядков, хотя время вычислений при этом может быть значительным. Кроме того, для многих существующих методов решения СЛАУ в настоящее время разрабатываются параллельные реализации для компьютерных систем как с общей, так и распределенной памятью.

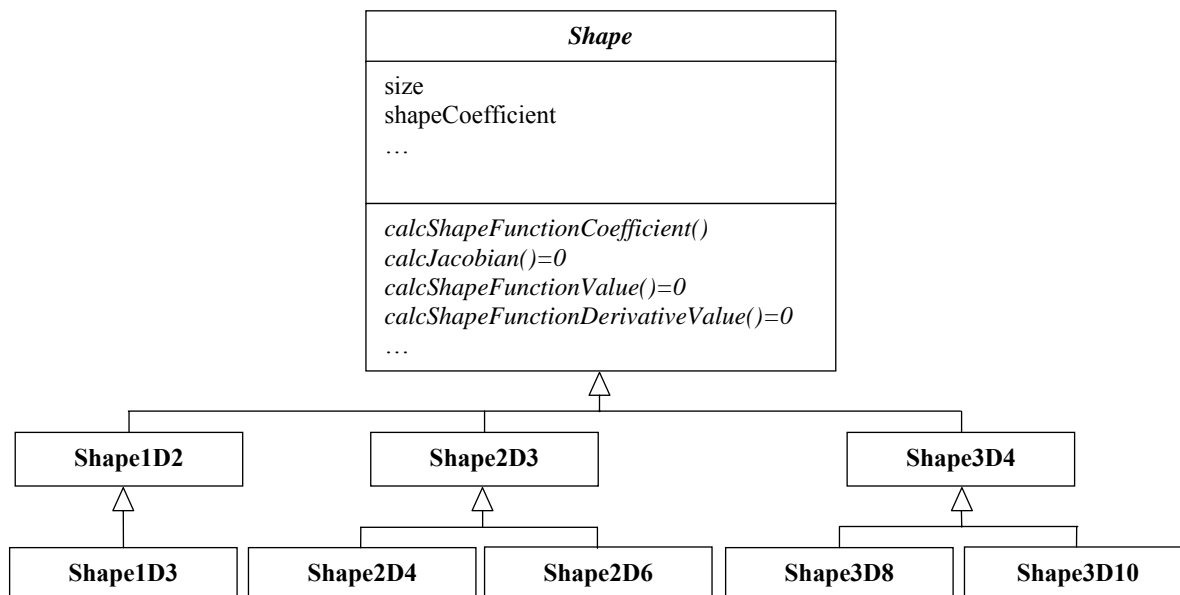


Рисунок 1 – Иерархия классов, описывающих форму конечного элемента

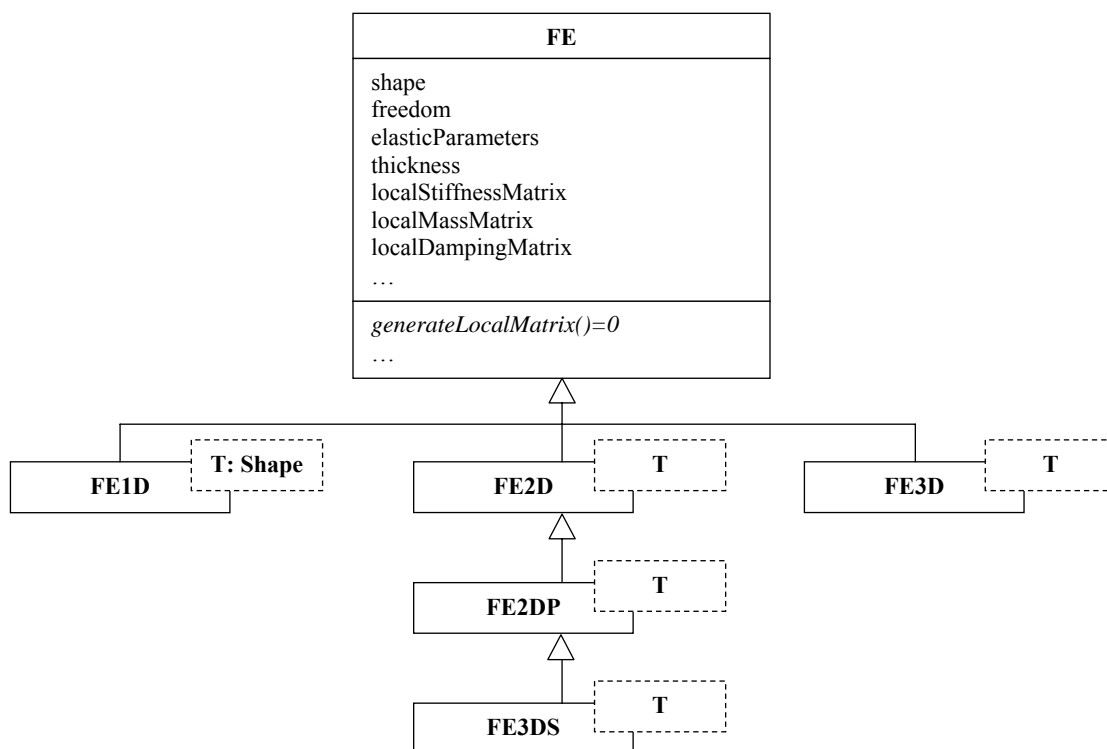


Рисунок 2 – Иерархия классов шаблонов, описывающих различные типы конечных элементов

Таким образом, возникает задача разработки такой реализации систем конечно-элементного анализа, в которой алгоритмы решения задачи с использованием МКЭ не зависели бы от метода решения СЛАУ. Для этого предлагается обобщенный подход, когда метод решения СЛАУ является параметром алгоритма расчета задачи с использованием МКЭ.

Для этого, в первую очередь, нужно создать абстрактный класс-решатель, инкапсулирующий все необходимые операции для формирования СЛАУ в МКЭ: добавление элемента в матрицу; добавление элемента в вектор столбец правой части; учета краевых условий и т.д.

Диаграмма иерархии классов, реализующих решатель СЛАУ, приведена на рис. 3.

Здесь параметром класса шаблона Solver<T> является обобщенный тип T, инстанцируемый как реализованный в каждом конкретном решателе контейнер для хранения матрицы коэффициентов СЛАУ.

Для использования классов шаблонов, производных от Solver<T>, предлагается следующая иерархическая структура, инкапсулирующая различные алгоритмы применения МКЭ (рис. 4).

Здесь абстрактный класс FEM описывает базовые для МКЭ атрибуты, такие как: параметры расчета; конечно-элементную сетку; конечный элемент, абстрактную процедуру расчета и т.п.

Производные от FEM классы шаблонов реализуют определенные виды расчетов (например, линейную статику) и в качестве параметра принимают шаблоны следующих классов: решателя СЛАУ, производного

от Solver; конечного элемента, производного от FE; формы конечного элемента, производного от Shape.

4 ЭКСПЕРИМЕНТЫ

Для изучения возможностей предложенного обобщенного подхода к организации конечно-элементных классов на языке C++ была программно реализована библиотека QFEM [23], с помощью которой было выполнено численное решение ряда практических задач.

Вначале была рассмотрена тестовая задача о нахождении прогибов квадратной пластины с жестко заземленными краями, находящейся под действием равномерно распределенной поверхностной нагрузки. Были выбраны следующие параметры пластины: сторона – 1 м; толщина – 0,01 м; модуль Юнга – 203200 МПа; коэффициент Пуассона – 0,27; нагрузка – 0,05 МН/м².

В простейшем случае программа решения такой задачи с использованием библиотеки QFEM может быть реализована таким образом:

```
#include "fem/femstatic.h"
#include "analyse/analyse.h"
#include "solver/eigensolver.h"

// Инициализация системы сообщений
TMessenger *msg = new Tmessenger();
int main()
{
```

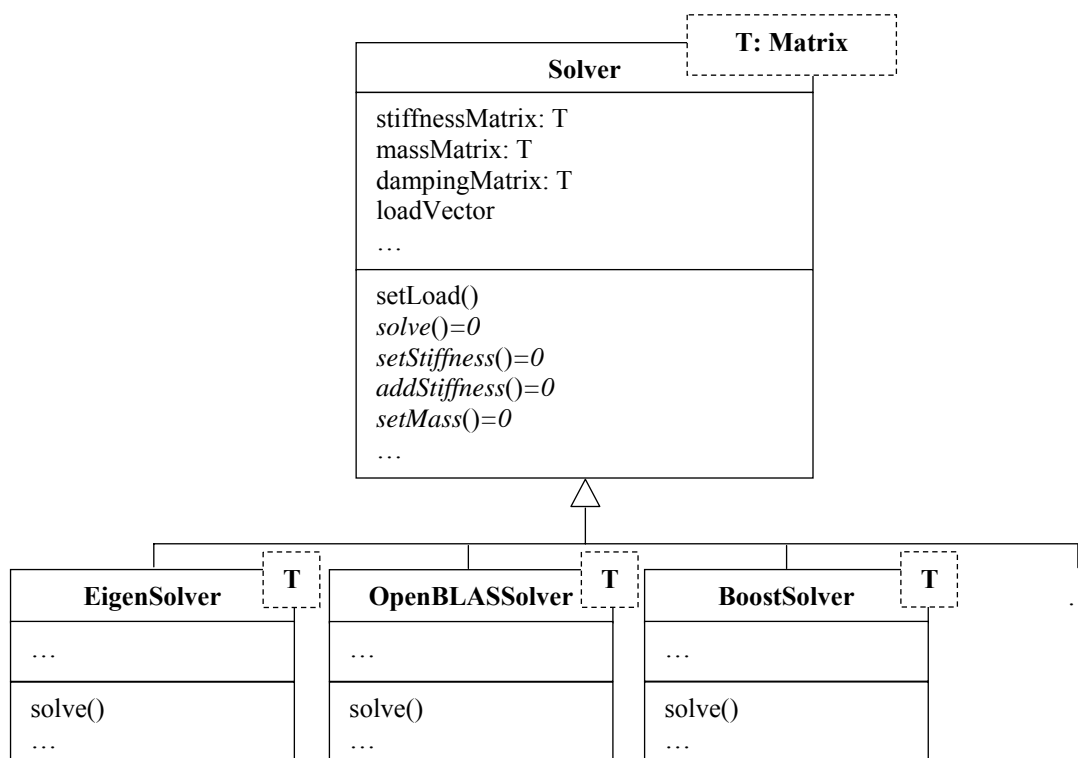


Рисунок 3 – Обобщенная реализация классов для решения СЛАУ

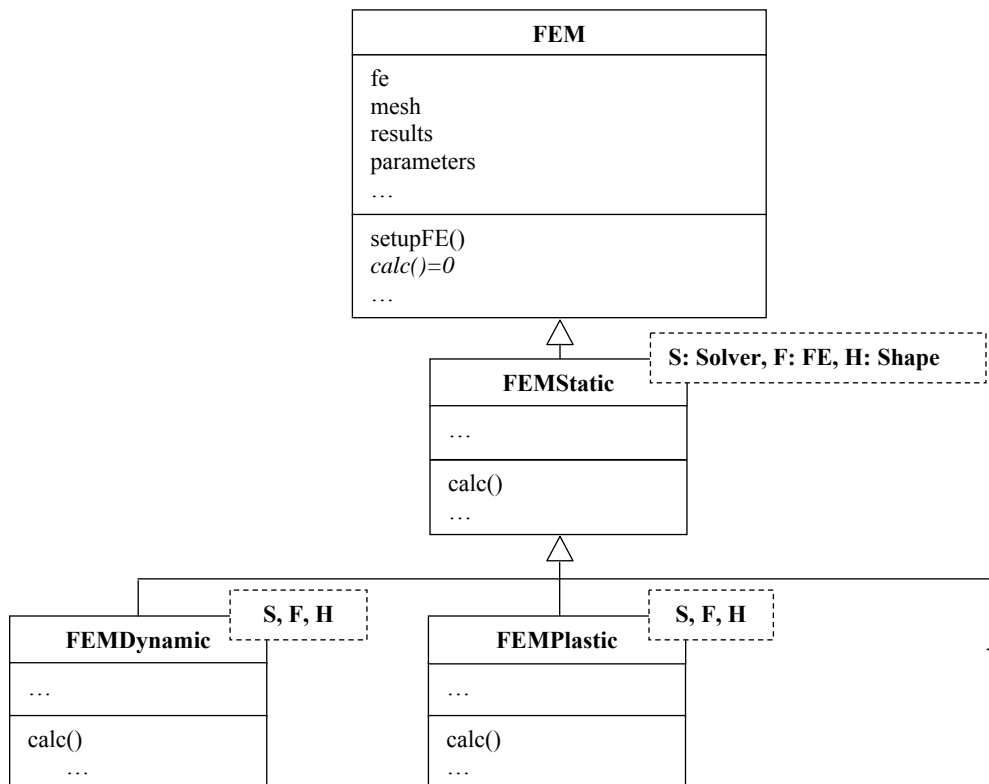


Рисунок 4 – Диаграмма классов шаблонов, реализующих различные алгоритмы МКЭ

```

// Конечно-элементная модель пластины
matrix<double> x = {{-0.5, -0.5}, {0, -0.5},
                  {0.5, -0.5}, {0.5, 0},
                  {0.5, 0.5}, {0, 0.5},
                  {-0.5, 0.5}, {-0.5, 0},
                  {0, 0}};
matrix<int> fe = {{0, 1, 8}, {0, 8, 7}, {1, 2, 8},
                 {2, 3, 8}, {3, 4, 8}, {4, 5, 8},
                 {5, 6, 8}, {6, 7, 8}};
be = fe;
TMesh mesh(FEType::fe2d3p, x, fe, be);
TResults result;
TFEMParams params;
TFEMStatic<TEigenSolver, TShape2D3,
           TFE2DP> fem("Plate", &mesh, &result);

params.addYoungModulus(203200);
params.addPoissonRatio(0.27);
params.addThickness(0.01);
params.addBoundaryCondition(Direction::X |
                             Direction::Y | Direction::Z, 0,
                             "x == -0.5 or x == 0.5 or y == -0.5 or y == 0.5");
params.addPressureLoad(0.05);

fem.setParams(params);
fem.startProcess();

delete msg;
return 0;
}
    
```

В приведенном примере выполняется расчет пластины, состоящей из 8 линейных треугольных КЭ. Для расчета пластины, состоящей из четырехуголь-

ных КЭ, вышеприведенную программу нужно переделать следующим образом:

```

...
matrix<int> fe = {{0, 1, 8, 7}, {1, 2, 3, 8},
                 {3, 4, 5, 8}, {5, 6, 7, 8}};
TFEMStatic<TEigenSolver, TShape2D4,
           TFE2DP> fem("Plate", &mesh, &result);
...
    
```

Расчет данной задачи выполнялся также при использовании сетки, состоящей из 2173 узлов и 1046 КЭ в форме квадратичного (шестиузлового) треугольника, а также сетки из 40401 узла и 40000 четырехугольных билинейных элементов. Полученное распределение прогибов пластины приведено на рис. 5.

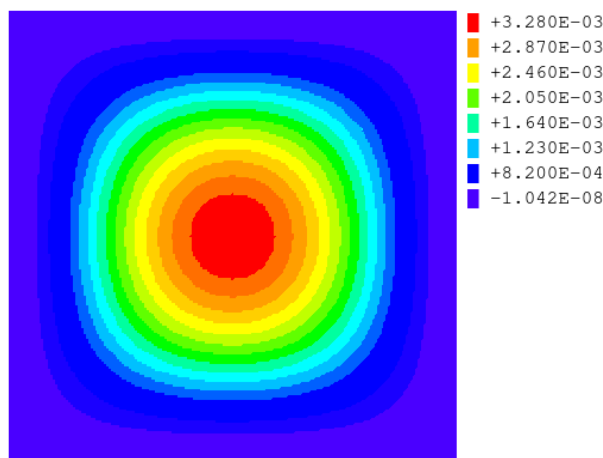


Рисунок 5 – Распределение прогибов по пластине

Расчетный максимальный прогиб пластины – 0,00343657 для шестиузловых треугольных элементов и 0,00327985 для четырехугольных элементов. Погрешность полученного численного результата по сравнению с имеющимся аналитическим решением [24] составляет 0,36%.

Следует отметить, что решение СЛАУ, возникающих при использовании МКЭ, является достаточно сложной задачей, т.к. в случае использования специальных типов КЭ могут возникать плохо обусловленные матрицы. На практике зачастую не всегда возможно заранее выбрать оптимальный метод решения СЛАУ, что приводит к необходимости проведения различных вычислительных экспериментов и подбору оптимального метода решения конкретной СЛАУ. Предложенный обобщенный подход к описанию класса, реализующего конечно-элементный расчет, где метод решения системы уравнений фактически является его параметром, позволяет существенно упростить задачу подбора оптимальных параметров решателя.

Так, например, при численном расчете вышеописанной задачи на компьютере с процессором AMD Ryzen 7 2700X Eight-Core Processor 3.70 GHz и объемом памяти 32 ГБ под управлением ОС Windows x64 при использовании сетки, состоящей из 40401 узлов и 40000 четырехугольных КЭ, время решения СЛАУ разными методами может значительно различаться (табл. 1).

Таким образом, обобщенная реализация решателя конечно-элементных классов позволяет без изменения структуры библиотеки легко добавлять новые реализации методов решения СЛАУ с последующей оптимизацией выбора метода решения для конкретной задачи.

5 РЕЗУЛЬТАТЫ

Для верификации предложенного подхода был выполнен численный анализ напряженно-деформированного состояния тонкостенной трубы с жестко заземленными краями, находящейся под действием внутреннего давления.

Задача решалась при следующих параметрах: радиус трубы – 1,99 м; длина – 4,014 м; толщина стенки – 0,0369 м; модуль Юнга – 203200 МПа; коэффи-

циент Пуассона – 0,27; равномерно распределенная по внутренней поверхности нагрузка – 0,05 МН/м².

Результаты проведенных экспериментов при использовании различных типов КЭ приведены в табл. 2.

На рис. 6 приведено распределение радиальных прогибов по трубе с визуализацией деформирования конструкции под действием нагрузки.

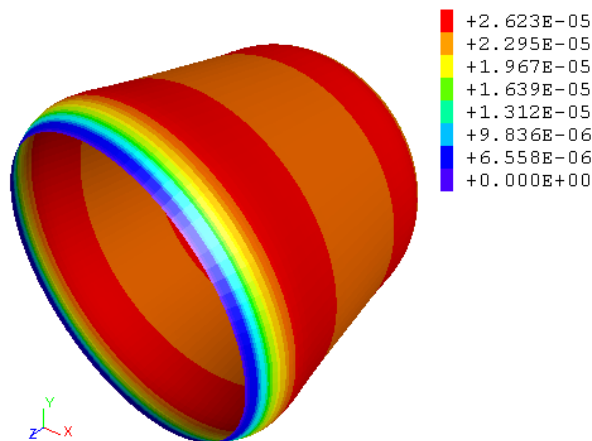


Рисунок 6 – Распределение прогибов по трубе

6 ОБСУЖДЕНИЕ

Как видно из табл. 2 использование различных типов КЭ для данной тестовой задачи дает расхождение между результатами не более 1,2%, что может свидетельствовать о достоверности полученного решения.

Предложенная структура классов шаблонов, описывающих КЭ, по сравнению с [19, 20, 22] позволяет существенно упростить процесс реализации новых типов элементов за счет разделения их формы и алгоритмов построения локальных матриц жесткости, массы и демпфирования.

По сравнению с [18, 21] предложенная структура классов шаблонов, разделяющая алгоритмы применения МКЭ и структуры данных для хранения разреженных матриц и решения соответствующих им СЛАУ, позволяет без изменения структуры ПО добавлять новые методы решения СЛАУ, что дает возможность, например, использовать параллельные компьютерные системы при решении задач высоких порядков.

Таблица 1 – Сравнение методов решения СЛАУ

Метод	Тип	Библиотека	Время расчета, с
Градиентный спуск	Итерационный	Eigen	1286
Метод Ланцоша	Итерационный	QFEM	870
LLT-факторизация	Прямой	Eigen	22
Прямое LLT-разложение	Прямой	Intel MKL	1

Таблица 2 – Результаты численных экспериментов

Тип КЭ	Количество узлов	Количество КЭ	Максимальный прогиб, м
Линейный тетраэдр	11845	34783	$2,63181 \times 10^{-5}$
Квадратичный тетраэдр	70317	34783	$2,61435 \times 10^{-5}$
Линейный оболочечный треугольник	28244	55984	$2,65075 \times 10^{-5}$
Билинейный оболочечный четырехугольник	5100	5000	$2,62280 \times 10^{-5}$
Квадратичный оболочечный треугольник	25752	12748	$2,62033 \times 10^{-5}$

Использование предложенной архитектуры ПО для конечно-элементного анализа по сравнению с [21], а также со стандартными объектно-ориентированными или функциональными подходами позволяет сократить затраты времени как на саму разработку программ, так и на последующее их сопровождение, а именно: добавление новых типов КЭ, методов решения СЛАУ, а также алгоритмов решения с помощью МКЭ новых классов задач.

При этом использование обобщенного подхода к разделению данных и алгоритмов позволяет уменьшить количество возникающих при разработке ошибок.

ВЫВОДЫ

Решена актуальная задача создания обобщенной архитектуры программ для конечно-элементного анализа, позволяющей уменьшить затраты времени на разработку ПО за счет разделения структур данных и алгоритмов их разработки, что дает возможность также снизить вероятность появления ошибок в программном коде.

Научная новизна полученных результатов состоит в том, что:

– впервые предложен обобщенный подход к разработке классов, реализующих КЭ, в которых форма элемента является параметром алгоритмов построения локальных матриц жесткости, массы и демпфирования, что позволяет значительно упростить процедуру добавления новых типов КЭ к уже имеющимся и уменьшить вероятность появления ошибок в их программной реализации;

– впервые предложен обобщенный подход к реализации процедуры решения СЛАУ, позволяющий за счет создания высокоуровневой обертки для различных методов решения систем уравнения унифицировать структуры данных и алгоритмы для хранения больших разреженных матриц, возникающих при применении МКЭ.

Таким образом, впервые предложена обобщенная архитектура ПО для конечно-элементного анализа, в которой тип и форма КЭ, а также метод решения СЛАУ являются параметрами алгоритмов решения задач с использованием МКЭ. Такой подход позволяет без изменения структуры ПО и алгоритмов расчета добавлять новые специализированные типы КЭ и методы решения СЛАУ, предназначенные как для решения новых классов задач, так и для использования конкретных возможностей вычислительной техники (параллельных вычисления, использования GPU и т.п.).

Практическая ценность полученных результатов состоит в том, что на языке C++ реализована библиотека классов шаблонов для решения различных задач механики (статика, динамика, пластичность), а также проведено экспериментальное исследование, подтвердившее работоспособность предложенных обобщенных подходов при решении ряда практических задач. Разработанное ПО может быть рекомендовано

для решения актуальных задач проектирования современной техники.

Перспективы дальнейших исследований состоят в том, чтобы расширить возможности разработанной библиотеки классов для использования параллельных компьютерных систем как с общей, так и распределенной памятью.

БЛАГОДАРНОСТИ

Работа выполнена в рамках государственной научно-исследовательской темы «Математическое и программное обеспечение автоматизированного проектирования аэрокосмической техники» (гос. рег. № 0118U000210) Запорожского национального университета.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Zienkiewicz O. C. The Finite Element Method: Its Basis and Fundamentals. Sixth edition / O. C. Zienkiewicz, R. L. Taylor, J. Z. Zhu. – Oxford : Butterworth-Heinemann, 2016. – 753 p. DOI: 10.1002/nme.1760
2. Madenci E. The Finite Element Method and Applications in Engineering Using ANSYS® / E. Madenci, I. Guven. – Boston : Springer, 2015. – 651 p. DOI: 10.1007/978-1-4899-7550-8
3. Zimmerman W. Multiphysics Modeling with Finite Element Methods / W. Zimmerman. – Singapore : World Scientific, 2006. – 432 p. DOI: 10.1142/6141
4. MSC Nastran – Multidisciplinary Structural Analysis. [Electronic resource] – Access mode: <https://www.mssoftware.com/product/msc-nastran>
5. Top Finite Element Analysis (FEA) Software: List, Reviews, Comparison & Price | TEC. [Electronic resource] – Access mode: <https://www3.technologyevaluation.com/sd/category/finite-element-analysis-fea>
6. FreeFEM – An open-source PDE Solver using The Finite Element Method: [Electronic resource] – Access mode: <https://freefem.org/>
7. GetFEM Homepage – GetFEM. [Electronic resource] – Access mode: <http://getfem.org/>
8. OpenCAD The Programmers Solid 3D CAD Modeller. [Electronic resource] – Access mode: <https://www.openscad.org/>
9. The deal.II Finite Element Library. [Electronic resource] – Access mode: <https://www.dealii.org/>
10. Netgen/NGSolve. [Electronic resource] – Access mode: <https://ngsolve.org/>
11. Stepanov A. A. From Mathematics to Generic Programming / A. A. Stepanov, D. E. Rose. – Crawfordsville : Addison-Wesley, 2014. – 292 p.
12. Meyers S. Effective Modern C++: 42 Specific Ways to Improve Your Use of C++11 and C++14 / S. Meyers. – Sebastopol : O'Reilly, 2017. – 316 p.
13. Galowicz J. C++17 STL Cookbook / J. Galowicz. – Birmingham : Packt Publishing Ltd, 2017. – 504 p.
14. Vandevoorde D. C++ Templates: The Complete Guide / D. Vandevoorde, N. M. Josuttis, D. Gregor. – Boston : Addison-Wesley, 2017. – 788 p.
15. Bittencourt M. L. Using C++ templates to implement finite element classes / M. L. Bittencourt // Engineering Computation. – 2000. – Vol. 17(7). – P. 775–788. DOI: 10.1108/02644400010352243
16. Bangerth W. Using Modern Features of C++ for Adaptive Finite Element Methods: Dimension-Independent Programming in deal.II / W. Bangerth // Proceedings of the 16th IMACS

- World Congress. – 2000. [Electronic resource] – Access mode: <https://www.dealii.org/7.3.0/>
17. FEMEngine: finite element method C++ code based on functional and template metaprogramming / [A. Gurin, A. Baykin, T. Polyansky et al] // IEEE Conference Proceedings. – 2019. – Vol. 2019. – P. 92–96. DOI: 10.1109/ISPRAS47671.2019.00020
18. Plagne L. Parallel expression template for large vectors / L. Plagne, F. Hülsemann // Conference: Workshop on Parallel/High-Performance Object-Oriented Scientific. – 2009. [Electronic resource] – Access mode: https://www.researchgate.net/publication/228965935_Parallel_expression_template_for_large_vectors
DOI: 10.1145/1595655.1595663
19. Schöberl J. C++11 Implementation of Finite Elements in NGSolve / J. Schöberl // ASC Report. – 2014. – No. 30. [Electronic resource] – Access mode: <https://www.asc.tuwien.ac.at/~schoeberl/wiki/publications/ngs-cpp11.pdf>
20. Rupp K. Increased efficiency in finite element computations through template metaprogramming / K. Rupp // Conference: Proceedings of the 2010 Spring Simulation Multiconference. – 2010. [Electronic resource] – Access mode: https://www.iue.tuwien.ac.at/pdf/ib_2010/CP2010_Rupp_2.pdf
DOI: 10.1145/1878537.1878633
21. Cirak F. Generic programming techniques for parallelizing and extending procedural finite element programs / F. Cirak, J. C. Cummings // Engineering with Computers. – 2008. [Electronic resource] – Access mode: https://www.researchgate.net/publication/220677624_Generic_programming_techniques_for_parallelizing_and_extending_procedural_finite_element_programs DOI: 10.1007/s00366-007-0058-x
22. Bastian P. Generic implementation of finite element methods in the Distributed and Unified Numerics Environment (DUNE) / P. Bastian, F. Heimann, S. Marnach // Kybernetika. – 2010. – Vol. 46. P. 294–315.
23. The Simple FEM Solver. [Electronic resource] – Access mode: <https://github.com/SeregaGomen/QFEM>
24. Timoshenko S. Theory of plates and Shells / S. Timoshenko, S. Woinowsky-Krieger. – Singapore : McGraw-Hill, Inc., 1964. – 594 p.
- Статья поступила в редакцию 15.02.2021.
После доработки 09.04.2021.

УДК 004.9:004.94:51

РЕАЛІЗАЦІЯ СКІНЧЕННО-ЕЛЕМЕНТНОЇ БІБЛІОТЕКИ КЛАСІВ З ВИКОРИСТАННЯМ УЗАГАЛЬНЕНОГО ПРОГРАМУВАННЯ

Чопоров С. В. – д-р техн. наук, професор, професор кафедри програмної інженерії Запорізького національного університету, Запоріжжя, Україна.

Ігнатченко М. С. – аспірант кафедри програмної інженерії Запорізького національного університету, Запоріжжя, Україна.

Кудін О. В. – канд. фіз.-мат. наук, доцент кафедри програмної інженерії Запорізького національного університету, Запоріжжя, Україна.

Кривохата А. Г. – канд. фіз.-мат. наук, старший викладач кафедри програмної інженерії Запорізького національного університету, Запоріжжя, Україна.

Гоменюк С. І. – д-р техн. наук, професор, декан математичного факультету, професор кафедри програмної інженерії Запорізького національного університету, Запоріжжя, Україна.

АНОТАЦІЯ

Актуальність. Для комп'ютерного моделювання складних об'єктів і явищ різної природи на практиці часто застосовують чисельний метод скінченних елементів. Його програмна реалізація (особливо для дослідження нових класів задач) є досить трудомістким процесом.

Висока вартість розробки програмного забезпечення обумовлює актуальність розробки нових підходів до підвищення ефективності програмування й супроводу (у т.ч. додавання нових функцій).

Мета. Мета роботи – створення нової ефективної архітектури програм скінченно-елементного аналізу проблем математичної фізики, що дозволяє легко розширювати їх функціональність для розв'язання нових класів задач.

Метод. Запропоновано метод розробки програм для скінченно-елементного аналізу з використанням узагальненого програмування, що дає можливість істотно спростити архітектуру програмного забезпечення й зробити його більш зручним для супроводу і модифікації за рахунок розділення алгоритмів і структур даних.

Запропоновано нову архітектуру класів, які реалізують скінченно-елементний розрахунок, що дозволяє легко розширювати функціональність програм за рахунок додавання нових типів скінченних елементів, методів розв'язання систем лінійних алгебраїчних рівняння, паралельних обчислень тощо.

Результати. Запропонований підхід було програмно реалізовано у вигляді бібліотеки класів на мові C++. Проведено ряд обчислювальних експериментів, які підтвердили його працездатність при вирішенні низки практичних задач.

Висновки. Розроблений підхід може бути використаний як для створення систем скінченно-елементного аналізу загальної призначення з відкритою архітектурою, так і для реалізації спеціалізованих програмних пакетів, орієнтованих на розв'язання конкретних класів задач (механіки руйнування, еластомерів, контактної взаємодії і т.д.).

КЛЮЧОВІ СЛОВА: метод скінченних елементів, скінченний елемент, об'єктно-орієнтоване програмування, узагальнене програмування, архітектура програмного забезпечення.

UDC 004.9:004.94:51

IMPLEMENTATION OF A FINITE ELEMENT CLASS LIBRARY USING GENERALIZED PROGRAMMING

Choporov S. V. – Dr. Sc., Professor, Professor of the Software Engineering Department, Zaporizhzhia National University, Zaporizhzhia, Ukraine.

Ihnatchenko M. S. – Postgraduate student of the Software Engineering Department, Zaporizhzhia National University, Zaporizhzhia, Ukraine.

Kudin O. V. – PhD, Associate Professor of the Software Engineering Department, Zaporizhzhia National University, Zaporizhzhia, Ukraine.

© Чопоров С. В., Ігнатченко М. С., Кудин О. В., Кривохата А. Г., Гоменюк С. І., 2021
DOI 10.15588/1607-3274-2021-2-17

Kryvokhata A. G. – PhD, Senior Lecturer of the Software Engineering Department, Zaporizhzhia National University, Zaporizhzhia, Ukraine.

Homeniuk S. I. – Dr. Sc., Professor, Dean of the Faculty of Mathematics, Professor of the Software Engineering Department, Zaporizhzhia National University, Zaporizhzhia, Ukraine.

ABSTRACT

Context. For computer modeling of complex objects and phenomena of various nature, in practice, the numerical finite element method is often used. Its software implementation (especially for the study of new classes of problems) is a rather laborious process.

The high cost of software development makes the development of new approaches to improving the efficiency of programming and maintenance (including the addition of new functions) urgent.

Objective. The aim of the work is to create a new effective architecture of programs for finite element analysis of problems in mathematical physics, which makes it easy to expand their functionality to solve new classes of problems.

Method. A method for developing programs for finite element analysis using generalized programming is proposed, which makes it possible to significantly simplify the architecture of the software and make it more convenient for maintenance and modification by separating algorithms and data structures.

A new architecture of classes that implement finite element calculation is proposed, which makes it possible to easily expand the functionality of programs by adding new types of finite elements, methods for solving systems of linear algebraic equations, parallel computations, etc.

Results. The proposed approach was implemented in software as a class library in C++. A number of computational experiments have been carried out, which have confirmed its efficiency in solving practical problems.

Conclusions. The developed approach can be used both to create general-purpose finite element analysis systems with an open architecture, and to implement specialized software packages focused on solving specific classes of problems (fracture mechanics, elastomers, contact interaction, etc.).

KEYWORDS: finite element method, finite element, object-oriented programming, generic programming, software architecture.

REFERENCES

1. Zienkiewicz O. C., Taylor R. L., Zhu J. Z. *The Finite Element Method: Its Basis and Fundamentals*. Sixth edition. Oxford, Butterworth-Heinemann, 2016, 753 p. DOI: 10.1002/nme.1760
2. Madenci E., Guven I. *The Finite Element Method and Applications in Engineering Using ANSYS®*. Boston, Springer, 2015, 651 p. DOI: 10.1007/978-1-4899-7550-8
3. Zimmerman W. *Multiphysics Modeling with Finite Element Methods*. Singapore, World Scientific, 2006, 432 p. DOI: 10.1142/6141
4. MSC Nastran – Multidisciplinary Structural Analysis. [Electronic resource]. Access mode: <https://www.mssoftware.com/product/msc-nastran>
5. Top Finite Element Analysis (FEA) Software: List, Reviews, Comparison & Price | TEC. [Electronic resource]. Access mode: <https://www3.technologyevaluation.com/sd/category/finite-element-analysis-fea>
6. FreeFEM – An open-source PDE Solver using The Finite Element Method. [Electronic resource]. Access mode: <https://freefem.org/>
7. GetFEM Homepage – GetFEM. [Electronic resource]. Access mode: <http://getfem.org/>
8. OpenCAD The Programmers Solid 3D CAD Modeller. [Electronic resource]. Access mode: <https://www.openscad.org/>
9. The deal.II Finite Element Library. [Electronic resource]. Access mode: <https://www.dealii.org/>
10. Netgen/NGSolve. [Electronic resource]. Access mode: <https://ngsolve.org/>
11. Stepanov A. A., Rose D. E. *From Mathematics to Generic Programming*. Crawfordsville, Addison-Wesley, 2014, 292 p.
12. Meyers S. *Effective Modern C++: 42 Specific Ways to Improve Your Use of C++11 and C++14* / S. Sebastopol, O'Reilly, 2017, 316 p.
13. Galowicz J. *C++17 STL Cookbook*. Birmingham, Packt Publishing Ltd, 2017, 504 p.
14. Vandevoorde D., Josuttis N. M., Gregor D. *C++ Templates: The Complete Guide*. Boston, Addison-Wesley, 2017, 788 p.
15. Bittencourt M. L. Using C++ templates to implement finite element classes, *Engineering Computation*, 2000, Vol. 17(7), pp. 775–788. DOI: 10.1108/02644400010352243
16. Bangerth W. Using Modern Features of C++ for Adaptive Finite Element Methods: Dimension-Independent Programming in deal.II, *Proceedings of the 16th IMACS World Congress*, 2000, [Electronic resource], Access mode: <https://www.dealii.org/7.3.0/>
17. Gurin A., Baykin A., Polyansky T. et al. FEMEngine: finite element method C++ code based on functional and template metaprogramming, *IEEE Conference Proceedings*, 2019, Vol. 2019, pp. 92–96. DOI: 10.1109/ISPRAS47671.2019.00020
18. Plagne L., Hülsemann F. Parallel expression template for large vectors, *Conference: Workshop on Parallel/High-Performance Object-Oriented Scientific*, 2009, [Electronic resource], Access mode: https://www.researchgate.net/publication/228965935_Parallel_expression_template_for_large_vectors DOI: 10.1145/1595655.1595663
19. Schöberl J. C++11 Implementation of Finite Elements in NGSolve, *ASC Report*, 2014, No. 30. [Electronic resource]. Access mode: <https://www.asc.tuwien.ac.at/~schoeberl/wiki/publications/ngs-cpp11.pdf>
20. Rupp K. Increased efficiency in finite element computations through template metaprogramming, *Conference: Proceedings of the 2010 Spring Simulation Multiconference*. 2010, [Electronic resource]. Access mode: https://www.iue.tuwien.ac.at/pdf/ib_2010/CP2010_Rupp_2.pdf DOI: 10.1145/1878537.1878633
21. Cirac F., Cummings J. C. Generic programming techniques for parallelizing and extending procedural finite element programs, *Engineering with Computers*, 2008. [Electronic resource]. Access mode: https://www.researchgate.net/publication/220677624_Generic_programming_techniques_for_parallelizing_and_extending_procedural_finite_element_programs DOI: 10.1007/s00366-007-0058-x
22. Bastian P., Heimann F., Marnach S. Generic implementation of finite element methods in the Distributed and Unified Numerics Environment (DUNE), *Kybernetika*, 2010, Vol. 46, pp. 294–315.
23. The Simple FEM Solver. [Electronic resource]. Access mode: <https://github.com/SeregaGomen/QFEM>
24. Timoshenko S., Woinowsky-Krieger S. *Theory of plates and Shells*. Singapore, McGraw-Hill, Inc., 1964, 594 p.

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

CONTROL IN TECHNICAL SYSTEMS

УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

UDC 004.023

GUIDED HYBRID GENETIC ALGORITHM FOR SOLVING GLOBAL OPTIMIZATION PROBLEMS

Avramenko S. E. – Ms. Sc., Department of System Analysis and Control, Dnipro University of Technology, Machine Learning Engineer, COMPARUS.UA, Dnipro, Ukraine.

Zheldak T. A. – PhD, Head of the Department of System Analysis and Control, Dnipro University of Technology, Dnipro, Ukraine.

Koriashkina L. S. – PhD, Associate Professor of the Department of System Analysis and Control, Dnipro University of Technology, Dnipro, Ukraine.

ABSTRACT

Context. One of the leading problems in the world of artificial intelligence is the optimization of complex systems, which is often represented as a nonlinear function that needs to be minimized. Such functions can be multimodal, non-differentiable, and even set as a black box. Building effective methods for solving global optimization problems raises great interest among scientists.

Objective. Development of a new hybrid genetic algorithm for solving global optimization problems, which is faster than existing analogues.

Methods. One of the crucial challenges for hybrid methods in solving nonlinear global optimization problems is the rational use of local search, as its application is accompanied by quite expensive computational costs. This paper proposes a new GBOHGA hybrid genetic algorithm that reproduces guided local search and combines two successful modifications of genetic algorithms. The first one is BOHGA that establishes a qualitative balance between local and global search. The second one is HGDN that prevents re-exploration of the previously explored areas of a search space. In addition, a modified bump-function and an adaptive scheme for determining one of its parameters – the radius of the “deflation” of the objective function in the vicinity of the already found local minimum – were presented to accelerate the algorithm.

Results. GBOHGA performance compared to other known stochastic search heuristics on a set of 33 test functions in 5 and 25-dimensional spaces. The results of computational experiments indicate the competitiveness of GBOHGA, especially in problems with multimodal functions and a large number of variables.

Conclusions. The new GBOHGA hybrid algorithm, developed on the basis of the integration of guided local search ideas and BOHGA and HGDN algorithms, allows to save significant computing resources and speed up the solution process of the global optimization problem. It should be used to solve global optimization problems that arise in engineering design, solving organizational and management problems, especially when the mathematical model of the problem is complex and multidimensional.

KEYWORDS: nonlinear optimization, global minimum, randomized search heuristics, hybrid approach, genetic algorithm, deflation operator, guided local search.

ABBREVIATIONS

GA is a genetic algorithm;

BFGS is a Broyden-Fletcher-Goldfarb-Shanno algorithm;

L-BFGS is a modification of the Broden-Fletcher-Goldfarb-Shanno algorithm (with limited memory);

L-BFGS-B is an extended modification of the Broden-Fletcher-Goldfarb-Shanno algorithm;

BOHGA is a best offspring hybrid genetic algorithm;

PSO is a particle swarm optimization algorithm;

HGDN is a hybrid genetic deflated Newton method;

GBOHGA is a guided best offspring hybrid genetic algorithm;

BH is basin-hopping, a technique for global optimization;

DA is a dual annealing algorithm;

DE is a differential evolution, a population-based metaheuristic search algorithm.

NOMENCLATURE

$f(\mathbf{x})$ is an objective function;

$\nabla f(\mathbf{x})$ is a gradient of a function $f(\mathbf{x})$;

$H(\mathbf{x})$ is a Hessian matrix $f(\mathbf{x})$;

\mathbf{x}^* is a global optimum of a function $f(\mathbf{x})$;
 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is a point of n -dimensional real space, current position;
 γ is an improvement from the current positions to the next;
 $\mathbf{x}_0^k, k = \overline{1, K}$ are deflation points of the function, local optimums found at the current moment of the algorithm;
 r is a radius around the deflation point;
 r_k is an adjustable radius around the deflation point;
 $b_{\mathbf{x}_0}(\mathbf{x})$ is a bump-function (a smooth function with compact support);
 $\lambda \in (1, 2]$ is a parameter of the GBOHGA;
 α is a coefficient for adjusting the shape of the bump function.

INTRODUCTION

Today in the world of artificial intelligence one of the leading problems is the optimization of complex systems, which are reduced to many socio-economic, technical, organizational and managerial tasks [1, 2]. It is often presented as a nonlinear function that needs to be optimized. Problems of nonlinear optimization also arise in engineering design and in the fundamental sciences, such as in physics, chemistry, molecular biology, and others [3, 4]. Numerical solutions to such problems can be associated with significant difficulties because the objective function can be multi-extreme, non-differentiable and, in general, given in the form of a black box. In addition, each calculation of a function may require significant computational resources.

Unfortunately, there is no universal method for determining the global solution of the nonlinear optimization problem in the general formulation. Most traditional optimization methods are deterministic and local. They are often unable to leave the areas of attraction of local optimums. The use of the found local solutions may be insufficient because the global optimum can give a significant advantage over local ones. But there is an alternative approach – the use of evolutionary methods of global optimization and the deliberate introduction of an element of randomness into the search algorithm. These methods are based on the natural behavior of biological and physical systems and are able to overcome the shortcomings of traditional methods of “stuck” in the local optimum [5]. In addition, randomness can serve the purpose of collecting information about the behavior of the object of study and management objectives. The advantages of such methods are considered to be their increased speed; low sensitivity to irregular behavior of the objective function and the presence of random errors in calculating the function; relatively simple internal implementation; low sensitivity to the growing dimensionality of the optimization problem; the possibility of natural introduction into the search process of learning and self-learning. Besides, within the known random search schemes, new algorithms are easily

built that implement various heuristic adaptation procedures.

A popular representative of evolutionary strategies is a genetic algorithm. In its classical form, it successfully explores the search space but has a slow convergence. To get rid of this shortcoming developed and widely used in the practice implementations that combine genetic algorithm and methods that in the process of their work use local information about promising areas of search. Such methods are called hybrid. One of the important issues in them is the balance between their components, and one of the well-known problems is the repeated exploration of the same area with the repeated finding of the already known local solution. Therefore, the development and research of new effective methods and algorithms for finding the global optimum of a multi extreme function continues.

This paper presents a new hybrid genetic algorithm GBOHGA for solving the global optimization problems, which integrates methods that can solve the above problems. One way to balance the components of the algorithm is to manage the proportion of the population that should perform a local search. For this purpose, the scheme laid down in the BOHGA method was used [6]. To prevent re-exploration of previously visited areas, the idea of “deflation” of the function from HGDN [7] was borrowed to “remove” already found solutions, but it is proposed to use it to the objective function, which is ideologically similar to guided local search [8] from combinatorial optimization. In addition, a new adaptive scheme is proposed to establish the size of the radius of the “deflation” area in order to avoid the unsuccessful selection of this parameter. As will be shown below, all this provides significant savings in computing resources and speeds up the algorithm, as it prevents the unnecessary use of local search.

The object of study is the problems of global optimization.

The subject of study is local search methods and randomized search heuristics for solving global optimization problems, their computational characteristics, and conditions of application.

The purpose of the work is to develop a new hybrid genetic algorithm for solving the problem of global optimization, devoid of the shortcomings of existing analogs.

1 PROBLEM STATEMENT

We will consider the problem of optimization in the next statement:

$$f^* = \min_{x \in \Omega} f(x), \quad (1)$$

where the feasible set is n -dimensional parallelepiped

$$\Omega = \{x \in R^n, a_i \leq x_i \leq b_i, i = \overline{1, n}\}. \quad (2)$$

Since the aim of the research is to develop an algorithm for solving the problem (1), (2) for the widest pos-

sible class of objective functions, we will assume that the function $f(x)$ is continuous, but can be multimodal, nonlinear, non-differentiable, have a ravine relief or high computational complexity.

2 REVIEW OF THE LITERATURE

The variety of global optimization problems entails a variety of approaches to their solution [1, 9]. An extensive overview of such problems and methods is given in [10]. Many studies are devoted to the study of the peculiarities of the use of methods for specific problems of global optimization. Let us recall only some of the known approaches, and review the latest developments in this field, without claiming to be complete.

There are so-called passive (non-adaptive) algorithms, in which the calculation points of the objective function are selected before the search from some a priori ideas about the problem and can not be changed during the operation of the algorithm. This approach usually requires a very large amount of computation to obtain a reliable estimate of the global solution. When using sequential (adaptive) schemes for placing the calculation points of the function, both a priori and information about the problem obtained during the search process are taken into account. Adaptive methods are generally characterized by a denser placement of the points of calculation of the function in the vicinity of the solution and rarer outside of it. Consistent are, for example, branch-and-bound algorithm, Bayesian methods, algorithms based on interval arithmetic, methods of multiple local descent, and many others [10]. Within the framework of interval analysis, the technique of multiple partitioning is intensively developed, as well as more complex partitioning strategies based on simplexes [11] and various auxiliary functions [12]. Each of the techniques is common enough to solve different classes of global optimization problems.

For many practical problems (solving systems of nonlinear equations and inequalities, optimization of hierarchical models associated with placement problems, service systems, etc.), the assumption that the functions characterizing the system of study are Lipschitz functions is typical. The development of the theory and methods of numerical solution of problems of this type is the subject of Lipschitz global optimization [13, 14].

Special mention should be made of adaptive stochastic search strategies, many of which underlie heuristic global optimization algorithms (simulated annealing methods, genetic algorithms, tabu search methods, and others). Different strategies for approaching and relaxing the original problem are also used, which involve the consistent construction of auxiliary subtasks: their solution allows to improve the estimation of the global optimum [15–17]. In [15, 16] methods and algorithms for solving optimization problems are studied, which do not use information about the derivatives of the objective function but require only knowledge of its values. In [15] a systematic comparison of existing implementations of such algorithms on a large set of test problems, convex and nonconvex, smooth and non-smooth is presented. The authors single out global solvers that are superior to local ones, even for convex problems, and demonstrate excellent performance in terms of improving suboptimal solutions.

Geometric and information bases for the construction of global optimization algorithms are considered in [16].

Here are some new ideas for speeding up the search so that accelerated optimization methods automatically implement local behavior in promising regions and do not stop the search process of a global minimum. The authors emphasize that a reasonable combination of new and traditional computational steps leads to two dozen different global optimization algorithms. All of them are studied and numerically compared on three test sets of functions and applied problems.

For costly, in terms of computing resources, black-box problems surrogate models are now widely used to reduce computation time and effort in finding the global optimum. Thus, in [17] a new global optimization method is proposed, which uses an ensemble of surrogates and reduction of hierarchical design. The search uses an ensemble of three representative surrogate methods with optimized weights to select promising sampling points, narrow the study space, and determine the global optimum. The new proposed global optimization method tested using eighteen representative functions and two engineering design optimization problems, demonstrated improved capabilities in identifying promising search areas and reducing design space, as well as excellent search efficiency and reliability in determining global optimum. Fundamental problems that arise during global optimization based on surrogates from the point of view of the practitioner, including the selection of concepts, methods, techniques, and engineering applications, are discussed in [18]. Providing a comprehensive discussion of the issues raised, the paper explores the latest achievements in the planning of experiments, methods of surrogate modeling, criteria for filling and reducing the project space.

New sampling points are needed to increase the accuracy of the surrogate model forecast. The decision-making strategies for the following promising samples are called the filling criteria, the so-called adaptive sampling methods. The filling criteria (which are conventionally classified as exploitation, exploration, combined exploitation, and research) may well guide the selection of new sampling points depending on the information from the optimization process to be used. On the one hand, operating methods may focus on regions located in the area of the best point that has been found so far, but which may not even be stationary for the function. This can only lead to a local approximation and a local optimum. On the other hand, exploration and combined exploitation usually explore rare regions or regions with high uncertainty. However, only the use of exploration strategy can lead to the depletion of computational resources when approaching the global minimum with a given accuracy. In general, high accuracy is needed in potentially promising regions. Therefore, the authors [19] propose to combine exploitation and exploration to balance competing goals between lower cost and more accurate assessment of the optimal solution. The work focuses on parallel algorithms of effective global optimization. Here, a multi-point fill criterion is developed, which uses entropy to accurately measure the uncertainty of the Kriging surrogate, and then balances global exploration and the local exploitation of multi-point selection criteria. A strategy for optimizing the decomposition of domains, which provides a small amount of training

data, is also proposed. According to the authors, compared to several other methods, it has an obvious advantage in solving complex optimization problems in a large-scale parallel computing environment.

Bayesian optimization, which is used to search for the global minimum of computationally expensive black-box functions when the evaluation of functions is costly and the optimum can be achieved with a relatively small number of evaluations, is becoming increasingly popular today. However, its application for multidimensional problems with several thousand observations remains a difficult task, and the improvement of appropriate methods and approaches is an important area of modern research. Thus, [20] proposed ensemble Bayesian optimization to simultaneously solve the problems of large-scale observations, large dimensionality of the input space, and the choice of batch requests that balance quality and diversity. In [21] the authors proposed the TuRBO algorithm based on a collection of local probability models that provide local search trajectories that can be used to quickly identify the best values of the objective function. This approach is complemented by a so-called gangster strategy, which distributes samples across trust regions, implicitly trading exploration and exploitation.

Therefore, deterministic methods have an advantage in the guaranteed quality of the solution found, but the computational complexity of such methods is exponential in the worst case. In fact, no method is better for the worst case of a function than passive grid search. Because optimization problems usually come from an arbitrary source and therefore tend to have unknown statistical characteristics, alternative approaches to solving such problems, including adaptive search strategies and statistical modeling tools, are becoming important tools, especially in the case of high-dimensional problems.

Significant results show evolutionary strategies that mimic biological and social models of evolution. Different deterministic and stochastic algorithms can be built on the basis of different evolutionary rules. This is evidenced, for example, by the successful application of the genetic algorithm to real problems in a wide range of industries [22]. Although metaheuristics are considered reliable, their convergence slows down when solving complex problems, because they do not involve the use of local information about promising areas of search. To combat this defect, the idea of combining local search with various metaheuristic algorithms is being increasingly introduced.

The following is a brief description of the genetic algorithm, together with its modifications and methods of constructing hybrids based on it.

3 MATERIALS AND METHODS

The genetic algorithm is an iterative optimization method based on the concept of natural selection known in genetics. Variables that characterize the solution are represented as a set of genes on a chromosome. For example, a point in n -dimensional space can be encoded by a characteristic vector $x = (x_1, \dots, x_n)$. GA operates with a finite set of solutions (population), generates new solutions as different combinations of parts of others, using operators such as selection, crossover and mutation [22]. The new solutions are posi-

tioned in the population according to their position on the surface of the studied function.

At the beginning of the algorithm, initialization is performed – the formation of the initial population randomly. Next, the selection of chromosomes is done, which consists of selecting those that will produce offspring for the next population. Genetic operators that form a new population are applied to the selected chromosomes. Crossover is a key operator based on the assumed high probability that the new solution obtained from two sufficiently good solutions of the problem will be good or even better than the previous ones. An important but still secondary role is played by the mutation operator, which introduces new genetic material into the population to maintain population diversity and prevents losses that could occur due to the exclusion of any significant gene as a result of crossbreeding.

The steps described above are repeated many times until the algorithm's stop criterion is met, for example, when the already achieved value of the objective function does not improve for some time. Or the operation of the algorithm can be strictly limited by a predetermined duration or number of iterations.

In order not to waste significant computing resources when solving complex problems, after crossover and mutation, an additional operator is used. It performs a local search, starting in a certain area of the promising area. Such algorithms are called hybrid. Methods and algorithms of local search most often find the nearest extremum, and the trajectory of their movement significantly depends on the choice of the starting point and the nature of the target function.

The choice of local search method for inclusion in the hybrid should be made taking into account the benefits and costs of using each of the approaches. Genetic algorithm operators also partially perform the role of local search with relatively low computational cost compared to more accurate local search methods. Therefore, it is rational to limit the usage of a local search operator. A specific scheme for determining individuals from the population to whom it is desirable to apply local search is offered by the BOHGA [6]. The algorithm's scheme proposes to perform a local search only when the best representative of the offspring population is also the best in the current parent population. When such best offspring appear, it is very likely that such offspring will explore a new area, and therefore local knowledge will be used to accelerate the search for the most promising region.

However, in the described scheme, there can still be a re-application of local search to individuals that fall into the same area of gravity of the search space, so in order to prevent unnecessary computational costs in [7], a deflated Newton modification was proposed that is a key idea of the HGDN method. This hybrid uses a genetic algorithm as a global search method and deflated Newton scheme as a local search operator. This modification of Newton's method effectively identifies several local optimums in the immediate vicinity of the starting point and accordingly changes the function. Here is the main idea of the method.

It is known that Newton's method calculates the gradient and Hessian of the objective function at the current point and uses this information to predict the new location of the solution by solving the equation $H(\mathbf{x})\gamma = -\nabla f(\mathbf{x})$. The deflated Newton scheme is that Newton's algorithm looks for station-

ary points of function $f(\mathbf{x})$. Let $\mathbf{x}_0^1, \mathbf{x}_0^2, \dots, \mathbf{x}_0^K$ – be the stationary points defined at the moment. Then other such points can be found by applying the deflation of the gradient of the function f by the formula:

$$\nabla f_{\mathbf{x}_0}(\mathbf{x}) = \frac{\nabla f(\mathbf{x})}{\prod_{k=1}^K \|\mathbf{x} - \mathbf{x}_0^k\|^2}. \quad (3)$$

The deflated function has no roots in the found optimums, and therefore, the local search method will no longer be able to converge to these points. The principle of operation of this approach is shown in Fig. 1 [7]. Thus, in Fig. 1, a the genetic algorithm applies the local search operator to each individual of the population, and in the next iteration, the same optimum is identified again. To combat this behavior, formula (3) is used to calculate the gradient of the function, i.e. the roots are removed (Fig. 1,b), and therefore the already found optimums cannot be identified again, and each individual can find new optimums (Fig. 1, c).

Since the proposed deflation formula is often unreliable in the computational sense, alternative deflation operators are being developed [7, 23].

The localized operator proposed in [7], in order to affect only the area close to the deflation, uses the bump-function (smooth function with compact support) in the following form:

$$b_{\mathbf{x}_0}(x) = \begin{cases} \prod_{i=1}^n \frac{\exp\left(\frac{-\alpha}{r^2 - (x_i - x_{0i})^2}\right)}{\exp\left(\frac{-\alpha}{r^2}\right)}, & \text{if } x_{0i} - r < x_i < x_{0i} + r, \\ 0 & \text{else,} \end{cases}$$

with \mathbf{x}_0 as a deflation point. And the deflated function itself has the form:

$$\nabla f_{\mathbf{x}_0}(\mathbf{x}) = \frac{\nabla f(\mathbf{x})}{1 - b_{\mathbf{x}_0}(\mathbf{x})}.$$

We will build a new global optimization algorithm that combines the above-described BOHGA and HGDN hybrids, as both seek to reduce the number of function evaluations by

eliminating the useless use of the local search operator to individuals who are either in non-promising regions or in the same one.

In addition, we develop the idea of a guided local search algorithm [8], which has successfully proven itself in combinatorial optimization [24–26]. The latter was proposed by Voudouris and Tsang and is a meta-heuristic method that during the search process constructs penalties and uses them to help local search algorithms get out of the minimum already found, and even from almost horizontal areas.

As soon as the search algorithm falls within the local minimum, the objective function is modified in a certain way, and the local search works with this modified function.

In the new algorithm GBOHGA to modify the objective function in the areas of attraction of already found local extrema, we will use the bump-function introduced in HGDN. That is, the function $f(\mathbf{x})$ in these areas is replaced by the following:

$$f_{\mathbf{x}_0}(\mathbf{x}) = \frac{f(\mathbf{x})}{1 - b_{\mathbf{x}_0}(\mathbf{x})}. \quad (5)$$

Thus, in this way, the algorithm forces the individual to look for a new optimum elsewhere, preventing re-convergence to the point \mathbf{x}_0 . However, with the long-term operation of an algorithm with a highly multimodal function, the set of local optimums found can become very large, which affects both the amount of memory consumed and computational costs. Besides, in practice, there is a question of choosing a value for the parameter r that is responsible for the size of the radius of the deflation point of the previously found optimum. If you choose too small a value, its positive effect on the algorithm will not be noticeable, in addition, it can lead to a significant number of local optimums in the same area of attraction of the search space. On the contrary, choosing too large a value of this parameter, we assume that the local search algorithm qualitatively explored this region and provided the best possible solution from it.

Although, of course, it may happen that this relatively small area itself is very complex, and, therefore, such a choice of the radius of deflation can cause premature convergence of the algorithm.

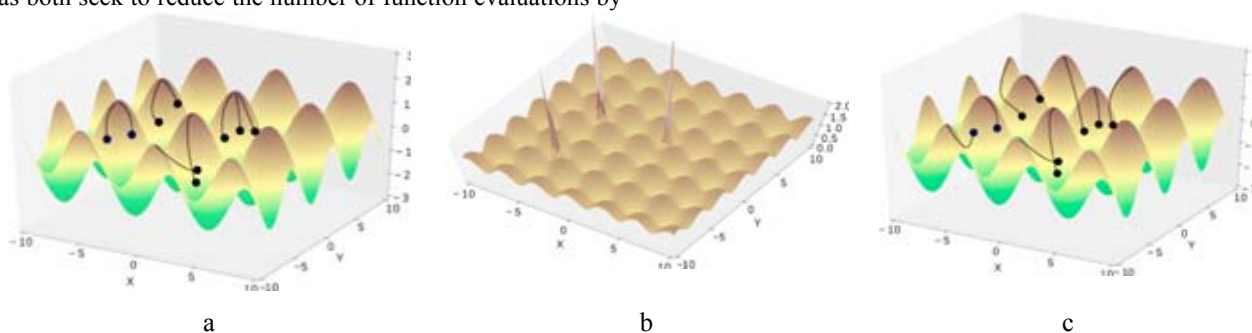


Figure 1 – Demonstration of deflated Newton’s method [7]: a – the hybrid method applies local optimization to each individual. In the next iteration, the same optima can be identified again; b – the gradient of the function is deflated at the optima of the function; c – HGDN removes any identified optima after each iteration. Therefore, already found optima cannot be identified again

To overcome this uncertainty, the new algorithm proposes to use an adaptive scheme for selecting the value of the parameter r , that responds to the frequency of falling into the area of attraction of the local optimum. Given the type of objective function, for each found point x_0 of the local optimum, a certain rather small initial value r_0 is chosen. Next, in the case of repeated convergence to the previously found optimum at a distance less than or equal to λr_i , from two points of the optimum – new x_0^k and x_0^i save the one that corresponds to the smaller value of the function, with radius $r = \lambda r_i$. The bump-function has the form:

$$b_{x_0^k}(x) = \begin{cases} \prod_{i=1}^n \frac{\exp\left(\frac{-\alpha}{r_k^2 - (x_i - x_{0i}^k)^2}\right)}{\exp\left(\frac{-\alpha}{r_k^2}\right)}, & \text{if } \|x - x_0^k\| < r_k, \\ 0 & \text{else.} \end{cases} \quad (6)$$

In order not to perform a costly operation of local search in those areas that are likely to be of no use, we will use the BOHGA scheme of a selection of individuals to which the deflation operator (5), (6) is applied with an adaptive choice of parameter r . Thus, the scheme of the GBOHGA is as follows.

GBOHGA algorithm

1. The operators of the genetic algorithm are determined; parameters r_0, α, λ are set.
2. The initial population is randomly generated.
3. Evaluation of the objective function for each individual of the initial population is performed.
4. A population of offspring is created using genetic algorithm operators.
5. The objective function of each individual of the population of offspring according to formulas (5), (6) is estimated.
6. It is determined whether the best offspring individual among the offspring population is also the best among the current parent population.
 - 6.1 If not, go to step 7 (local search is not performed).
 - 6.2 In the case of such a descendant, the local search operation is applied only to it and the suitability of the new solution x_0^k is estimated.
 - 6.3 It is checked whether the new solution is similar to any previously found. That is, whether the new solution x_0^k is located at a distance less than λr_i , from the i -th previously found local optimum.

6.3.1 If not, then a new solution x_0^k with a radius r_k is added to the set of local optimums $\{(x_0^1, r_1), (x_0^2, r_2), \dots, (x_0^{k-1}, r_{k-1})\}$.

6.3.2 Otherwise, from the two points of the local optimum – the new x_0^k and x_0^i – we keep the one that corresponds to the smaller value of the function, with a radius $r = \lambda r_i$.

7. It is decided which individuals to include in the next population and which parents should also be left in the new generation according to the elitist approach.

8. If the stop criterion is met, the procedure stops, otherwise, the transition to step 4 is performed.

4 EXPERIMENTS

To test and compare the efficiency of the developed and existing, the best for today, optimization algorithms among the known modern artificial landscapes [27] selected 33 functions with different properties in terms of modality, separability, scalability, and landscape of the valley. The choice of functions for testing is justified by the following considerations.

Multimodal functions with many local minima form one of the most complex classes of optimization problems for many algorithms and are used to test the ability of the algorithm to proceed from any local minimum. If the algorithm has a poorly designed search process, it cannot effectively investigate the landscape of functions and is stuck at the local minimum. Also for algorithms, functions with flat surfaces can cause difficulties, because their invariance does not give the algorithm any information to direct the search process to the minima (for example, Stepint or Powell Sum).

Among the test functions, there are separable and non-separable. Usually, it is much easier to optimize the former, since each variable of such a function is independent of the others, it is possible to perform a sequence of n independent optimization processes to find the optimal value of the respective variables.

An important indicator of the effectiveness of optimization algorithms is the ability to overcome the problem of dimensionality because it is known that with the increasing number of parameters, the search space also increases, and exponentially. Therefore, the set of test functions includes functions with the ability to increase the number of their components.

For some functions, the area containing global minima is very small compared to the whole search space, for example, Powell or Schaffer problems. For the last problem, the global minimum is very close to the local minima. And if the algorithm does not have time to change direction in functions with a narrow curved valley, or inefficiently explores the search space, in the case where the functions have several global minima, such an algorithm will not be able to cope with these problems. The choice of test functions is made taking into account the variety of the functions' hypersurfaces [28].

Table 1 contains a general description of a set of selected artificial landscapes listed above. The list of all test functions used in the work with their mathematical notation is given in the Table 2, and in Fig. 2 are presented the landscapes of some of them.

Table 1 – General characteristics of a set of artificial landscapes

A property that is typical for a function from the set of artificial landscapes	Percentage of functions in the set
continuity	91
differentiability	73
separability	36
unimodality	58
convexity	42

Table 2 – Artificial landscapes

N	Name, mathematical notation, optimal solution	N	Name, mathematical notation, optimal solution
1	Ackley 1 $f(x) = -20e^{-0.02\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2}} - e^{\frac{1}{n}\sum_{i=1}^n \cos(2\pi x_i)} + 20 + e$ $-35 \leq x_i \leq 35, x^* = (0,0,\dots,0), f(x^*) = 0$	11	Schwefel 2.21 $f(x) = \max_{1 \leq i \leq n} x_i , -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
2	Alpine 1 $f(x) = \sum_{i=1}^n x_i \sin(x_i) + 0.1x_i , -10 \leq x_i \leq 10, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$	12	Schwefel 2.22 $f(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i , -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
3	Chung Reynolds $f(x) = \left(\sum_{i=1}^n x_i^2\right)^2, -100 \leq x_i \leq 100,$ $x^* = (0,0,\dots,0), f(x^*) = 0$	13	Schwefel 2.23 $f(x) = \sum_{i=1}^n x_i^{10}, -10 \leq x_i \leq 10, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
4	Exponential $f(x) = -\exp\left(-0.5\sum_{i=1}^n x_i^2\right), -1 \leq x_i \leq 1, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 1$	14	Sphere $f(x) = \sum_{i=1}^n x_i^2, 0 \leq x_i \leq 10, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
5	Griewank $f(x) = \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1, -100 \leq x_i \leq 100,$ $x^* = (0,0,\dots,0), f(x^*) = 0$	15	Step $f(x) = \sum_{i=1}^n (\lfloor x_i \rfloor), -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
6	Hapy Cat $f(x) = \left[\left(\ x\ ^2 - n \right)^2 \right]^\alpha + \frac{1}{n} \left(\frac{1}{2} \ x\ ^2 + \sum_{i=1}^n x_i \right) + \frac{1}{2},$ $-2 \leq x_i \leq 2, \alpha - \text{real}; x^* = (-1,-1,\dots,-1), f(x^*) = 0$	16	Step 2 $f(x) = \sum_{i=1}^n (\lfloor x_i + 0.5 \rfloor)^2, -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0.5,0.5,\dots,0.5), f(x^*) = 0$
7	Periodical $f(x) = 1 + \sum_{i=1}^n \sin^2(x_i) - 0.1e^{\left(\sum_{i=1}^n x_i^2\right)}, -10 \leq x_i \leq 10,$ $x^* = (0,0,\dots,0), f(x^*) = 0.9$	17	Step 3 $f(x) = \sum_{i=1}^n (\lfloor x_i^2 \rfloor), -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
8	Powell Sum $f(x) = \sum_{i=1}^n x_i ^{i+1}, -1 \leq x_i \leq 1, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$	18	Stretched V Sine Wave $f(x) = \sum_{i=1}^{n-1} (x_{i+1} + x_i^2)^4 \left[\sin^2 \left\{ 50(x_{i+1} + x_i^2)^{0.1} \right\} + 0.1 \right],$ $-10 \leq x_i \leq 10, i = \overline{1, n}, x^* = (0,0,\dots,0), f(x^*) = 0$
9	Rastrigin $f(x) = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i)),$ $-5.12 \leq x_i \leq 5.12, i = \overline{1, n}, x^* = (0,\dots,0), f(x^*) = 0$	19	Sum Squares $f(x) = \sum_{i=1}^n ix_i^2, -10 \leq x_i \leq 10, i = \overline{1, n},$ $x^* = (0,0,\dots,0), f(x^*) = 0$
10	Rosenbrock $f(x) = \sum_{i=1}^{n-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \right],$ $-30 \leq x_i \leq 30, i = \overline{1, n}, x^* = (1,1,\dots,1), f(x^*) = 0$	20	Trigonometric 1 $f(x) = \sum_{i=1}^n \left[n - \sum_{j=1}^n \cos x_j + i(\cos(x_i) - \sin(x_i)) \right]^2$ $0 \leq x_i \leq \pi, i = \overline{1, n}, x^* = (0,0,\dots,0), f(x^*) = 0$

Continuation of the Table 2

21	Salomon $f(x) = 1 - \cos\left(2\pi\sqrt{\sum_{i=1}^n x_i^2}\right) + 0.1\sqrt{\sum_{i=1}^n x_i^2},$ $-100 \leq x_i \leq 100, i = \overline{1, n}, x^* = (0, \dots, 0), f(x^*) = 0$	28	Trigonometric 2 $f(x) = 1 + \sum_{i=1}^n 8 \sin^2(7(x_i - 0.9)^2) + 6 \sin^2[14(x_1 - 0.9)^2] +$ $+ (x_i - 0.9)^2, -500 \leq x_i \leq 500, i = \overline{1, n}$ $x^* = (0.9, 0.9, \dots, 0.9), f(x^*) = 1$
22	Sargan $f(x) = \sum_{i=1}^n n \left(x_i^2 + 0.4 \sum_{j \neq i} x_i x_j \right),$ $-100 \leq x_i \leq 100, i = \overline{1, n}, x^* = (0, \dots, 0), f(x^*) = 0$	29	W / Wavy Function $f(x) = 1 - \frac{1}{n} \sum_{i=1}^n \cos(kx_i) e^{-x_i^2/2}, k - \text{дійсне число}$ $-\pi \leq x_i \leq \pi, i = \overline{1, n}, x^* = (0, 0, \dots, 0), f(x^*) = 0$
23	Schaffer F6 $f(x) = \sum_{i=1}^n \left(0.5 + \frac{\sin^2 \sqrt{x_i^2 + x_{i+1}^2} - 0.5}{\left[1 + 0.001(x_i^2 + x_{i+1}^2)\right]^2} \right)$ $-100 \leq x_i \leq 100, i = \overline{1, n}, x^* = (0, \dots, 0), f(x^*) = 0$	30	Weierstrass $f(x) = \sum_{i=1}^n \left[\sum_{k=0}^{k_{\max}} a^k \cos(2\pi b^k (x_i + 0.5)) - n \sum_{k=0}^{k_{\max}} a^k \cos(\pi b^k) \right]$ $-0.5 \leq x_i \leq 0.5, i = \overline{1, n}, x^* = (0, 0, \dots, 0), f(x^*) = 0$
24	Schumer Steiglitz $f(x) = \sum_{i=1}^n x_i^4, x^* = (0, 0, \dots, 0), f(x^*) = 0$	31	Whitley $f(x) = \sum_{i=1}^n \sum_{j=1}^n \left[\frac{(100(x_i^2 - x_j)^2 + (1 - x_j)^2)^2}{4000} - \right.$ $\left. - \cos(100(x_i^2 - x_j)^2 + (1 - x_j)^2 + 1) \right]$ $x^* = (1, 1, \dots, 1), f(x^*) = 0$
25	Schwefel 2.20 $f(x) = -\sum_{i=1}^n x_i , -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0, 0, \dots, 0), f(x^*) = 0$		
26	Schwefel $f(x) = \left(\sum_{i=1}^n x_i^2 \right)^\alpha, \alpha \geq 0 - \text{дійсний параметр}$ $-100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0, 0, \dots, 0), f(x^*) = 0$	32	Xin-She Yang 3 $f(x) = e^{-\sum_{i=1}^n (x_i/\beta)^{2m}} - 2e^{-\sum_{i=1}^n x_i^2} \cdot \prod_{i=1}^n \cos^2(x_i)$ $-20 \leq x_i \leq 20, m = 5, \beta = 15, x_i^* = 0 \forall i, f(x^*) = -1$
27	Schwefel 1.2 $f(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j^2 \right)^2, -100 \leq x_i \leq 100, i = \overline{1, n},$ $x^* = (0, 0, \dots, 0), f(x^*) = 0$	33	Xin-She Yang (Function 4) $f(x) = \left(\sum_{i=1}^n \sin^2(x_i) - e^{-\sum_{i=1}^n x_i^2} \right) e^{-\sum_{i=1}^n \sin^2 \sqrt{ x_i }},$ $-10 \leq x_i \leq 10, i = \overline{1, n}, x^* = (0, 0, \dots, 0), f(x^*) = -1$

For all the problems consider an objective function subject to bound constraints on the values of the variables. All functions are scalable, which allows us to check the behavior of algorithms in large-scale problems. Two-thirds of all test functions are not separable, which is also a challenge for most algorithms, but perhaps the most difficult is multimodality, especially in multidimensional spaces.

The testing was performed in comparison with other popular evolutionary algorithms, including the BH, DA, PSO, DE. For correct analysis, instead of independent implementation of these methods, their open-source versions from the SciPy library were used.

A key aspect of developing a new algorithm is the desire to reduce the required number of calculations of the

objective function with an adequate level of reliability of the algorithm. Instead of counting the number of calls to the objective function, we can also use the physical time spent on solving problems, but this characteristic strongly depends on the quality of the implementation of the algorithm and the level of hardware, and therefore it is less universal.

In order to obtain reliable comparison results, the experiment was performed 50 times with random initial populations, and the results were averaged. Fig. 3 showed the performance profiles for BH, DA, DE, PSO, and GBOHGA on the test data set using function evaluations as performance criteria.

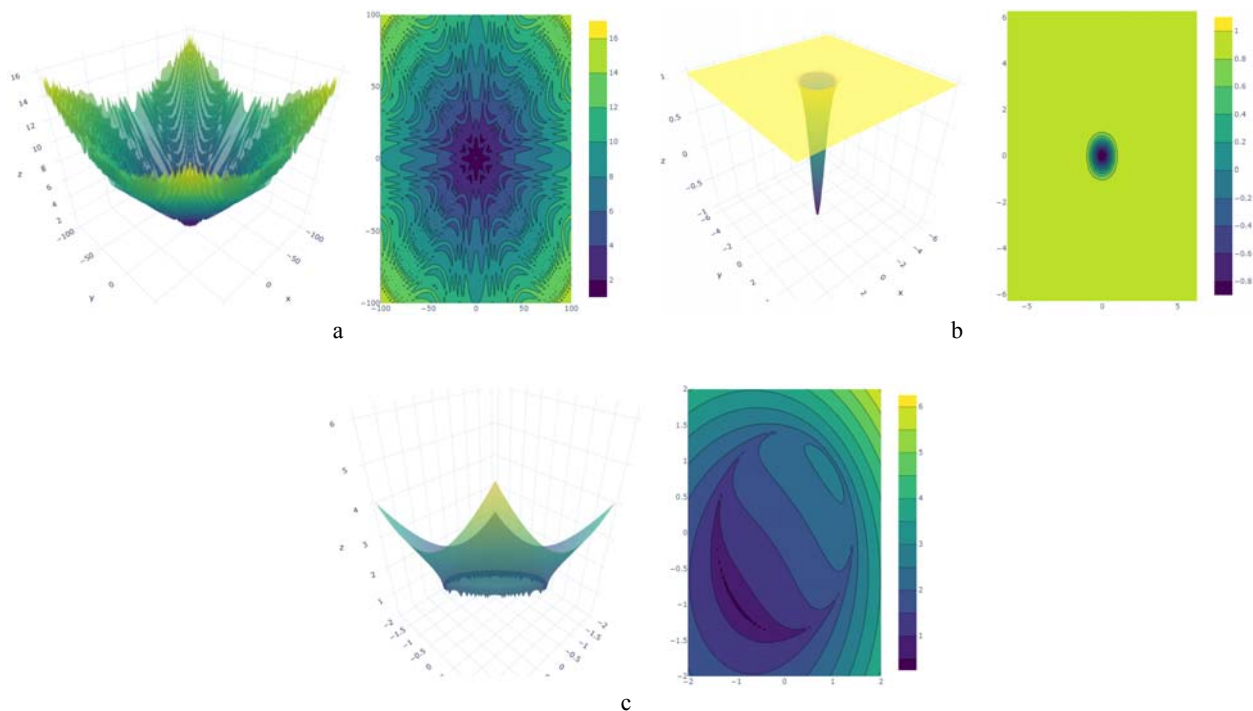


Figure 2 – 3d-Surface and Contour Map of some test functions:
 a – Salomon Function; b – Xin-She Yang Function; c – Hapy Cat Function

5 RESULTS

Fig. 3 shows that for the test set of problems GBOHGA is the best algorithm in terms of performance, followed by DA and DE. The GBOHGA algorithm requires 3 times fewer evaluations of the objective function than DA and 6 times less than DE. At the same time, these algorithms tend to outperform GBOHGA, solving more problems by 5% and 6%, respectively. A detailed description of the results obtained in terms of each function is given separately in the Table 3. For each of the 5 optimization algorithms in the first column (μ) is given the average value of the number of calls to the objective function for each of the 33 test functions based on the results of 50 runs of the algorithm. The second column (σ), according to the same scheme, indicates the standard deviation, and the third (%) indicates the success level (values from 0.0 to 1.0) of the algorithm on this problem, which characterizes the robustness of the algorithm. The results of this table are also summarized for all functions and given in the last line. For greater clarity, in Fig. 4 a comparison of algorithms by their performance is demonstrated. Thus, GBOHGA is in 62.5% of cases faster than competing algorithms.

When solving optimization problems with 25 variables, the GBOHGA algorithm is still the best in terms of performance, and it can only compete with DA, which solves 3% more problems from the test suite, but con-

sumes 5 times more computing resources (see Table 4 for more details). In terms of performance speed, GBOHGA shows even better results than in the previous experiment, namely in 69.7% of cases GBOHGA is faster than competing methods.

Obviously, with the increase in the number of variables, GBOHGA shows not worse, but even better results compared to other methods.

It should be noted that when designing the algorithm, separately, ten different crossover operators were compared, of which the best results were found by a discrete crossover. Also, five mutation operators and two selection operators were tested and it was shown that the best combination is the rank selection with the non-uniform mutation.

Further, in combination with the genetic algorithm, seven different methods of local optimization were used. According to the obtained results, the advantage is attributed to quasi-Newton methods (L-BFGS-B turned out to be the best option).

Therefore, the above-mentioned operators were chosen as the key components of the proposed hybrid algorithm, other parameters of the genetic algorithm were chosen arbitrarily, just as all competing algorithms were performed without prior optimization of parameters, ie the values were chosen either arbitrarily or most used.

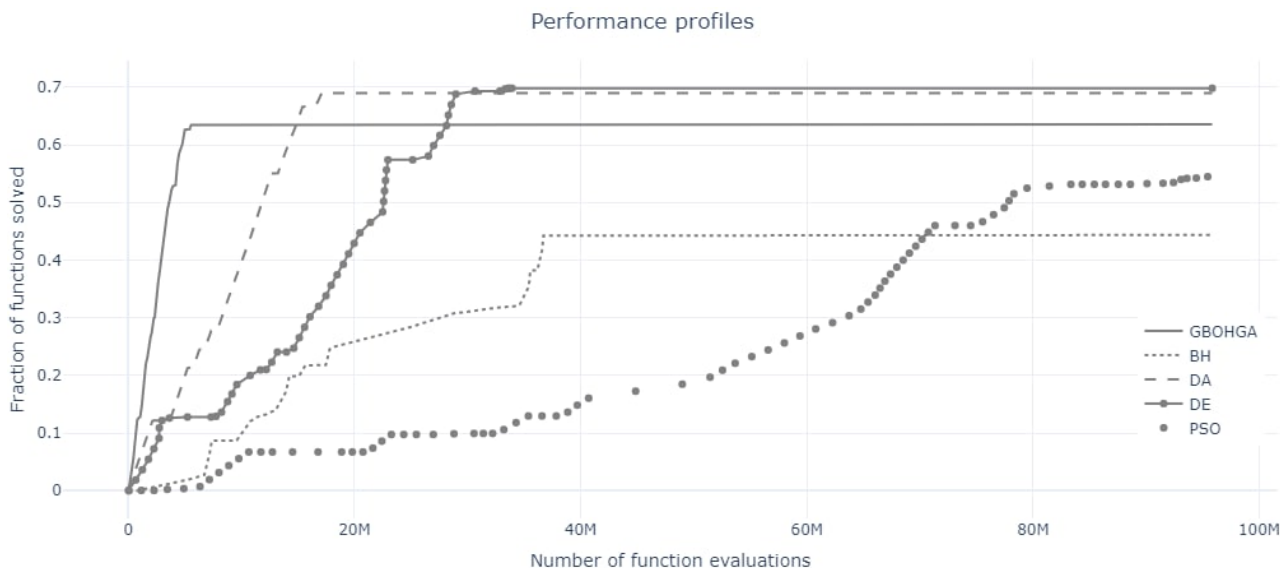


Figure 3 – Algorithms performance profiles

Table 3 – Mean (μ) and standard deviation (σ) of the number of function evaluations and success rate (%) on the results of 50 runs for each function from the test set for each method

Function name	BH			DA			DE			GBOHGA			PSO		
	μ	σ	%	μ	σ	%	μ	σ	%	μ	σ	%	μ	σ	%
Ackley1	14692	560	0.0	10701	131	1.0	20829	3097	0.96	4022	1050	0.76	53297	3732	0.0
Alpine1	119324	7602	0.88	11347	538	1.0	17643	1118	1.0	4602	627	1.0	72846	14714	0.16
Chung Reynolds	8712	154	1.0	10356	43	1.0	15525	554	1.0	2444	224	1.0	42601	2534	1.0
Exponential	3572	44	1.0	10032	5	1.0	968	97	1.0	2688	1931	1.0	43793	2147	1.0
Griewank	10105	305	0.0	10376	96	0.04	20303	5524	0.2	3184	884	0.36	54546	10131	0.0
HappyCat	27875	5363	0.0	10857	311	0.0	69822	8046	0.0	4054	1407	0.0	108498	6777	0.0
Periodic	5659	52	0.0	10038	6	0.0	14710	3839	0.04	2628	485	0.34	47473	2551	0.0
PowellSum	20654	470	1.0	10255	39	1.0	17677	536	1.0	2258	198	1.0	40994	3348	1.0
Rastrigin	15567	1078	0.48	10366	82	1.0	12508	2928	0.8	3587	692	1.0	54366	3089	0.0
Rosenbrock	31886	13068	0.34	10480	101	1.0	42053	9270	0.94	2838	424	1.0	92979	5267	0.02
Salomon	19196	5393	1.0	10453	124	0.02	15849	1596	0.0	3140	797	0.32	41447	2331	0.0
Sargan	5864	111	1.0	10057	7	1.0	16165	667	1.0	2394	274	1.0	53666	3147	1.0
SchafferF6	18194	3525	0.0	10860	255	0.06	26292	5890	0.0	3764	797	0.24	62897	14239	0.0
Schumer Steiglitz	12213	5943	0.76	10348	29	1.0	16150	625	1.0	2468	231	1.0	42139	2148	1.0
Schwefel	36485	1342	0.0	10811	233	0.0	15537	538	1.0	2841	287	0.0	204860	9943	1.0
Schwefel12	7494	52	1.0	10088	14	1.0	25597	837	1.0	2375	261	1.0	54120	3811	1.0
Schwefel220	125013	1957	1.0	11077	409	1.0	15243	548	1.0	3481	414	1.0	73152	5418	0.98
Schwefel221	93100	6074	1.0	11548	409	1.0	18376	698	1.0	4100	982	0.98	69960	4544	1.0
Schwefel222	115859	11836	0.46	12324	568	1.0	15709	511	1.0	4943	635	1.0	74771	8138	0.94
Schwefel223	16868	686	1.0	10450	41	1.0	16888	577	1.0	2393	148	1.0	32622	1430	1.0
Sphere	921	67	1.0	10013	0	1.0	36873	1075	1.0	3524	472	1.0	21472	307	1.0
Step	707	0	0.0	10091	16	1.0	2775	320	1.0	3835	786	0.46	27776	650	1.0
Step2	707	0	0.0	10093	17	1.0	2830	406	1.0	3556	756	0.6	27970	760	1.0
Step3	707	0	0.0	10101	20	1.0	2764	282	1.0	3958	705	0.56	27786	613	1.0
StretchedV SineWave	12512	3271	0.0	11485	513	0.0	70214	8716	0.0	4868	2056	0.02	70630	21089	0.0
SumSquares	8757	103	1.0	10069	4	1.0	15352	558	1.0	2427	225	1.0	48238	2280	1.0
Trigonometric1	1334	381	0.96	10211	97	0.74	19919	4577	0.94	2379	312	0.8	21115	317	1.0
Trigonometric2	61209	2235	0.0	12641	477	1.0	5823	686	1.0	6383	1432	0.58	99812	22224	0.28
WWavy	6590	1751	0.0	10340	76	1.0	13221	907	1.0	3454	592	1.0	52355	8667	0.0
Weierstrass	67323	2585	0.0	13265	471	0.0	75201	0	0.0	8846	2923	0.0	54581	13629	0.0
Whitley	44508	3693	0.0	11484	448	0.32	14540	3778	0.14	3895	706	0.2	71142	12189	0.02
XinSheYangN3	1152	237	0.0	10289	53	0.54	282	42	0.16	2330	61	0.02	30334	12807	0.32
XinSheYangN4	7806	1467	0.0	10080	27	0.0	14059	1027	0.0	2414	310	0.04	41851	6388	0.02
Всього в середньому:	27956	36013	0.45	10697	836	0.69	20839	18627	0.7	3517	1631	0.64	58063	34567	0.54

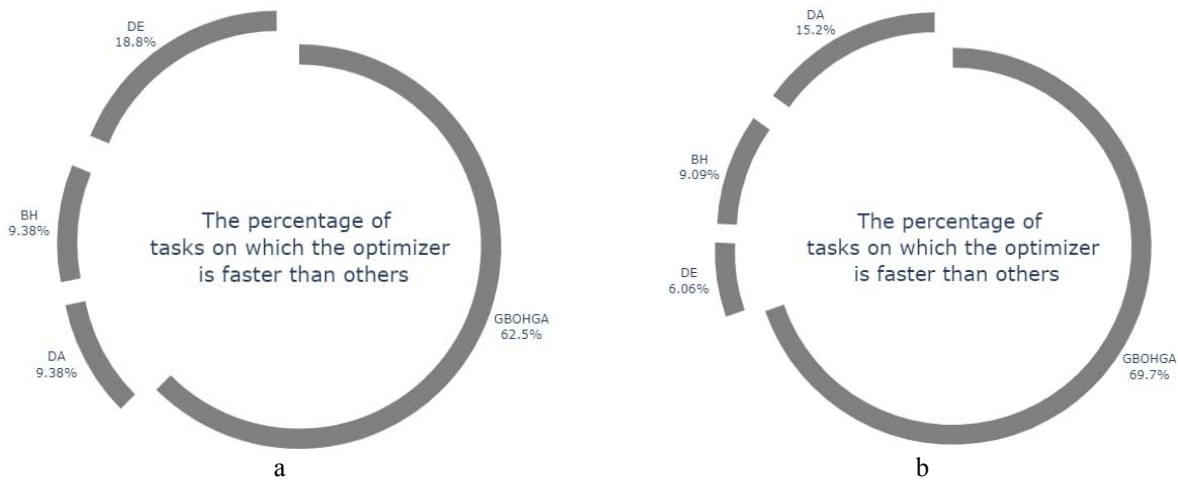


Figure 4 – The percentage of problems on which the optimizer is faster than others: a – with 5; b – with 25 variables



Figure 5 – Performance profiles in problems with 25 variables

6 DISCUSSION

Therefore, based on the results of numerical experiments of solving known global optimization problems, the proposed GBOHGA algorithm demonstrated the following advantages:

- due to the reduction in the number of evaluations of the objective function, the performance speed of the algorithm is higher compared to any other considered evolutionary algorithm (almost 70% of problems with 25 variables were solved faster using GBOHGA than using other solvers);
- by the criterion of reliability, GBOHGA is better than the basin-hopping algorithm and particle swarm optimization in problems with 5 variables;
- by the criterion of the robustness of solving problems with 25 variables, the algorithm is better than three of the four other solvers among the considered.

The work [29] shows the importance of the development and wide applicability of mobile neural networks. Such networks, as a rule, have parameters, the choice of

which is necessary to achieve the minimum value of the loss function. Genetic algorithms (including ours) have the potential to improve the quality of such neural networks.

CONCLUSIONS

The paper proposes a new hybrid genetic algorithm GBOHGA, an optimization scheme that combines elements of BOHGA and HGDN methods and which integrates the strengths of local search with powerful global search capabilities. The advantage of the proposed method can be considered the simplicity of its implementation, as the key operators of the genetic algorithm and local search methods remain unchanged. Due to the prevention of a significant part of the unnecessary use of the local search operator, significant savings were achieved in the number of evaluations of the objective function and execution time of the algorithm, which made the method competitive with other widely used methods of global optimization.

Table 4 – Mean value and standard deviation of the number of function evaluations and success rate after 10 runs for each function with 25 variables from the test set for each method

Function name	BH			DA			DE			GBOHGA			PSO		
	μ	σ	%	μ	σ	%	μ	σ	%	μ	σ	%	μ	σ	%
Ackley1	69614	696	0.0	57876	690	1.0	375609	32	0.0	12222	3230	1.0	89650	4805	0.0
Alpine1	679007	9705	0.9	106564	13930	1.0	375921	0	1.0	42749	11233	1.0	96936	5266	0.0
Chung Reynolds	71201	74140	0.9	53552	193	1.0	327326	17048	1.0	3446	74	1.0	112826	7586	0.0
Exponential	19350	162	1.0	50222	39	1.0	19226	2552	1.0	4270	1350	1.0	78228	4271	0.0
Griewank	95882	260	1.0	50921	208	1.0	71432	72252	0.3	2934	464	1.0	84713	4532	0.0
HappyCat	117521	2317	0.0	54454	1293	0.0	377231	852	0.0	7539	2368	0.0	105716	4896	0.0
Periodic	27205	47	0.0	50248	46	0.0	375531	0	0.0	2925	291	0.0	83399	6435	0.0
PowellSum	126551	2321	1.0	52819	342	1.0	375401	0	1.0	2572	153	1.0	50547	5022	1.0
Rastrigin	206042	6159	0.0	55903	872	1.0	375658	25	0.0	7739	2091	1.0	98980	5377	0.0
Rosenbrock	299921	43662	0.1	54974	505	1.0	288349	141600	0.9	7858	1246	1.0	114993	9423	0.0
Salomon	34913	12131	0.8	54005	2108	0.0	340540	39914	0.0	5742	1442	0.2	45660	3652	0.0
Sargan	50216	60667	0.9	50315	59	1.0	375401	0	1.0	3862	973	1.0	112947	5184	0.0
SchafferF6	338304	36814	0.0	106756	10820	0.1	384706	4170	0.0	52606	11817	0.8	96166	7029	0.0
Schumer Steiglitz	17080	2347	0.0	54072	205	1.0	372626	5500	1.0	3277	128	1.0	108191	9569	0.0
Schwefel	188737	9032	0.0	53854	1343	0.0	328746	10168	1.0	5557	1477	0.0	82255	5516	0.0
Schwefel12	68272	10944	0.9	52554	339	1.0	377137	227	1.0	5503	721	1.0	103254	4828	0.0
Schwefel220	695733	6717	1.0	56755	2396	1.0	329758	14548	1.0	10814	1445	1.0	99846	9536	0.0
Schwefel221	498875	17899	1.0	75031	4684	1.0	375785	89	1.0	26047	7896	0.3	88771	7744	0.0
Schwefel222	341734	107144	0.0	73008	4702	1.0	366096	8306	1.0	23501	8166	1.0	116924	18060	0.0
Schwefel223	34150	39015	0.3	55713	296	1.0	375401	0	1.0	3638	160	1.0	96984	9145	0.0
Sphere	5043	236	1.0	50053	0	1.0	375427	0	1.0	7031	2890	1.0	25165	1330	0.9
Step	2727	0	0.0	51316	109	1.0	58713	5417	1.0	9532	2276	1.0	37924	1620	0.0
Step2	2727	0	0.0	51233	101	1.0	60326	7288	1.0	9625	2363	0.9	40472	3364	0.0
Step3	2727	0	0.0	51462	113	1.0	57626	6490	1.0	8432	2868	0.8	45182	5523	0.0
StretchedV SineWave	266810	83587	0.0	149045	41857	0.0	381953	1510	0.0	28011	41761	0.0	95680	5320	0.0
SumSquares	134815	12197	0.2	50775	95	1.0	329988	14109	1.0	4710	1188	1.0	96382	5384	0.0
Trigonometric1	52392	35191	0.6	52520	753	0.0	101946	96678	0.1	4426	1028	0.0	50013	24921	0.4
Trigonometric2	345622	12423	0.0	153122	17328	1.0	227821	45852	1.0	58542	11799	0.5	154987	15666	0.0
WWavy	81105	4064	0.0	54506	487	1.0	375991	123	0.0	10902	1623	1.0	80479	6991	0.0
Weierstrass	309193	12501	0.0	101761	5373	0.0	375921	0	0.0	41843	26637	0.0	65630	17056	0.0
Whitley	267065	168786	0.0	79505	3630	0.2	333035	46053	0.0	19983	1315	0.0	145926	7956	0.0
XinSheYangN3	6185	941	0.0	54670	165	0.0	2814	486	0.0	4129	79	0.0	23832	551	0.0
XinSheYangN4	6515	1496	0.0	55450	432	0.0	373601	5692	0.0	2956	126	0.0	30242	117	0.0
Total average:	165552	190266	0.35	65909	28012	0.68	292213	133191	0.58	13483	17486	0.65	83603	33631	0.07

According to the results of testing on a set of 33 functions with different landscapes, the proposed algorithm showed high results in terms of performance, and it is important to emphasize that the greatest benefits from its use can be obtained in problems of large dimensionality of search space.

The practical value of the work lies in the possibility of using the developed algorithm to solve global optimization problems that arise in engineering design, solving organizational and management problems, especially when the mathematical model of the problem has a high dimensionality.

The direction of further research is the analysis of the duration of local search, necessary to find a global solution to the optimization problem with acceptable accuracy, as well as the choice of truncation criterion in the proposed scheme.

ACKNOWLEDGEMENTS

The work was performed within the framework of the international cooperation program between the Dnipro University of Technology and the University of Koblenz-Landau (Germany). The authors would like to thank Doctor of Sciences, Professor of the Department of Modeling © Avramenko S. E., Zheldak T. A., Koriashkina L. S., 2021 DOI 10.15588/1607-3274-2021-2-18

Simulation & Optimization of Complex Systems University of Koblenz-Landau Michael Hinze for support and a fruitful paper discussion.

REFERENCES

1. Locatelli M., Schoen F. Global optimization: theory, algorithms, and applications, *Society for Industrial and Applied Mathematics*, 2013, 432 p. DOI: 10.1137/1.9781611972672
2. Floudas C. A., Gounaris C. E. A review of recent advances in global optimization, *Journal of Global Optimization*, 2009, Vol. 45, No. 3, pp. 3–38. DOI 10.1007/s10898-008-9332-8
3. Fasano G., Pinter J. D. Space engineering: modeling and optimization with case studies. Switzerland, Springer, 2016, 492 p. DOI 10.1007/978-3-319-41508-6_18
4. Rao S. S. Engineering optimization: theory and practice. Hoboken, New Jersey, John Wiley & Sons, 2009, 813 p.
5. Siddique N., Adeli H. Nature inspired computing: an overview and some future directions, *Cognitive Computation*, 2015, Vol. 7, No. 6, pp. 706–714.
6. Wan W., Birch J. B. An improved hybrid genetic algorithm with a new local search procedure, *Journal of Applied Mathematics*, 2013, Vol. 2013, Article ID 103591, 10 p. DOI: 10.1155/2013/103591
7. Noack M. M., Funke S. W. Hybrid genetic deflated newton method for global optimisation, *Journal of Computational*

- and Applied Mathematics, 2017, Vol. 325, pp. 97–112. DOI: 10.1016/j.cam.2017.04.047
8. Voudouris C., Tsang E.P., Alsheddy A. Guided local search, *Handbook of Metaheuristics. International Series in Operations Research & Management Science* Eds.: M. Gendreau, J.Y. Potvin. Boston, Springer, 2010, Vol. 146, pp. 321–361. DOI: 10.1007/978-1-4419-1665-5_11
 9. Pintér J. D. Global Eds.: P. M. Pardalos, H. E. Romeijn Optimization: Software, Test Problems, and Applications, *Handbook of Global Optimization. Nonconvex Optimization and Its Applications*. Boston, Springer, 2002, Vol. 62, pp. 515–569. DOI: 10.1007/978-1-4757-5362-2_15
 10. Pardalos P. M., Zhigljavsky A. A., Žilinskas J. Advances in Stochastic and Deterministic Global Optimization. New York, Springer, 2016, 296 p. DOI: 10.1007/978-3-319-29975-4
 11. Paulavičius R., Žilinskas J. Simplicial Global Optimization, *Journal of Global Optimization*, 2014, Vol. 60, No. 4, pp. 61–86. DOI: 10.1007/978-1-4614-9093-7_3
 12. Beliakov G. Cutting angle method – A tool for constrained global optimization, *Optimization Methods and Software*, 2004, Vol. 19, № 2, pp. 137–151.
 13. Kvasov D. E., Sergeyev Y. D. Univariate geometric Lipschitz global optimization algorithms, *Numerical Algebra, Control and optimization*, 2012, Vol. 2, No. 1, pp. 69–90.
 14. Malherbe C., Vayatis N. Global optimization of Lipschitz functions, *Machine Learning : 34th International Conference, Sydney, 6–11 August 2017 : proceedings*. Australia, PMLR, 2017, Vol. 70, pp. 2314–2323.
 15. Rios L. M., Sahinidis N. V. Derivative-free optimization: a review of algorithms and comparison of software implementations, *Journal of Global Optimization*, 2013, Vol. 56, pp. 1247–1293. DOI 10.1007/s10898-012-9951-y
 16. Sergeyev Y. Mukhametzhano M., Kvasov D. et al. Derivative-free local tuning and local improvement techniques embedded in the univariate global optimization, *Journal of Optimization Theory and Applications*, 2017, Vol. 171, pp. 186–208. DOI: 10.1007/s10957-016-0947-5
 17. Ye P., Pan G., Dong Z. Ensemble of surrogate based global optimization methods using hierarchical design space reduction, *Structural and Multidisciplinary Optimization*, 2018, Vol. 58, pp. 537–554. DOI: 10.1007/s00158-018-1906-6
 18. Pengcheng Ye. A Review on Surrogate-Based Global Optimization Methods for Computationally Expensive Functions, *Software Engineering*, 2019, Vol. 7, No. 4, pp. 68–84. DOI: 10.11648/j.se.20190704.11
 19. Li Z., Ruan Sh., Gu J. et al. Investigation on parallel algorithms in efficient global optimization based on multiple points infill criterion and domain decomposition, *Structural and Multidisciplinary Optimization*, 2016, Vol. 54, No. 4, pp. 747–773. DOI: 10.1007/s00158-016-1441-2
 20. Wang Z., Gehring C., Kohli P. et al. Batched large-scale Bayesian optimization in high-dimensional spaces, *Artificial Intelligence and Statistics : Twenty-First International Conference, Lanzarote, 9–11 April 2018 : proceedings*. Spain, PMLR, 2018, Vol. 84, pp. 745–754.
 21. Eriksson D., Pearce M., Gardner J. R. et al. Scalable Global Optimization via Local Bayesian Optimization, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver*. Canada, BC, 2019, pp. 5497–5508.
 22. Mitchell M. An introduction to genetic algorithms. Cambridge, Mass., MIT Press, 1998, 221 p.
 23. Farrell P. E., Biriksson Á., S. W. Funke Deflation techniques for finding distinct solutions of nonlinear partial differential equations, *Society for Industrial and Applied Mathematics. Journal on Scientific Computing*, 2015, Vol. 37, No. 4, pp. A2026–A2045.
 24. Nieto-Fuentes R., Segura C., Valdez S. I. A Guided Local Search Approach for the Travelling Thief Problem, *IEEE Congress on Evolutionary Computation, Rio de Janeiro, 8–13 July 2018. Brazil, 2018*, pp. 1–8. DOI: 10.1109/CEC.2018.8477821
 25. Touat M. Tayeb F. B., Benhamou B. et al. An Integrated Guided Local Search considering Human Resource Constraints for the Single-machine Scheduling problem with Preventive Maintenance, *IEEE International Conference on Systems, Man and Cybernetics, Bari, 6–9 October 2019. Italy, 2019*, pp. 3799–3804. DOI: 10.1109/SMC.2019.8914261
 26. Peh S. C. W., Hong J. L. Eds.: Gervasi O. et al. GLSDock – Drug Design Using Guided Local Search, *Computational Science and Its Applications. Lecture Notes in Computer Science: 16th International Conference, Beijing, July 4–7 2016 : proceedings*. China, Springer, 2016, Vol. 9788, pp. 11–21. DOI: 10.1007/978-3-319-42111-7_2
 27. Jamil M., Yang X. S. A literature survey of benchmark functions for global optimisation problems, *International Journal of Mathematical Modelling and Numerical Optimisation*, 2013, Vol. 4, No. 2, pp. 150–194. DOI: 10.1504/IJMMNO.2013.055204
 28. Dall’Igna Júnior A., Silva R. S., Mundim K. C. et al. Performance and parameterization of the algorithm simplified generalized simulated annealing, *Genetics and Molecular Biology*, 2004, Vol. 27, № 4, pp. 616–622.
 29. Khabaralak K., Koriashkina L. Fast Facial Landmark Detection and Applications: A Survey, arXiv: 2101.10808, 2021, 20 p.

Received 25.02.2021.
Accepted 30.04.2021.

УДК 004.023

КЕРОВАНИЙ ГІБРИДНИЙ ГЕНЕТИЧНИЙ АЛГОРИТМ РОЗВ’ЯЗАННЯ ЗАДАЧ ГЛОБАЛЬНОЇ ОПТИМІЗАЦІЇ

Авраменко С. Є. – магістр кафедри системного аналізу і управління Національного технічного університету «Дніпровська політехніка», інженер з машинного навчання, COMPARUS.UA, Дніпро, Україна.

Желдак Т. А. – канд. техн. наук, завідувач кафедри системного аналізу і управління Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.

Коряшкіна Л. С. – канд. фіз.-мат. наук, доцент кафедри системного аналізу і управління Національного технічного університету «Дніпровська політехніка», Дніпро, Україна.

АНОТАЦІЯ

Актуальність. Провідною проблемою в світі штучного інтелекту є оптимізація складних систем, що нерідко подається у вигляді нелінійної функції, яку необхідно мінімізувати. Такі функції можуть виявитися мультимодальними, недиференційованими, і навіть, заданими у формі чорної скриньки. Побудова ефективних методів розв'язання задач глобальної оптимізації і сьогодні викликає значний інтерес серед науковців.

Мета. Розробка нового гібридного генетичного алгоритму розв'язання задачі безумовної глобальної оптимізації, швидшого за існуючі аналоги.

Методи. Одним з важливих викликів, з якими стикаються гібридні методи під час розв'язування задач нелінійної глобальної оптимізації, є раціональне використання локального пошуку, оскільки його реалізація супроводжується достатньо коштовними обчислювальними витратами. В даній роботі запропонований новий гібридний генетичний алгоритм GBOHGA, який відтворює керований локальний пошук, і поєднує дві успішні модифікації генетичних алгоритмів, перша з яких BOHGA встановлює якісний баланс між локальним та глобальним пошуком, друга – HGDN – запобігає повторному дослідженню раніше вже розвіданої області пошукового простору. Окрім цього, було запропоновано модифіковану *intra*-функцію та адаптивну схему визначення її параметру – радіусу області «дефляції» цільової функції в околі вже знайденого локального мінімуму – задля прискорення роботи алгоритму.

Результати. Роботу GBOHGA з іншими відомими стохастичними пошуковими евристичними алгоритмами на наборі із 33 тестових функцій в 5 та 25 вимірному просторах. Результати обчислювальних експериментів свідчать про конкурентну спроможність GBOHGA, особливо в задачах з мультимодальними функціями і великою кількістю змінних.

Висновки. Новий гібридний алгоритм GBOHGA, розроблений на основі інтеграції ідей керованого локального пошуку і алгоритмів BOHGA та HGDN, дозволяє значно економити обчислювальні ресурси та прискорювати розв'язання задачі глобальної оптимізації. Його варто застосовувати для розв'язання задач глобальної оптимізації, що виникають в інженерному проектуванні, вирішення організаційно-управлінських проблем, особливо коли математична модель задачі є складною і має високу розмірність.

КЛЮЧОВІ СЛОВА: нелінійна оптимізація, глобальний мінімум, стохастичні пошукові евристичні, гібридний підхід, генетичний алгоритм, оператор дефляції, керований локальний пошук.

УДК 004.023

УПРАВЛЯЕМЫЙ ГИБРИДНЫЙ ГЕНЕТИЧЕСКИЙ АЛГОРИТМ РЕШЕНИЯ ЗАДАЧ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

Авраменко С. С. – магистр кафедры системного анализа и управления Национального технического университета «Днепропетровская политехника», инженер по машинному обучению, COMPARUS.UA, Днепр, Украина.

Желдак Т. А. – канд. техн. наук, заведующий кафедрой системного анализа и управления Национального технического университета «Днепропетровская политехника», Днепр, Украина.

Коряшкина Л. С. – канд. физ.-мат. наук, доцент кафедры системного анализа и управления Национального технического университета «Днепропетровская политехника», Днепр, Украина.

АННОТАЦИЯ

Актуальность. Одной из ведущих проблем в мире искусственного интеллекта является оптимизация сложных систем, нередко представляемая в виде нелинейной функции, которую необходимо минимизировать. При этом функции могут оказаться многоэкстремальными, негладкими, и даже, заданными в форме черного ящика. Построение эффективных методов решения таких задач и сегодня вызывает значительный интерес в научном сообществе.

Цель. Разработка нового гибридного генетического алгоритма решения задачи глобальной оптимизации, который по скорости превосходил бы существующие аналоги.

Метод. Одним из важных вызовов, с которыми сталкиваются гибридные методы при решении задач нелинейной глобальной оптимизации, является рациональное использование локального поиска, поскольку его реализация сопровождается высокими вычислительными затратами. В данной работе предложен новый гибридный генетический алгоритм GBOHGA, который воспроизводит управляемый локальный поиск, и объединяет две успешные модификации генетических алгоритмов, первая – BOHGA – устанавливает качественный баланс между локальным и глобальным поиском, вторая – HGDN – предотвращает повторное исследование ранее уже разведанных областей поискового пространства. Кроме этого, для ускорения работы алгоритма предложена модифицированная *intra*-функция и адаптивная схема определения ее параметра, а именно, радиуса области «дефляции» целевой функции в окрестности уже найденного локального минимума.

Результаты. Проведено сравнение работы GBOHGA с другими известными стохастическими поисковыми эвристичными алгоритмами на наборе из 33 тестовых функций в 5 и 25 мерном пространствах. Результаты вычислительных экспериментов свидетельствуют о конкурентной способности GBOHGA, особенно при решении задач с мультимодальными функциями и большим количеством переменных.

Выводы. Новый гибридный алгоритм GBOHGA, разработанный на основе интеграции идей управляемого локального поиска и алгоритмов BOHGA и HGDN, позволяет значительно экономить вычислительные ресурсы и ускорять решение задачи глобальной оптимизации. Его целесообразно применять для решения задач глобальной оптимизации, возникающих в инженерном проектировании, решения организационно-управленческих проблем, особенно когда математическая модель задачи является сложной и имеет высокую размерность.

КЛЮЧЕВЫЕ СЛОВА: нелинейная оптимизация, глобальный минимум, стохастические поисковые эвристичные, гибридный подход, генетический алгоритм, оператор дефляции, управляемый локальный поиск.

ЛІТЕРАТУРА / LITERATURE

1. Locatelli M. Global optimization: theory, algorithms, and applications / M. Locatelli, F. Schoen. – Society for Industrial and Applied Mathematics, 2013. – 432 p. DOI: 10.1137/1.9781611972672
2. Floudas C.A. A review of recent advances in global optimization / C.A. Floudas, C.E. Gounaris // *Journal of Global Optimization*. – 2009. – Vol. 45, № 3. – P. 3–38. DOI 10.1007/s10898-008-9332-8
3. Fasano G. Space engineering: modeling and optimization with case studies / G. Fasano, J.D. Pinter. – Switzerland: Springer, 2016. – 492 p. DOI 10.1007/978-3-319-41508-6_18
4. Rao S. S. Engineering optimization: theory and practice / S. S. Rao. – Hoboken, New Jersey : John Wiley & Sons, 2009. – 813 p.
5. Siddique N. Nature inspired computing: an overview and some future directions / N. Siddique, H. Adeli // *Cognitive Computation*. – 2015. – Vol. 7, № 6. – P. 706–714.
6. Wan W. An improved hybrid genetic algorithm with a new local search procedure / W. Wan, J. B. Birch // *Journal of Applied Mathematics*. – 2013. – Vol. 2013. – Article ID 103591, 10 p. DOI:10.1155/2013/103591
7. Noack M. M. Hybrid genetic deflated newton method for global optimisation / M. M. Noack, S. W. Funke // *Journal of Computational and Applied Mathematics*. – 2017. – Vol. 325. – P. 97–112. DOI: 10.1016/j.cam.2017.04.047
8. Voudouris C. Guided local search / C. Voudouris, E. P. Tsang, A. Alsheddy // *Handbook of Metaheuristics. International Series in Operations Research & Management Science / Eds.: M. Gendreau, J.Y. Potvin*. – Boston : Springer, 2010. – Vol. 146. – P. 321–361. DOI: 10.1007/978-1-4419-1665-5_11
9. Pintér J. D. Global Optimization: Software, Test Problems, and Applications / János D. Pintér // *Handbook of Global Optimization. Nonconvex Optimization and Its Applications / Eds.: P. M. Pardalos, H. E. Romeijn*. – Boston : Springer, 2002. – Vol. 62. – P. 515–569. DOI: 10.1007/978-1-4757-5362-2_15
10. Pardalos P. M. Advances in Stochastic and Deterministic Global Optimization / P. M. Pardalos, A. A. Zhigljavsky, J. Žilinskas. – New York : Springer, 2016. – 296 p. –DOI: 10.1007/978-3-319-29975-4
11. Paulavičius R. Simplicial Global Optimization / R. Paulavičius, J. Žilinskas // *Journal of Global Optimization*. – 2014. – Vol. 60, № 4. – P. 61–86. DOI: 10.1007/978-1-4614-9093-7_3
12. Beliakov G. Cutting angle method – A tool for constrained global optimization / G. Beliakov // *Optimization Methods and Software*. – 2004. – Vol. 19, № 2. – P. 137–151.
13. Kvasov D. E. Univariate geometric Lipschitz global optimization algorithms / D. E. Kvasov, Y. D. Sergeyev // *Numerical Algebra, Control and optimization*. – 2012. – Vol. 2, № 1. – P. 69–90.
14. Malherbe C. Global optimization of Lipschitz functions. / C. Malherbe, N. Vayatis // *Machine Learning : 34th International Conference, Sydney, 6–11 August 2017 : proceedings*. – Australia : PMLR, 2017. – Vol. 70. – P. 2314–2323.
15. Rios L. M. Derivative-free optimization: a review of algorithms and comparison of software implementations / L. M. Rios, N. V. Sahinidis // *Journal of Global Optimization*. – 2013. – Vol. 56. – P. 1247–1293. – DOI 10.1007/s10898-012-9951-y
16. Sergeyev Y. Derivative-free local tuning and local improvement techniques embedded in the univariate global optimization / [Y. Sergeyev, M. Mukhametzanov, D. Kvasov et al.] // *Journal of Optimization Theory and Applications*. – 2017. – Vol. 171. – P. 186–208. DOI: 10.1007/s10957-016-0947-5
17. Ye P. Ensemble of surrogate based global optimization methods using hierarchical design space reduction / P. Ye, G. Pan, Z. Dong // *Structural and Multidisciplinary Optimization*. – 2018. – Vol. 58. – P. 537–554. DOI: 10.1007/s00158-018-1906-6
18. Pengcheng Ye. A Review on Surrogate-Based Global Optimization Methods for Computationally Expensive Functions / Ye. Pengcheng // *Software Engineering*. – 2019. – Vol. 7, № 4. – P. 68–84. DOI: 10.11648/j.se.20190704.11
19. Investigation on parallel algorithms in efficient global optimization based on multiple points infill criterion and domain decomposition / [Z. Li, Sh. Ruan, J. Gu et al.] // *Structural and Multidisciplinary Optimization*. – 2016. – Vol. 54, № 4. – P. 747–773. DOI: 10.1007/s00158-016-1441-2
20. Wang Z. Batched large-scale Bayesian optimization in high-dimensional spaces / [Z. Wang, C. Gehring, P. Kohli et al.] // *Artificial Intelligence and Statistics : Twenty-First International Conference, Lanzarote, 9–11 April 2018 : proceedings*. – Spain : PMLR, 2018. – Vol. 84. – P. 745–754.
21. Eriksson D. Scalable Global Optimization via Local Bayesian Optimization / [D. Eriksson, M. Pearce, J. R. Gardner et al.] // *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver*. – Canada: BC, 2019. – P. 5497–5508.
22. Mitchell M. An introduction to genetic algorithms / M. Mitchell. – Cambridge, Mass. : MIT Press, 1998. – 221 p.
23. Farrell P. E. Deflation techniques for finding distinct solutions of nonlinear partial differential equations / P. E. Farrell, Á. Birkisson, S. W. Funke // *Society for Industrial and Applied Mathematics. Journal on Scientific Computing*. – 2015. – Vol. 37, № 4. – P. A2026–A2045.
24. Nieto-Fuentes R. A Guided Local Search Approach for the Travelling Thief Problem / R. Nieto-Fuentes, C. Segura, S. I. Valdez // *IEEE Congress on Evolutionary Computation, Rio de Janeiro, 8–13 July 2018*. – Brazil, 2018. – P. 1–8. DOI: 10.1109/CEC.2018.8477821
25. An Integrated Guided Local Search considering Human Resource Constraints for the Single-machine Scheduling problem with Preventive Maintenance / [M. Touat, F. B. Tayeb, B. Benhamou et al.] // *IEEE International Conference on Systems, Man and Cybernetics, Bari, 6–9 October 2019*. – Italy : 2019. – P. 3799–3804. – DOI: 10.1109/SMC.2019.8914261
26. Peh S. C. W. GLSDock – Drug Design Using Guided Local Search / S. C. W. Peh, J. L. Hong // *Computational Science and Its Applications. Lecture Notes in Computer Science*. / Eds.: Gervasi O. et al. // *16th International Conference, Beijing, China, July 4–7 2016 : proceedings*. – Springer, 2016. – Vol 9788. – P. 11–21. – DOI: 10.1007/978-3-319-42111-7_2
27. Jamil M. A literature survey of benchmark functions for global optimisation problems / M. Jamil, X. S. Yang // *International Journal of Mathematical Modelling and Numerical Optimisation*. – 2013. – Vol. 4, № 2. – P. 150–194. DOI: 10.1504/IJMMNO.2013.055204
28. Dall’Igna Júnior A. Performance and parameterization of the algorithm simplified generalized simulated annealing / [A. Dall’Igna Júnior, R. S. Silva, K. C. Mundim et al.] // *Genetics and Molecular Biology*. – 2004. – Vol. 27, № 4. – P. 616–622.
29. Khabaralak K. Fast Facial Landmark Detection and Applications: A Survey / K. Khabaralak, L. Koriashkina // *arXiv: 2101.10808*. – 2021. – 20 p.

CRITERIA FOR ESTIMATING THE SENSORIMOTOR REACTION TIME BY THE SMALL UAV OPERATOR

Vakaliuk T. A. – Dr. Sc., Professor, Professor at the Department of software engineering, Zhytomyr Polytechnic State University, Zhytomyr, Ukraine.

Pilkevych I. A. – Dr. Sc., Professor, Professor at the Department of computer information technologies, Zhytomyr military institute named after S. P. Korolov, Zhytomyr, Ukraine.

Tokar A. M. – PhD, Head of the Division of science center, Zhytomyr military institute named after S. P. Korolov, Zhytomyr, Ukraine.

Loboda R. I. – Research Officer of the Division of science center, Zhytomyr military institute named after S. P. Korolov, Zhytomyr, Ukraine.

ABSTRACT

Context. The rapid development of science and technology predetermines a significant expansion of the fields of application of UAVs different purposes. The key to the effective use UAVs is high-quality training of operators, an important element of which is the PPS of candidates, in particular, the assessment of their sensorimotor reactions. This can be achieved by selecting and justifying appropriate criteria.

Objective. The goal of the work is the justification criteria for estimating the time sensorimotor reactions of a small UAV operator by analyzing the density distribution of statistical data.

Method. A method has been developed to determine criteria for evaluating the time of sensorimotor reactions a small UAV operator based on the accumulation statistical material and its mathematical processing based on the results of a field experiment. The method allows to estimate numerical characteristics the distribution of the average reaction time in three modes: training production, in the conditions overload, in the conditions of overtraining and to obtain a generalized estimation. It was possible, by analyzing the occasional non-interruptible values, which take values within a certain range of values, to establish standards against which the obtained values the sensorimotor reaction time of the small UAV operator are compared and a decision is made on their suitability for training.

Results. We obtained statistical series for the modes of assessment: skill development, under obstacle conditions, under conditions skill restructuring. For a visual representation of the series the corresponding histograms the distribution of the average reaction time duration were constructed. In order to eliminate the representativeness error, statistical series alignment was carried out by selecting a theoretical distribution curve for each series, which displays only essential features of the statistical material. For this purpose, we approximated the histogram of distribution by the polynomial fourth degree. The interval theoretical density of distribution, in which the time sensomotor reaction of an arbitrary person is considered normal, with a given probability reliability such event – 0.95 has been established. To verify the effectiveness of the proposed method, algorithms for estimating the sensorimotor reaction time of a small UAV operator in three modes have been synthesized and the corresponding software that implements the proposed algorithms has been developed.

Conclusions. The criteria for evaluating the sensorimotor reaction time for UAV operator to a visual stimulus using specialized software were substantiated. This allowed the previous PPS training candidates to take into account the requirements to the motor skills of the small UAV operator and the specificity his movements. The conducted experiments confirmed the validity of decisions made. Prospects for further research may include expansion of testing modes with justification for appropriate evaluation criteria.

KEYWORDS: evaluation criteria, small UAV operator, sensorimotor reaction time, distribution density.

ABBREVIATIONS

PC is a personal computer;
PPS is a professional psychological selection;
UAV is an unmanned aerial vehicle;
USB is an Universal Serial Bus.

NOMENCLATURE

a_m is a value of the minimal sensorimotor reaction time operators in m -th testing mode;
 b_m is a value of maximum allowable operator's sensorimotor reaction time in m -th testing mode;
 $f(t)$ is a distribution density of average reaction time duration;
 $f_m(t)$ is a distribution density of average reaction time duration in m -th testing mode;
 K is a coefficient on 9-th point scale;
 m is a number of evaluation modes of sensorimotor reaction, $m=3$;

m_i is a mathematical expectation of mean sensorimotor reaction time estimates;
 n is a number of people being tested;
 t is a time interval;
 t_{mn} is a sensorimotor reaction time of n -th operator in m -th testing mode;
 T_N is a normative operator's sensorimotor reaction time;
 T_m is a measured operator's sensorimotor reaction time;
 $X(t)$ is a number of people whose reaction time falls within the time interval t ;
 σ is a standard deviation of mean sensorimotor reaction time estimates;
 $\hat{\sigma}$ is an estimation of standard deviation of mean sensorimotor reaction time estimates.

INTRODUCTION

To date, the list of industries where UAVs are used is rapidly expanding. The effectiveness of their application depends both on the technical characteristics the UAV and on the actions operators. Today, methods research technical systems allow successfully establishing the laws of occurrence device failures and methods of their forecasting. This makes it possible to design and manufacture optimal devices, of which the onboard equipment for UAVs is completed.

At the same time, it has been established that one of the causes reduced effectiveness of small UAV application is insufficient operator professional reliability [1–5], which is understood as the reliability, accuracy and timeliness their actions aimed at performing the task during UAV control [6]. Operator's professional reliability directly depends on his psycho-physiological qualities, high manifestations which is one factors of operators' reliable work [7, 8]. The main direction to ensure the necessary level of reliability operators is the formation of their professional suitability, which is understood as the availability human psychophysiological and psychological properties, which correspond to a particular profession and provide sufficient efficiency of professional activity [6, 7]. The process of professional suitability formation involves creating a system for professional selection and training for a certain profile of operator activity (methods, tools, programs, criteria, simulators and the like) [6, 9]. An important component in this system is the professional psychological selection (PPS) of candidates for training, which consists in carrying out a set to ensure the quality selection the candidates, by assessing the compliance of the level development necessary psychophysiological qualities and personality properties with the requirements of professional activity [10, 11]. One of main stages PPS small UAV operators is evaluation evaluation of their sensorimotor reactions, which involves the use specialized complex (hardware, hardware-software, software). The efficiency of this process largely depends on the selected criteria for evaluating the obtained reaction time on the small UAV operator.

The object of study is the process evaluation of sensorimotor reaction time a small UAV operator.

The subject of study is the method of evaluating criteria for the sensorimotor reaction time of a small UAV operator.

The purpose of the work is to improve PPS quality of small UAV operators by evaluating their sensorimotor reaction time.

1 PROBLEM STATEMENT

To estimate the level of operator fatigue for different activity profiles, their readiness for action and fitness for professional activity, the results of sensomotor reaction duration measurement (a single discrete movement an operator to the appearance (cessation of action) a stimulus) are used [12–19]. The use methods for

© Vakaliuk T. A., Pilkevych I. A., Tokar A. M., Loboda R. I., 2021
 DOI 10.15588/1607-3274-2021-2-19

measuring the duration is based on the fact that each elementary motor response is the result of a complex interaction different human brain systems. Reaction duration is a manifestation of both individual peculiarities a nervous system and a person's state at a given moment. Depending on the profile of professional activity, the purpose and conditions of the study, the type of reaction being monitored, different complexes (devices) for evaluating the operator's sensorimotor reactions are used. At the same time appropriate methods of processing and evaluating measurement results are used [14, 15].

The development for computing technology has led to the use specialized software products for the evaluation of sensorimotor reactions operators for different activity profiles, in particular UAV operators [16, 19]. The use such systems involves performing a certain action using the keyboard, manipulator, external devices in response to stimuli that are displayed on the monitor.

Taking into account current trends, we will evaluate the sensorimotor reaction time of small UAV operators, according to the scheme, which is shown in Fig. 1.

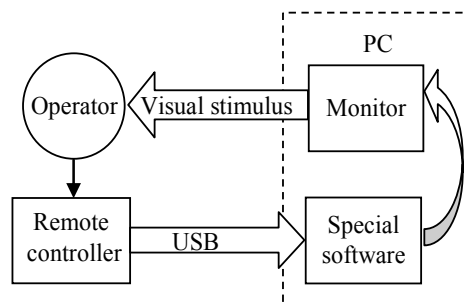


Figure 1 – The scheme for evaluating the time of sensorimotor reactions by small UAV operators

The result of measuring the time of sensorimotor reactions is an amount of data:

$$\begin{bmatrix} t_{11}, t_{12}, \dots, t_{1n}, \\ t_{21}, t_{22}, \dots, t_{2n}, \\ t_{m1}, t_{m2}, \dots, t_{mn}. \end{bmatrix}$$

In this case, the task of justification criteria for estimating the sensorimotor reaction time the small UAV operator is to find such values a_m i b_m , for which the condition is met

$$\int_{a_m}^{b_m} f_m(t) dt = 0.95. \quad (1)$$

2 REVIEW OF THE LITERATURE

As yet, little is known about the specific demands on operating UAV personnel. The different unmanned aerial systems that are currently in use differ in operational concepts and control systems [11].

Analysis of published shows that to assess the required abilities and skills from cognitive, psychomotor, physical,

sensory and interactive/social domains needs to take into account the specifics of small UAV operator activity.

Thus, in the study the factors that affect the stable work and fatigue of the operator several UAV used the task of observing signals and counting symbols for 2-th hours. The subjective states were assessed with the Dundee Stress State Questionnaire [17]. This approach requires a lot of time to conduct the study, process and evaluate the results.

The psychomotor reaction time operator MQ-1 Predator was evaluated using the MicroCog test. The analyses in this study were generated using SAS software, version 9.3 (SAS Institute, Cary, NC), taking into account daily assessments and dividing into two groups based on threshold. To determine which aptitudes were predictive of high performers, stepwise binary logistic regressions were used to retain predictor effects titrating from $p < 0.05$ to $p < 0.30$, by incremental units of 0.05 [16].

When evaluating psychomotor reactions for operators of military UAVs, the specifics of activities (UAV operators and sensor/payload operators), concepts of operation and control system of UAVs, in particular Heron 1 in the German Air Force and LUNA, KZO and Aladin in the German Army [11].

Consequently, it is possible to take into account the specifics a small UAV functioning when evaluating the operator's sensorimotor reaction time, by selecting and justifying the appropriate criteria. It is important that the criterion can be applied to estimate reactions, which are measured on the same complex, which was used to obtain output data for its justification. This will allow for static errors that may differ between complexes (stimulus output time, inertia of hardware levers, duration of response arrival, and the like).

3 MATERIALS AND METHODS

The most common approach to substantiating the criteria and standards for evaluating sensorimotor reactions is the accumulation for statistical material and its mathematical mathematical processing [16, 11, 18–20]. To get the necessary data, we developed a program complex in LabVIEW, which allows estimating sensorimotor reaction time in the modes: skill development, under obstacle conditions, and under conditions of skill rearrangement. The program provides formation of visual stimuli and displaying them on the PC monitor depending on the testing mode, reading and processing of hardware data, processing of measurement results and archiving of obtained data. The appearance of the program working window is shown on Fig. 2.

In order to substantiate the criterion for evaluating the sensorimotor reaction time by a small UAV operator, a control group of 67 students was selected. It should be noted that the students had no skills of UAV piloting. As a result of testing the selected persons with the help of the developed complex, statistical data was obtained

concerning the average time of sensorimotor reactions of each person in conditions a skill, obstacles, rearrangement skill. The sensorimotor reaction time is a random continuous value that can take a value from a certain interval. In order to establish a norm against which to compare the values obtained and decide on the suitability of candidates for study, let us estimate the numerical characteristics of the data obtained in each of the three modes.

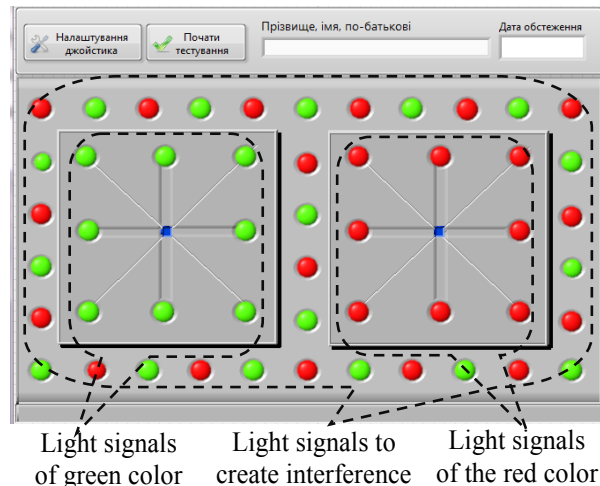


Figure 2 – The external appearance a working window

To this end, let us construct statistical series for the three evaluation modes (Table 1). To do this, divide the entire range of obtained values into intervals and count the number X_i , alues that fell into each i -th interval [21].

On the basis these statistical series, histograms for the distribution of the duration of the average reaction time (Fig. 3).

Table 1 – Statistical series of obtained values

t, s	$X(t)$ in the conditions of the skill development, the person	$X(t)$ in the conditions of disadvantages, the person	$X(t)$ in the context of skill restructuring, the person
0.8–0.9	2	1	5
0.9–1.0	1	1	5
1.0–1.1	11	4	7
1.1–1.2	17	9	7
1.2–1.3	10	13	8
1.3–1.4	6	11	10
1.4–1.5	9	8	4
1.5–1.6	3	4	5
1.6–1.7	4	7	4
1.7–1.8	1	7	6
1.8–1.9	1	1	2
1.9–2.0	0	1	1
2.0–2.1	0	0	1
2.1–2.2	1	0	1
2.2–2.3	1	0	1

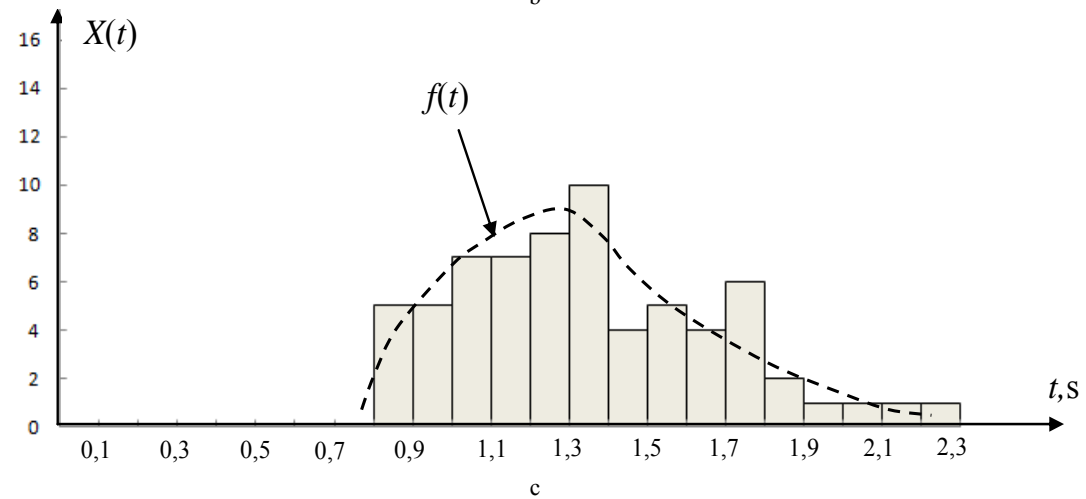
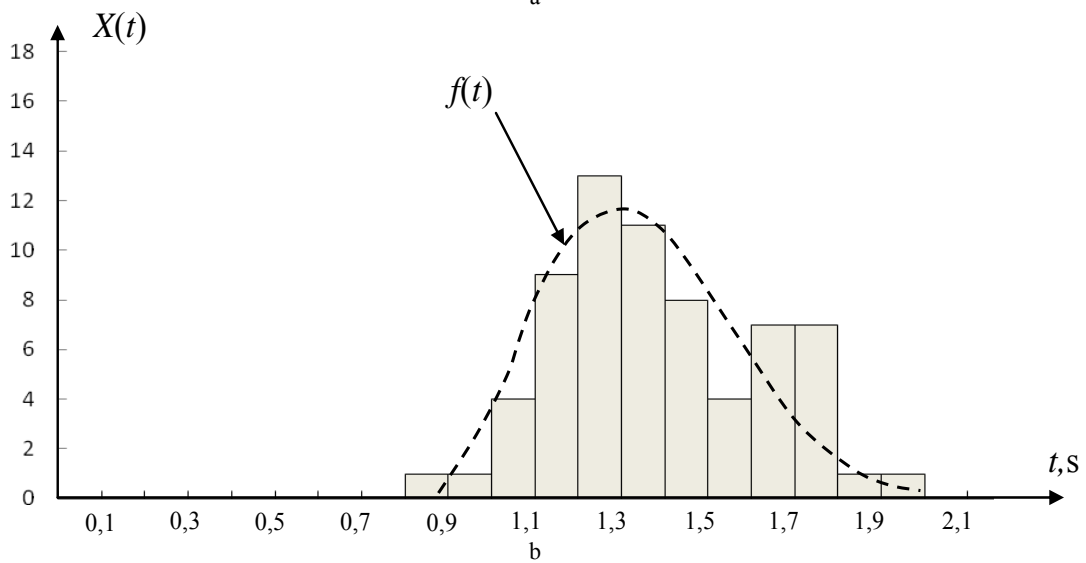
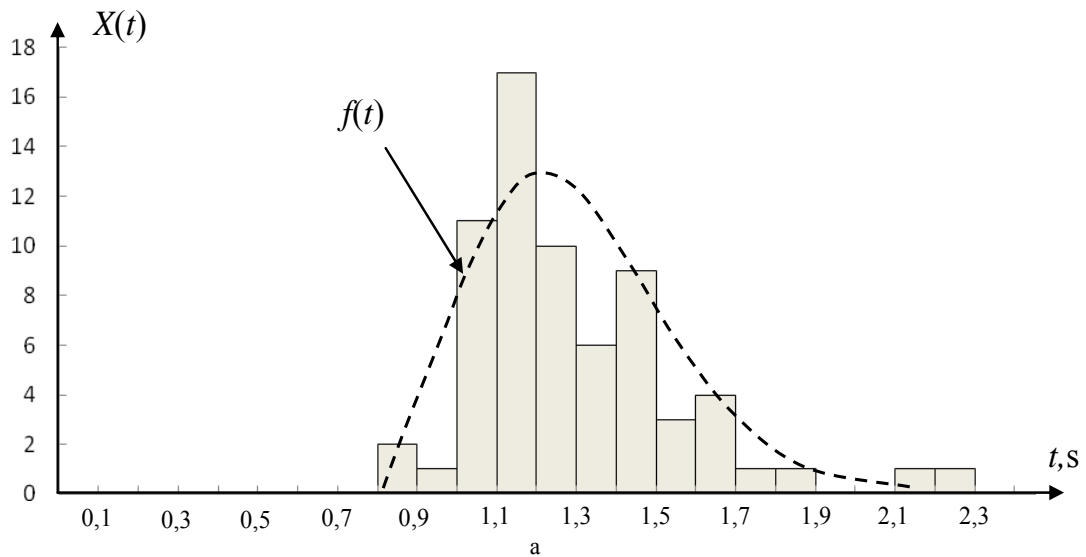


Figure 3 – Histograms showing the distribution of average reaction times:
 a – in conditions of skill development; b – in conditions of obstacles; c – in conditions of skill reconstruction

In the given statistical distributions there is an error of representativeness. That is why during processing this statistical material there was a decoupled task on statistical series alignment, which consisted in selection of theoretical distribution curve for each series, which will display only essential features of statistical material and not randomness, connected with insufficient volume of experimental data [21]. For this purpose, the histogram of the distribution was approximated by a fourth-degree polynomial in Microsoft Office Excel. As a result, we obtained theoretical densities of the distribution of the average reaction time – $f(t)$ (see Fig. 3).

A priori, most researchers assume that the distribution of sensorimotor reaction times is close to normal. However, the obtained plot for the density distribution of the mean reaction time has an asymmetry, which agrees with numerous experimental works, in which it is shown that the distribution of reaction times has a significant positive asymmetry in most cases [18, 20, 22]. This phenomenon has a convincing explanation: reaction times are characterized by the presence a rigid lower limit, determined by the physiological capabilities of the person and the complex on which the study is conducted, while the upper limit varies over a wide range and is usually associated with the features tested [18].

Determination of the distribution law with the density values that are presented in Fig. 3 and the calculation its parameters requires additional research. Therefore, taking into account the need to determine the standard for the upper limit of measured values only, in our further calculations we restrict ourselves to the data that describe the right branch of the distribution density relative to its mode, to the data describing the right branch of distribution density relative to its mode. With increasing of observations, this branch will approach corresponding branch of normal law distribution density [21] with mathematical expectation m_t , which is equal to the modality of the distribution. In this case formula (1) it will look like

$$\int_{m_t}^{T_N} f_m(t) dt = 0.48.$$

Probability of a random variable distributed according to the normal law in the interval m_t+T_N will have a value of 0,48 if

$$T_N = m_t + 2\sigma. \quad (2)$$

To obtain an estimate of the standard deviation we use the expression [21]

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (X_i - m_t)^2}{n-1}}. \quad (3)$$

Using the level statistical data, which are described by the theoretical graph of the density distribution of the average reaction time, we performed calculations for expressions (2), (3) and obtained standards for the upper limit of measured values of the duration of sensorimotor reactions by small UAV operators in conditions: skill development, obstacles, skill restructuring (Table 2).

Table 2 – Results of the calculations

indicator	Values of the indicators		
	in the development of a skill, s	in the face of obstacles, s	in the context of skill restructuring, s
m_t	1.22	1.33	1.30
$\hat{\sigma}$	0.33	0.29	0.4
T_N	1.88	1.91	2.10

4 EXPERIMENTS

Based on the results obtained, the sensorimotor reaction time for candidates to study the appropriate training program and to master the specialty UAV operator in the PPS was evaluated. It should be noted that candidates were pre-selected for the level of general training, motivation to study, experience in UAV piloting. The total number of persons in the training groups was 208 persons.

During the evaluation, the developed software package was used (see Fig. 1). As the hardware that provides specific movements in response to stimuli was used a UAV control panel, which is connected to the PC through the USB interface. The task of the candidates was to respond as quickly as possible with movements using the control levers to the flashes of light signals that are located in the program window. The light signals light up until the right movement was made.

In the right square there are 8-th red light signals that need movements of the right arm control panel. In the left square there are eight green light signals, which need to perform movements with the left manipulator of the control panel.

When evaluating the sensorimotor reactions by small UAV operators under conditions a skill is developed, the light signals that are located in the squares light up. The light signals can be extinguished by making a movement, with the appropriate remote control arm, in the direction the light signals that lit up.

When evaluating the sensorimotor reactions by small UAV operators under obstacle conditions, together with the light signals that are located in the squares, the light signals located on the outer perimeter, which are designed to create obstacles (distracting visual signals), are lit. It is possible to extinguish the light signals by making a movement, with the appropriate remote control arm, in the direction of the light bulb, which is lit in the squares.

The peculiarity in evaluating sensorimotor reactions of small UAV operators in conditions of skill rearrangement is the need to make mirror movements by manipulators of the remote control in response to the ignition light signals. That is, when red light signals, which are located in the right square, ignite, it is necessary to respond with movements of the left manipulator, when green light signals, which are located in the left square, ignite, it is necessary to respond with movements of the right manipulator.

In each test mode 50-th flashes are performed. Hardware calibration is performed before testing begins. Evaluation results are recorded in a text file that can be reviewed in a Microsoft Office Excel or Word environment.

5 RESULTS

According to the results of the experiment, the distribution of candidates by time of testing is plotted, which is shown in Fig. 4.

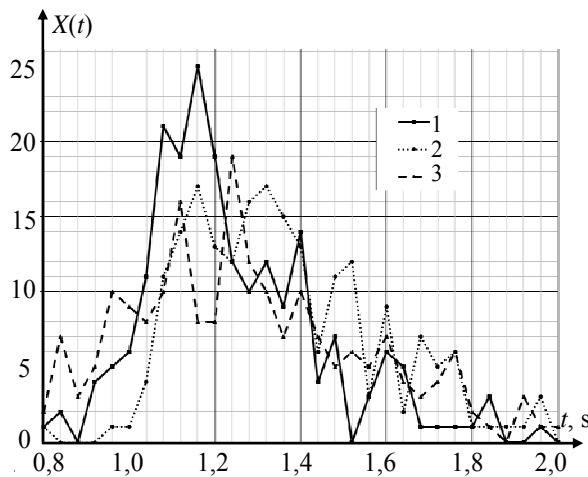


Figure 4 – Distribution of candidates by time of testing: 1 – in conditions of skill development; 2 – in conditions of obstacles; 3 – in conditions of skill restructuring

In order to obtain an integral assessment for candidates' suitability for study, the results obtained from all tests, including sensorimotor reaction time, were reduced to a single 9-th point scale. For this purpose, the coefficient was calculated by the expression

$$K = 10 \frac{T_N}{T_m}$$

The transition to the 9-th grade scale was carried out according to Table 3.

Table 3 – Grading standards with a 9-point scale

score	1	2	3	4	5	6	7	8	9
K	≤7	7	8	9	10	11	12	13	≥13

The final score was derived as the arithmetic mean value of the scores obtained in the three test modes. If a fractional number was obtained, it was rounded to a whole number. Candidates who scored less than 5-th points in total were considered unfit for further study, there were 4 such out of 208 examinees. According to the results of other tests (spatial operation of geometrical grudges; memory for visual grudges; long-term concentration of attention under difficult conditions, responsibility and neatness) these persons turned out to be fit for study. According to the results from the study, they successfully passed the examinations. According to the instructors' evaluations, the marked persons had a slightly harder time mastering practical skills from piloting.

6 DISCUSSION

The results show that those who were pre-selected for the level a general training, motivation to study got better indicators of sensorimotor reaction time, in comparison with the control group. At the same time, we can conclude that the sensorimotor reaction time cannot be regarded as the only indicator of candidates' suitability for the training program for small UAV operators. It is reasonable to use reaction time estimation as an additional factor during PPS. The obtained distribution densities of the average reaction time duration need additional study.

It is necessary to investigate the causes of the nature of failures at certain hour intervals in the conditions of obstacles and skill restructuring. Based on the results of the accumulation of statistical material, it is advisable to refine the criteria for evaluating the sensorimotor reaction time of a small UAV operator.

In further research, it is advisable to increase the number of criteria for evaluating the operator's sensorimotor reaction time. This will make it possible to introduce additional levels of candidate suitability: suitable, suitable with restriction, conditionally suitable, unsuitable, and so on.

CONCLUSIONS

The methods of justification criteria for evaluating the sensorimotor reaction time of a small UAV operator have been further developed in this work. This allowed improving the quality of PPS of UAV operators.

For the first time, specialized software has been developed, the implementation of which allowed to increase the efficiency and simplify the procedure for evaluating study candidates. The proposed approach ensured that the peculiarities involved in the functioning the small UAV and the specifics of the operator's nimble movements were taken into account.

Improved the process of integral assessment the candidates' suitability for training according to the results obtained by passing many tests, by reducing the obtained estimates the sensorimotor reaction time of the small UAV operator to a single scale.

ACKNOWLEDGEMENTS

The work is supported scientific research project of Zhytomyr military institute named after S. P. Korolov "Development of a hardware and software system for evaluating sensorimotor reactions of unmanned aerial vehicle operators" (state registration number 0118U000023).

REFERENCES

- Williams K.W. Human Factors Implications of Unmanned Aircraft Accidents: Flight-Control Problems, *Human Factors of Remotely Operated Vehicles (Advances in Human Performance and Cognitive Engineering Research)*, 2006, Vol. 7, pp. 105–116. DOI: 10.1016/S1479-3601(05)07008-6
- Wiegmann D., Shappell S. A human error approach to aviation accident analysis: the Human Factors Analysis and Classification System. Burlington, VT, Ashgate, 2003, 165 p. DOI: 10.4324/9781315263878-3
- Goodrich M. A., Cummings M. L. Human Factors Perspective on Next Generation Unmanned Aerial Systems, *Handbook of Unmanned Aerial Vehicles*, Springer-Verlag, 2015, pp. 2405–2423. DOI: 10.1007/978-90-481-9707-1_23
- Oncu M., Yildiz S. An analysis of human causal factors in unmanned aerial vehicle (UAV) accidents. alifornia, Monterey, December 2014, 91 p. DOI: 10.21236/ada620843
- Qi S., Wang F., Jing L. Aircraft System Pilot/Operator Qualification Requirements and Training Study, *MATEC Web Conference: Second International Conference on Mechanical, Material and Aerospace Engineering, 2MAE*, 2018, Vol. 179, 03006. DOI: 10.1051/mateconf/201817903006
- Havlikova M., Jirglb M., Bradac Z. Human Reliability in Man-Machine Systems, *Procedia Engineering*, 2015, Vol. 100, pp. 1207–1214. DOI: 10.1016/j.proeng.2015.01.485
- Chung W. K. Reliability evaluation of a human operator under various levels of stress, *Microelectronics Reliability*, 1991, Vol. 31, № 6, pp. 1251–1255. DOI: 10.1016/0026-2714(91)90315-X
- Petrov A., Volodina K., Belyaevam T. The role of the psychophysiological characteristics of a person in his professional development, *Education and Self Development*, 2019. Vol. 14, No. 4, pp. 63–71. DOI: 10.26907/esd14.4.06
- Hryshchenko Y. Reliability problem of ergatic control systems in aviation, *Methods and Systems of Navigation and Motion Control (MSNMC)*, Fourth International Conference. Kiev, 18–20 October 2016, pp. 126–129. DOI: 10.1109/msnmc.2016.7783123
- Lancaster R., Baseman E., Smolinski L. A quality improvement project: Defining and operationalizing a holistic admission selection policy, *Journal of Professional Nursing*, 2020, Vol. 36, No. 4, pp. 259–263. DOI: 10.1016/j.profnurs.2020.01.003
- Melcher W., Neumann M., Eißfeldt H., Schwab A. Cognitive and psychomotor requirements for operators of military UAS, *61st Conference of the International Military Testing Association in Tallinn*, 7–11 October 2019, 11 p.
- Ermakov A. Simple and Complex Sensomotor Reaction for Choice when Teaching Protection Against Armed Attacker, *First International Volga Region Conference on Economics, Humanities and Sports (FICEHS 2019)*, 18 January 2020, pp. 772–774. DOI: 10.2991/aebmr.k.200114.185
- Tadema J., Theunissen E. A concept for UAV operator involvement in airborne conflict detection and resolution, *27th Digital Avionics Systems Conference, St. Paul, MN, USA, 26–30 October 2008*, pp. 4.C.1–1–4.C.1–12. DOI: 10.1109/DASC.2008.4702829
- Chappelle W., Tran N., Thompson W., Goodman T., Hyde K. Intelligence and neuropsychological aptitude testing of U.S. Air Force MQ-1 Predator pilot training candidates, *Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine*, 2012. Technical Report AFRL-SA-WP-TR-2013-0003. DOI: 10.21236/ada577826
- Chappelle W., Heerema B., Thompson W. Factor analysis of computer-based Multidimensional Aptitude Battery-Second Edition intelligence testing from rated U. S. Air Force pilots. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine, 2012, Technical Report AFRL-SA-WP-TR-2013-0005. DOI: 10.21236/ada583710
- Chappelle W., Goodman T., Swearingen J., Thompson W. A Preliminary Investigation into Cognitive Aptitudes Predictive of Overall MQ-1 Predator Pilot Qualification Training Performance, *Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine*, 2015. Technical Report AFRL-SA-WP-SR-2015-0025, 14 p.
- Ryan W., Matthews G., Lin J., Szalma J., Calhoun G., Funke G., Chiu C-Y., Ruff H. Vigilance and Automation Dependence in Operation of Multiple Unmanned Aerial Systems (UAS): A Simulation Study, *Hum Factors*, May 2019, Vol. 61, No. 3, pp. 488–505. DOI: 10.1177/0018720818799468
- Kukushkin Ju., Ajvazjan S. Metodika avtomatizirovannoj obrabotki upravljajushhij dvizhenij operatora v prikladnyh issledovaniyah nadezhnosti jergaticeskijh sistem, *Kibernetika i programirovanie*, 2018, No. 5, pp. 15–23. DOI: 10.25136/2306-4196.2018.5.1817
- Gusev D., Klimov R. Programmnyj kompleks apriornogo ocenivaniya pokazatelej kachestva professional'noj dejatel'nosti operatora jergaticeskoy sistemy, *Programmnye sistemy i vychislitel'nye metody*, 2015, No. 4, pp. 374–389. DOI: 10.7256/2305-6061.2015.4.17965
- Ignatova Y., Makarova I., Yakovleva K., Aksenova A. Visual-motor reactions as an indicator of CNS functional state, *Ulyanovsk Medico-biological Journal*, 2019, No. 3, pp. 38–51. DOI: 10.34014/2227-1848-2019-3-38-51
- Werner L. Probability Theory. De Gruyter Textbook, 2016, 395 p. DOI: 10.1515/9783110466195
- Kulakov A. Osobennosti prostojs psihofiziologicheskoy reakcii, *Fiziologija cheloveka*, 2018, Tom 44, No. 4, pp. 60–66. DOI: 10.1134/s0131164618040069

Received 28.01.2021.
Accepted 15.05.2021.

УДК 004.67: 37.042.2

КРИТЕРІЇ ОЦІНЮВАННЯ ЧАСУ СЕНСОМОТОРНОЇ РЕАКЦІЇ ОПЕРАТОРА МАЛИХ БПЛА

Вакалюк Т. А. – доктор педагогічних наук, професор, професор кафедри інженерії програмного забезпечення, Державний університет «Житомирська політехніка», Житомир, Україна.

Пількевич І. А. – доктор технічних наук, професор, професор кафедри комп'ютерних інформаційних технологій, Житомирський військовий інститут імені С. П. Корольова, Житомир, Україна.

Токар А. М. – кандидат технічних наук, начальник науково-дослідного відділу, Житомирський військовий інститут імені С. П. Корольова, Житомир, Україна.

Лобода Р. І. – науковий співробітник науково-дослідного відділу, Житомирський військовий інститут імені С. П. Корольова, Житомир, Україна.

АНОТАЦІЯ

Актуальність. Стрімкий розвиток науки та техніки зумовлює значне розширення сфер застосування БпЛА різного призначення. Запорукою ефективного використання БпЛА є якісна підготовка операторів, важливим елементом якої є професійний психологічний відбір (ППВ) кандидатів, зокрема, оцінювання їх сенсомоторних реакцій. Це можна забезпечити шляхом вибору та обґрунтування відповідних критеріїв.

Мета. Метою роботи є обґрунтування критеріїв оцінювання часу сенсомоторних реакцій оператора малих БпЛА шляхом аналізу щільності розподілу статистичних даних.

Метод. Запропоновано метод, що дозволив визначити критерії оцінювання часу сенсомоторних реакцій оператора малих БпЛА на основі накопичення статистичного матеріалу та його математичного оброблення за результатами проведення натурального експерименту. Метод дозволяє оцінити числові характеристики розподілу тривалості середнього часу реакції у трьох режимах: вироблення навички, в умовах перешкод, в умовах перебудови навички та отримати узагальнену оцінку. Вдалося, шляхом аналізу випадкових неперервних величин, що приймають значення з деякого проміжку значень, встановити нормативи, з якими порівнюються отримані значення часу сенсомоторної реакції оператора малих БпЛА та приймається рішення щодо їх придатності до навчання.

Результати. У роботі отримано статистичні ряди для режимів оцінювання: вироблення навички, в умовах перешкод, в умовах перебудови навички. Для наочного представлення рядів побудовано відповідні гістограми розподілу тривалості середнього часу реакції. З метою усунення помилки репрезентативності здійснено вирівнювання статистичних рядів шляхом підбору для кожного ряду теоретичної кривої розподілу, що відображає лише суттєві риси статистичного матеріалу. Для цього здійснено апроксимація гістограми розподілу поліномом четвертого ступеня. Встановлено інтервал теоретичної щільності розподілу, попадання в який часу сенсомоторної реакції довільної особи вважається нормою, при заданій ймовірності достовірності такої події – 0,95. Для перевірки дієвості запропонованого методу синтезовано алгоритми оцінювання часу сенсомоторної реакції оператора малих БпЛА у трьох режимах та розроблено відповідне програмне забезпечення, що реалізує запропоновані алгоритми.

Висновки. Обґрунтовано критерії оцінювання часу сенсомоторної реакції оператора БпЛА на зоровий подразник з використанням спеціалізованого програмного забезпечення. Це дозволило проводити попередній ППВ кандидатів на навчання з урахуванням вимог до моторики оператора малих БпЛА і специфіки його рухів. Проведені експерименти підтвердили обґрунтованість прийнятих рішень. Перспективи подальших досліджень можуть включати розширення режимів тестування з обґрунтуванням відповідних критеріїв оцінювання.

КЛЮЧОВІ СЛОВА: критерії оцінювання, оператор малих БпЛА, час сенсомоторної реакції, щільність розподілу.

УДК 004.67: 37.042.2

КРИТЕРИИ ОЦЕНКИ ВРЕМЕНИ СЕНСОМОТОРНОЙ РЕАКЦИИ ОПЕРАТОРА МАЛЫХ БПЛА

Вакалюк Т. А. – доктор педагогических наук, профессор, профессор кафедры инженерии программного обеспечения, Государственный университет «Житомирская политехника», Житомир, Украина.

Пількевич І. А. – доктор технических наук, профессор, профессор кафедры компьютерных информационных технологий, Житомирский военный институт имени С. П. Королева, Житомир, Украина.

Токар А. Н. – кандидат технических наук, начальник научно-исследовательского отдела, Житомирский военный институт имени С. П. Королева, Житомир, Украина.

Лобода Р. И. – научный сотрудник научно-исследовательского отдела, Житомирский военный институт имени С. П. Королева, Житомир, Украина.

АНОТАЦІЯ

Актуальность. Стремительное развитие науки и техники приводит к значительному расширению сфер применения БпЛА различного назначения. Залогом эффективного использования БпЛА является качественная подготовка операторов, важным элементом которой является профессиональный психологический отбор (ППО) кандидатов, в частности, оценка их сенсомоторных реакций. Это можно обеспечить путем выбора и обоснования соответствующих критериев.

Цель. Целью работы является обоснование критериев оценки времени сенсомоторных реакций оператора малых БпЛА путем анализа плотности распределения статистических данных.

Метод. Предложен метод, который позволил определить критерии оценки времени сенсомоторных реакций оператора малых БпЛА на основе накопления статистического материала и его математической обработки по результатам проведения натурального эксперимента. Метод позволяет оценить числовые характеристики распределения продолжительности среднего времени реакции в трех режимах: выработки навыка, в условиях помех, в условиях перестройки навыка и получить обобщенную оценку. Удалось, путем анализа случайных непрерывных величин, принимающих значения из некоторого промежутка значений, установить нормативы, с которыми сравниваются полученные значения времени сенсомоторной реакции оператора малых БпЛА и принимается решение об их пригодности к обучению.

Результаты. В работе получены статистические ряды для режимов оценки: выработки навыка, в условиях помех, в условиях перестройки навыка. Для наглядного представления рядов построены соответствующие гистограммы распределения продолжительности среднего времени реакции. С целью устранения ошибки репрезентативности осуществлено выравнивание статистических рядов путем подбора для каждого ряда теоретической кривой распределения, которая отражает только существенные черты статистического материала. Для этого осуществлена аппроксимация гистограммы распределения полиномом четвертой степени. Установлено интервал теоретической плотности распределения, попадание в который времени сенсомоторной реакции произвольного лица считается нормой, при заданной вероятности

© Vakaliuk T. A., Pilkevych I. A., Tokar A. M., Loboda R. I., 2021
DOI 10.15588/1607-3274-2021-2-19

достоверности такого события – 0,95. Для проверки действенности предложенного метода синтезированы алгоритмы оценки времени сенсомоторной реакции оператора малых БПЛА в трех режимах и разработано соответствующее программное обеспечение, реализующее предложенные алгоритмы.

Выводы. Обоснованы критерии оценки времени сенсомоторной реакции оператора БПЛА на зрительный раздражитель с использованием специализированного программного обеспечения. Это позволило проводить предварительный ППО кандидатов на обучение с учетом требований к моторике оператора малых БПЛА и специфике его движений. Проведенные эксперименты подтвердили обоснованность принимаемых решений. Перспективы дальнейших исследований могут включать расширение режимов тестирования с обоснованием соответствующих критериев оценки.

КЛЮЧЕВЫЕ СЛОВА: критерии оценки, оператор малых БПЛА, время сенсомоторной реакции, плотность распределения.

ЛІТЕРАТУРА / LITERATURA

1. Williams K.W. Human Factors Implications of Unmanned Aircraft Accidents: Flight-Control Problems / K. W. Williams // Human Factors of Remotely Operated Vehicles (Advances in Human Performance and Cognitive Engineering Research). – 2006. – Vol. 7. – P. 105–116. DOI: 10.1016/S1479-3601(05)07008-6
2. Wiegmann D. A human error approach to aviation accident analysis: the Human Factors Analysis and Classification System / D. Wiegmann, S. Shappell. – Burlington, VT: Ashgate, 2003. – 165 p. DOI: 10.4324/9781315263878-3
3. Goodrich M. A. Human Factors Perspective on Next Generation Unmanned Aerial Systems / M. A. Goodrich, M. L. Cummings // Handbook of Unmanned Aerial Vehicles, Springer-Verlag. – 2015. – P. 2405–2423. DOI: 10.1007/978-90-481-9707-1_23
4. Oncu M. An analysis of human causal factors in unmanned aerial vehicle (UAV) accidents / M. Oncu, S. Yildiz. – California, Monterey, December 2014. – 91 p. DOI: 10.21236/ada620843
5. Qi S. Aircraft System Pilot/Operator Qualification Requirements and Training Study / S. Qi, F. Wang, L. Jing // MATEC Web Conference: Second International Conference on Mechanical, Material and Aerospace Engineering, 2MAE. – 2018. – Vol. 179. – P. 03006. DOI: 10.1051/mateconf/201817903006
6. Havlikova M. Human Reliability in Man-Machine Systems / M. Havlikovaa, M. Jirglb, Z. Bradac // Procedia Engineering. – 2015. – Vol. 100. – P. 1207–1214. DOI: 10.1016/j.proeng.2015.01.485
7. Chung W. K. Reliability evaluation of a human operator under various levels of stress / W. K. Chung // Microelectronics Reliability. – 1991. – Vol. 31, № 6. – P. 1251–1255. DOI: 10.1016/0026-2714(91)90315-X
8. Petrov A. The role of the psycho-physiological characteristics of a person in his professional development / A. Petrov, K. Volodina, T. Belyaevam // Education and Self Development. – 2019. – Vol. 14, № 4. – P. 63–71. DOI: 10.26907/esd14.4.06
9. Hryshchenko Y. Reliability problem of ergatic control systems in aviation / Y. Hryshchenko // Methods and Systems of Navigation and Motion Control (MSNMC) : Fourth International Conference, Kiev, 18–20 October 2016. – P. 126–129. DOI: 10.1109/msnmc.2016.7783123
10. Lancaster R. A quality improvement project: Defining and operationalizing a holistic admission selection policy / R.Lancaster E.Baseman L.Smolinski // Journal of Professional Nursing. – 2020. – Vol. 36, № 4. – P. 259–263. DOI: 10.1016/j.profnurs.2020.01.003
11. Cognitive and psychomotor requirements for operators of military UAS / [W. Melcher, M. Neumann, H. Eißfeldt, A. Schwab]. – 61st Conference of the International Military Testing Association in Tallinn, 7–11 October 2019. – 11 p.
12. Ermakov A. Simple and Complex Sensomotor Reaction for Choice when Teaching Protection Against Armed Attacker / A. Ermakov // First International Volga Region Conference on Economics, Humanities and Sports (FICEHS 2019), 18 January 2020. – P. 772–774. DOI: 10.2991/aebmr.k.200114.185
13. Tadema J. A concept for UAV operator involvement in airborne conflict detection and resolution / J. Tadema, E. Theunissen // 27th Digital Avionics Systems Conference, St. Paul, MN, USA, 26–30 October 2008. – P. 4.C.1–1–4.C.1–12. DOI: 10.1109/DASC.2008.4702829
14. Chappelle W. Intelligence and neuropsychological aptitude testing of U.S. Air Force MQ-1 Predator pilot training candidates. / [W. Chappelle, N. Tran, W. Thompson et al] // Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine. – 2012. Technical Report AFRL-SA-WP-TR-2013-0003. DOI: 10.21236/ada577826
15. Chappelle W. Factor analysis of computer-based Multidimensional Aptitude Battery-Second Edition intelligence testing from rated U.S. Air Force pilots / W. Chappelle, B. Heerema, W. Thompson. – Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine. – 2012. Technical Report AFRL-SA-WP-TR-2013-0005. DOI: 10.21236/ada583710
16. A Preliminary Investigation into Cognitive Aptitudes Predictive of Overall MQ-1 Predator Pilot Qualification Training Performance / [W. Chappelle, T. Goodman, J. Swearingen, W. Thompson]. – Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine. – 2015. – Technical Report AFRL-SA-WP-SR-2015-0025. – 14 p.
17. Vigilance and Automation Dependence in Operation of Multiple Unmanned Aerial Systems (UAS): A Simulation Study / [W. Ryan, G. Matthews, J. Lin et al] // Hum Factors. – May 2019. – Vol. 61, № 3. – P. 488–505. DOI: 10.1177/0018720818799468
18. Кукушкин Ю. Методика автоматизированной обработки управляющих движений оператора в прикладных исследованиях надежности эргатических систем / Ю. Кукушкин, С. Айвазян // Кибернетика и программирование. – 2018. – № 5. – С. 15–23. DOI: 10.25136/2306-4196.2018.5.1817
19. Гусев Д. Программный комплекс априорного оценивания показателей качества профессиональной деятельности оператора эргатической системы / Д. Гусев, Р. Климов // Программные системы и вычислительные методы. – 2015. – № 4. – С. 374–389. DOI: 10.7256/2305-6061.2015.4.17965
20. Visual-motor reactions as an indicator of CNS functional state / [Y. Ignatova, I. Makarova, K. Yakovleva, A. Aksenova] // Ulyanovsk Medico-biological Journal. – 2019. – № 3. – P. 38–51. DOI: 10.34014/2227-1848-2019-3-38-51
21. Werner L. Probability Theory / L. Werner. – De Gruyter Textbook, 2016. – 395 p. DOI: 10.1515/9783110466195
22. Кулаков А. Особенности простой психофизиологической реакции / А. Кулаков // Физиология человека. – 2018. – Том 44, № 4. – С. 60–66. DOI: 10.1134/s0131164618040069

Наукове видання

**Радіоелектроніка,
інформатика,
управління**

№ 2/2021

Науковий журнал

Головний редактор – д-р техн. наук С. О. Субботін

Заст. головного редактора – д-р техн. наук Д. М. Піза

Комп'ютерне моделювання та верстання
Редактор англійських текстів

С. В. Зуб
С. О. Субботін

Оригінал-макет підготовлено у редакційно-видавничому відділі НУ «Запорізька політехніка»

Свідоцтво про державну реєстрацію
КВ № 24220-14060 ПР від 19.11.2019.

*Підписано до друку 15.06.2021. Формат 60×84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 23,02.
Тираж 300 прим. Зам. № 462.*

69063, м. Запоріжжя, НУ «Запорізька політехніка», друкарня, вул. Жуковського, 64

Свідоцтво суб'єкта видавничої справи
ДК № 6952 від 22.10.2019.